Batanero, C., Valenzuela-Ruiz, S.M. y Gea, M.M. (2020). Significados institucionales y personales de los estadísticos de orden en la Educación Secundaria. *Matemáticas, Educación y Sociedad*, *3*(2), 21-39

ISSN: 2603-9982

SIGNIFICADOS INSTITUCIONALES Y PERSONALES DE LOS ESTADÍSTICOS DE ORDEN EN LA EDUCACIÓN SECUNDARIA

Carmen Batanero, Universidad de Granada, España

Silvia M. Valenzuela-Ruiz, Universidad de Granada, España

María Magdalena Gea, Universidad de Granada, España

Resumen

El objetivo de este trabajo es caracterizar el significado institucional de los estadísticos de orden en la Educación Secundaria (ESO y Bachillerato), identificando los objetos matemáticos elementales involucrados. Para ello presentamos un análisis semiótico de los estadísticos de orden, utilizando el Enfoque Ontosemiótico del Conocimiento y la Instrucción matemáticos (EOS), que muestra su complejidad, incluso limitándonos a la estadística descriptiva. También resumimos las dificultades descritas, respecto a cada uno de los tipos de objetos primarios identificados en el EOS, con el fin de caracterizar los posibles significados personales que los estudiantes pueden asignar a los mismos. Esta información puede ser útil para identificar estas dificultades en los estudiantes y ayudarles a superarlas.

Palabras clave: *estadísticos de orden, significados institucionales, dificultades de los estudiantes.*

Institutional and personal meaning of order statistics

Abstract

The aim of this work is to characterize the institutional meaning of order statistics in Secondary Education (ESO and High School), by identifying the elementary mathematical objects required in the work with the same. To achieve this goal, we present a semiotic approach using Ontosemiotic Approach to Mathematical Knowledge and Instruction (OSA), which reveals the complexity of its meaning, even in descriptive analysis. We also summarize the main difficulties identified in previous research to characterize the personal meaning that students may assign to these statistics. This information is useful in order to identify these difficulties and help students to overcome them.

Keywords*: order statistics, institutional meanings, students' difficulties.*

Recibido: 28/07/2020; Aceptado: 17/09/2020

INTRODUCCIÓN

La enseñanza de la estadística se privilegia actualmente, debido a la necesidad de dotar al estudiante de herramientas que le permita interpretar los numerosos mensajes con información estadística que se presentan en los medos de comunicación (Ridgway, Nicholson, y McCusker, 2011). La necesidad de implicar activamente a los ciudadanos en las decisiones que afectan a campos como la salud, política o educación, requiere adquirir una *estadística cívica*, o conjunto de conocimientos estadísticos y actitudes hacia la estadística que permitan la participación democrática de los ciudadanos en la sociedad (Engel, 2019).

Uno de los temas que se estudia en diversos ciclos escolares son los estadísticos de orden, que nos indican la posición que un cierto valor ocupa dentro de un conjunto de datos ordenados, y por ello nos informan del porcentaje de datos que tienen un valor de la variable menor o igual que dicho estadístico (Hoaglin, Mosteller y Tukey, 1983). Una de las características más importantes de un dato en una lista ordenada es su distancia al principio o final de la lista; de ahí el interés de los estadísticos de orden (Velleman y Hoaglin, 1981).

Otras razones principales que justifican su estudio son, en primer lugar, el análisis exploratorio de datos, introducido por Tukey (1977), que se apoya en gran parte en los estadísticos de orden, en los que se incluyen la mediana, deciles, percentiles y sus rangos, porque son estadísticos robustos. Esta propiedad implica que su valor no se modifica sustancialmente cuando se producen ligeras variaciones en los datos y están poco afectados por los valores atípicos. En el análisis exploratorio de datos también se han introducido una amplia variedad de gráficos basados en los estadísticos de orden, entre ellos el gráfico de cajas y el gráfico cuantil-cuantil (Chambers, 2018).

La mediana, que es uno de estos estadísticos, también se considera como medida de tendencia central, y es la más apropiada para el estudio de datos que no siguen el modelo de la curva normal, cuando existen valores atípicos dentro del conjunto de datos o en variables ordinales (Betanzos y López, 2017).

Por otro lado, los estadísticos de orden son base de la inferencia estadística no paramétrica, que puede aplicarse en condiciones menos restrictivas que la paramétrica o en pequeñas muestras. Entre estos métodos encontramos el test de la mediana o el test de rangos de Wilcoxon, que permite contrastar si dos muestras independientes provienen de la misma población (equivalente al contraste “t” para muestras independientes en poblaciones no normales) o la prueba Kruskal-Wallis (alternativa al análisis de varianza de un factor paramétrico) (Verma, 2019). También existen medidas de asociación específicas para variables ordinales, como el coeficiente de correlación por rangos de Spearman, que se suele utilizar para comparar dos clasificaciones independientes de un mismo conjunto de objetos (Zar, 2005).

Además, existen un gran número de herramientas estadísticas que utilizan este concepto y están enfocadas a realizar análisis específicos. Por ejemplo, el método de Monte Carlo usado para aproximar expresiones matemáticas o la función de cuantiles en la teoría de valores extremos y, por ende, en el análisis de riesgos.

Las directrices curriculares en España refuerzan en los últimos años la enseñanza de la estadística en los diferentes niveles de educación obligatoria. En concreto, la normativa vigente recoge el concepto de mediana y otros estadísticos de orden en el decreto de enseñanzas mínimas de matemáticas tanto para el ciclo de la Educación Secundaria Obligatoria (ESO) como para el Bachillerato. La mediana y el intervalo mediano aparecen por primera vez en el currículo en los cursos primero y segundo de ESO (MECD, 2015). En tercer curso, tanto en las matemáticas orientadas a las enseñanzas aplicadas como en las orientadas a las enseñanzas académicas, se contemplan los estadísticos de orden, su cálculo, interpretación y propiedades, así como el diagrama de caja, que se debe utilizar para comparar dos distribuciones en cuarto curso. No hay una mención explícita en el currículo de Bachillerato a estos estadísticos, aunque en primer curso de las ramas de Ciencias Sociales y de Ciencias se incluye la estadística descriptiva bidimensional, donde los estudiantes tendrán ocasión de utilizarlos. Los estudiantes también trabajan implícitamente con los estadísticos de orden cuando estudian, calculan o representan gráficamente las frecuencias acumuladas. En inferencia estadística, estudiada en segundo curso del Bachillerato de Ciencias Sociales, el concepto de percentil se aplica, tanto en la lectura de las tablas de la distribución normal, como en el uso de dicha distribución para el cálculo de intervalos de confianza. Todos estos temas, así como la estadística no paramétrica, se estudian con frecuencia en los cursos universitarios de estadística, tanto de pregrado como de grado.

Aunque esta amplia presencia curricular llevaría a pensar que los estadísticos de orden son conceptos bien comprendidos, algunas investigaciones sugieren la existencia de dificultades frecuentes en estudiantes de Educación Secundaria Obligatoria (Cobo, 2003), Bachillerato (Mayén, 2009; Mayén, Cobo, Batanero y Balderas, 2007) e incluso futuros profesores (Gea, Batanero, Fernández y Arteaga, 2016).

El objetivo de este trabajo es caracterizar el significado institucional de los estadísticos de orden en la Educación Secundaria (ESO y Bachillerato), identificando los objetos matemáticos elementales involucrados en el trabajo con los mismos. Además, presentamos un resumen de las principales investigaciones relacionadas con la comprensión de estos estadísticos, que indican que el significado personal construido sobre los mismos no siempre corresponde al significado institucional pretendido en la enseñanza.

FUNDAMENTOS

Nuestra investigación parte del Enfoque Ontosemiótico del Conocimiento y la Instrucción matemáticos (EOS) (Godino, 2002; Godino, Batanero y Font, 2007; 2019), que considera las situaciones-problemas como origen de la actividad matemática. En particular, utilizamos la noción de significado de un objeto matemático que se propone dentro de este enfoque, como el sistema de prácticas que realiza una persona, o se lleva a cabo en el seno de una institución, para resolver las situaciones-problemas de donde surge dicho objeto matemático. Los autores diferencian entre prácticas operativas, que se realizan al operar con objetos matemáticos, y discursivas, que incluyen las justificaciones y argumentos que apoyan la resolución de problemas o su justificación; en nuestro caso, nos interesamos por las situaciones-problemas que dan origen a los estadísticos de orden.

En el EOS se entiende por institución a un conjunto de personas que se interesan por resolver una misma clase de situaciones-problemas y donde se aplican herramientas compartidas. En nuestro caso, serían instituciones los matemáticos que aplican métodos basados en estadísticos de orden en su trabajo o las instituciones de enseñanza donde se estudian estos estadísticos. Puesto que el estudio de los estadísticos de orden puede llevarse a cabo con diferentes grados de complejidad y en las diferentes ramas de la estadística, en este trabajo nos limitamos a su estudio dentro de la estadística descriptiva al nivel de la Educación Secundaria.

Al considerar en el EOS una dimensión personal (subjetiva, mental) y otra institucional (objetiva, epistémica) del significado de un objeto matemático (Godino y Batanero, 1998), en nuestro estudio, el significado institucional del tema será el fijado en los currículos para la Educación Secundaria, mientras que el significado personal sería el conjunto de prácticas que los estudiantes realizan cuando se enfrentan a los estadísticos de orden.

Asumimos también del EOS una clasificación de objetos primarios que se ponen en juego dentro la actividad matemática y que se consideran como los elementos de significado (Godino, Batanero y Font, 2007). Dichos objetos matemáticos primarios, a su vez, se organizan en entidades más complejas (sistemas conceptuales o teorías). Por ejemplo, la mediana es un percentil, y los percentiles forman parte de los estadísticos de orden y estos de los resúmenes estadísticos de una distribución. Los objetos elementales considerados en el EOS son los siguientes:

- *Campos de problemas:* son las situaciones o aplicaciones que inducen actividad matemática de donde surge el objeto y pueden ser propias de la misma matemática o extra-matemáticas. En nuestra investigación, nos interesamos por la clase de problemas asociados a los estadísticos de orden.
- *Lenguaje:* los objetos matemáticos son inmateriales, y por tanto se necesitan representaciones materiales de los mismos utilizadas en la actividad matemática (términos, expresiones, símbolos, tablas, gráficos). Por ejemplo, para trabajar con la mediana usamos el término "mediana" o las notaciones Me, P_{50}, D_5, Q_2 o C_2.
- *Algoritmos y procedimientos:* acciones y operaciones que se emplean para resolver las situaciones o tareas. Por ejemplo, es característico en la solución de problemas relacionados con los estadísticos de orden hallar el valor de la variable que deja por debajo de él un porcentaje dado de datos (la población o de la muestra).
- *Definiciones y propiedades:* ideas matemáticas y sus relaciones (conceptos, proposiciones). En nuestro caso, existen distintas definiciones equivalentes de los estadísticos de orden que, junto con sus propiedades, se detallarán más adelante.
- *Argumentos:* finalmente, todas estas acciones y objetos se ligan entre sí mediante argumentos o razonamientos que se usan para resolver o comprobar las soluciones de los problemas o explicar a otro la solución y que pueden ser deductivas, inductivas, analógicas, o de otro tipo.

Utilizando esta clasificación, a continuación, caracterizamos el significado institucional de los estadísticos de orden y en la siguiente sección los significados personales que le asignan algunos estudiantes.

SIGNIFICADOS INSTITUCIONALES DE LOS ESTADÍSTICOS DE ORDEN

Como se ha indicado, nos restringiremos a la estadística descriptiva y análisis exploratorio de datos, aunque algunos de los objetos matemáticos identificados en relación con los estadísticos de orden pueden generalizarse para utilizarse en probabilidad, inferencia o análisis de asociación entre variables.

Problemas de los que surgen los estadísticos de orden

En primer lugar, analizamos los campos de problemas, cuya resolución requiere de los estadísticos de orden, entre los que se encuentra la mediana. Cobo (2003) describe tres problemas principales que se resuelven utilizando la mediana y que son los siguientes:

- *CP1. Determinar una medida de centralización cuando la media no es representativa*. Esta situación se da cuando la distribución de la variable en estudio tiene valores atípicos o varias modas y también cuando la distribución de la variable es muy asimétrica. En estos casos, es más representativo aquel valor de la variable que deja por encima y por debajo el mismo número de datos, en comparación con la media aritmética.
- *CP2. Hallar una medida de centralización para un conjunto de datos ordinales.* En las variables ordinales (por ejemplo, el número de orden de nacimiento de los niños en una familia), las diferencias numéricas de la variable no corresponden a diferencias proporcionales en alguna magnitud subyacente (en el ejemplo, la diferencia de edad de los niños). La media aritmética no tiene sentido en estas variables y, por tanto, la mediana sería una medida de centralización más adecuada.
- *CP3. Comparar dos o más conjuntos de datos ordinales.* Por la misma razón anterior, no deben compararse las medias de los conjuntos de datos, sino sus medianas.

Además, podemos encontrar los siguientes campos de problemas asociados a otros estadísticos de orden, cuando se analiza el contenido de los documentos curriculares (MECD, 2015) así como libros de texto de estadística descriptiva (Calot, 1988) o análisis exploratorio de datos (Tukey, 1977; Velleman, y Hoaglin, 1981).

- *CP4. Identificar el porcentaje de valores de la distribución con valor menor o mayor a uno dado.* De este campo de problemas surge la idea de percentil, decil o cuartil, de los que la mediana es un caso particular. Se asocia en general al control de calidad de un producto o de desarrollo de un sujeto; por ejemplo, se sitúa a un niño en el percentil de su altura y peso para comprobar su crecimiento adecuado.
- *CP5. Identificar el valor al que corresponde un porcentaje de casos menor o igual en una distribución.* Es el problema inverso al anterior y da lugar al concepto de rango de percentil. Uno de los usos más frecuentes en Bachillerato es la lectura de la tabla de la distribución normal estándar, que se aplica, tanto para la resolución de problemas de cálculo de probabilidades con la curva normal, como en el cálculo del intervalo de confianza, que se estudia en el segundo curso del Bachillerato de Ciencias Sociales.
- *CP6. Resumir gráficamente la distribución representando los porcentajes de casos centrales y extremos de la misma.* Este resumen se consigue mediante el gráfico de cajas en el que se representan cinco puntos de las distribuciones: mínimo, máximo, mediana y cuartiles, así como los valores atípicos. La determinación de estos estadísticos es, entonces, necesaria para la construcción del gráfico. Este campo de problemas se puede ampliar para comparar dos o más distribuciones.

- *CP7. Comparar gráficamente la diferencia de distribución empírica con otra teórica.* Esto se consigue con el gráfico cuantil-cuantil. Para ello se representa en un sistema de coordenadas cartesianas un gráfico de dispersión donde la primera coordenada es un cuantil de la distribución teórica modelo y la segunda el correspondiente cuantil de la distribución empírica.

Lenguaje asociado

En el EOS (Godino et al., 2007), la importancia del lenguaje es que se considera como mediador de las prácticas personales o institucionales, debido a su carácter representacional y operativo. Puesto que el objeto abstracto no es tangible ni visible, para trabajar con él es representado mediante palabras y notaciones. Igualmente, el lenguaje matemático sirve para representar las propiedades, así como para describir los problemas y sus datos. Siguiendo a Ortiz, Batanero y Serrano (2001) vamos a distinguir tres tipos de lenguaje:

- *Lenguaje verbal:* que incluye términos matemáticos o vocablos y expresiones del lenguaje natural. En el caso de la "mediana" además de con este nombre también podemos referirnos a ella mediante el percentil 50, quinto decil o segundo cuartil. Todos estos términos son sinónimos y el alumno debe saber emplearlos indistintamente. Igualmente ocurre con otros estadísticos, por ejemplo, hablamos de primer cuartil o percentil del 25%. Además, se utilizan expresiones como "estadísticos de orden", "medidas de posición", "resúmenes estadísticos" o "porcentaje de casos mayor/menor", entre otros. El posible problema con el lenguaje verbal es que es poco preciso y una misma palabra puede tener diferentes significados, que el alumno llega a diferenciar únicamente con la experiencia y el aprendizaje (Barwell, 2005).
- *Lenguaje simbólico:* un elemento característico del lenguaje matemático son los símbolos. Las notaciones más usuales para referirnos a los estadísticos de orden son Q_i para los cuantiles, C_i para los cuartiles, D_i para los deciles y P_i para los percentiles. En el caso particular de la mediana podemos referirnos a ella mediante las notaciones Me, P_{50}, D_5 y Q_2. Para el cálculo de estos estadísticos intervienen, además, otros elementos entre los que destacamos las siguientes notaciones: x_i para los datos, f_i para las frecuencias relativas, n_i para las frecuencias absolutas, I_{i-1} e I_i para los extremos inferior y superior de un intervalo respectivamente, a_i para la amplitud de dicho intervalo y *n* para referirnos al número total de datos, así como desigualdades, paréntesis o corchetes y símbolos como %.
- *Lenguaje gráfico y tabular:* este tipo de lenguaje nos permite resumir y visualizar la información de forma eficiente, aunque no está exento de complejidad, debido a los diferentes tipos de tablas y gráficos, cada uno de los cuáles requiere de la comprensión de otros objetos matemáticos para su interpretación (Pallauta, Gea y Batanero, 2020). Destacamos en relación con los estadísticos de orden los listados de datos aislados, que pueden presentarse ordenados (ascendente o descendente) o desordenados. También tenemos las tablas de frecuencia (datos aislados o agrupados en intervalos) y gráficos como los de barras o líneas, el gráfico de caja o el gráfico cuantil-cuantil.

Definiciones y propiedades

Existen diferentes definiciones equivalentes para los estadísticos de orden, que analizaremos con el ejemplo de la mediana, que se pueden generalizar fácilmente a otros estadísticos. Cada una de ellas resalta diferentes aspectos del significado del concepto:

- Si queremos enfatizar la idea de posición del estadístico dentro de la distribución, tenemos la siguiente definición: "La mediana, valor mediano o valor central de una serie numérica es el valor que ocupa el punto central cuando la serie está ordenada creciente o decrecientemente" (Nortes Checa, 1993, p. 69) o bien "Mediana es el valor de los datos que ocupa la posición media, cuando los datos están clasificados en orden, de acuerdo con su tamaño" (Johnson y Kuby, 2008, p. 75).
- Si se pretende resaltar el porcentaje de casos situados por encima y por debajo de la mediana, tomamos la siguiente: "Se define la mediana como el valor de la variable, tal que, supuestos ordenados todos los valores numéricos de la variable en orden creciente, la mitad son inferiores a él y la otra mitad iguales o superiores" (Nortes Checa, 1993, p. 69). O esta otra definición equivalente: en un conjunto finito de valores, la mediana es el valor que divide a la distribución en dos partes iguales, de tal forma que la mitad de los valores son mayores que la mediana y la otra mitad menores.
- Si relacionamos la mediana con la frecuencia acumulada: "La mediana es el valor de la variable estadística tal que la ordenada del diagrama acumulativo de frecuencias absolutas es igual a n/2, siendo n el número total de datos" (Calot, 1988, p. 56).
- También podemos definir la mediana como solución de una ecuación en la que interviene la función de distribución: "La mediana es el valor de la variable estadística tal que la ordenada de la curva de distribución empírica (representación gráfica de las frecuencias relativas acumuladas) es igual a ½" (Calot, 1988, p. 56).

Indiferentemente de la definición que utilicemos, los estadísticos de orden tienen las siguientes propiedades, que en Cobo (2003) aparecen estudiadas para el caso de la mediana y clasificadas en tres bloques: algebraicas, numéricas y estadísticas.

- En primer lugar, podemos considerar que el estadístico de orden es una función que asigna al conjunto de datos dado un valor numérico. Las *propiedades algebraicas* se refieren a la operación realizada con dicho conjunto de datos, que no es una operación interna (puede tomar un valor distinto al de los datos que componen el conjunto), no tiene elemento neutro ni simétrico, ni es asociativa. Conserva los cambios de escala y origen, las unidades de medida son las mismas que las de la variable y, además, es conmutativa (el orden de los datos no afecta al valor del estadístico).
- Por otro lado, cuando se analiza el valor numérico obtenido mediante la operación anterior, encontramos las *propiedades numéricas.* Una de estas propiedades es que un estadístico de orden está comprendido entre el mínimo y el máximo del conjunto de datos, pudiendo ser igual a uno de estos dos valores, pero no tiene por qué coincidir siempre con uno de los valores de los datos. Esto ocurre, por ejemplo, cuando se calcula la mediana en un conjunto par de datos y los dos datos centrales tienen diferente valor, en cuyo caso, se suele dar como valor de la

mediana la media aritmética de los dos valores que ocupan las posiciones centrales. Por otro lado, el valor del percentil del r% es una función creciente del valor r, es decir, cuanto mayor sea el porcentaje para el que se desea calcular el percentil, más alto será el valor de dicho percentil. No obstante, hay que tener en cuenta que no es una función lineal. Además, el valor numérico del estadístico de orden no tiene por qué cambiar cuando se cambia cualquier dato, pues en su cálculo no se tienen en cuenta los valores de los datos sino la posición que ocupan.

- Finalmente, cuando consideramos los estadísticos de orden en función de la información que nos entregan los datos, se tienen las *propiedades estadísticas.* Todos los estadísticos de orden son *robustos,* es decir, su valor no se ve afectado por los valores extremos. Además, la mediana es preferible como medida de tendencia central a la media o la moda en distribuciones asimétricas, con valores atípicos y no unimodales. También cuando la variable está agrupada en intervalos y existen intervalos abiertos y cuando se trabaja con variables cualitativas ordinales.

Procedimientos de cálculo

Los estadísticos de orden no tienen un único procedimiento de cálculo, sino que este depende del tipo de datos y la forma en que se presentan. Cobo y Batanero (2000) describen los siguientes procedimientos de cálculo de la mediana, que se generalizan de forma sencilla al resto de estadísticos de orden:

- *Cuando los datos no están agrupados y el número de datos es impar*, se calcula la mediana simplemente tomando el valor de la variable situado en el centro de la lista de datos ordenados por valores crecientes o decrecientes.
- *Para un número par de datos no agrupados,* la mediana es la media aritmética de los dos valores que se encuentren en el centro de la lista de datos ordenados por valores crecientes o decrecientes de la variable.
- *Si los datos son discretos y se da su distribución*, la mediana puede obtenerse de forma gráfica a partir del diagrama acumulativo de frecuencias, buscando el valor de la variable al que corresponde la frecuencia relativa acumulada 1/2. La ecuación F(Me)=1/2, en general, no tiene solución, puesto que la distribución acumulada es una función que varía por saltos, por lo que se derivan diferentes procedimientos de cálculo que se muestran en la Figura 1, siguiendo a Calot (1988, p. 57). Si el número de datos es impar, el valor n/2 (respectivamente ½, en frecuencia relativa) corta a la gráfica en el salto del diagrama acumulativo para uno de los valores de la variable. Dicho valor de la variable será la mediana. Si el número de datos es par, la mediana está indeterminada entre los valores x_i y $x_{i+1},$ ya que todos los valores del intervalo cumplen la definición de mediana (Figura 1). En este caso, se toma como mediana la media aritmética de estos dos valores.

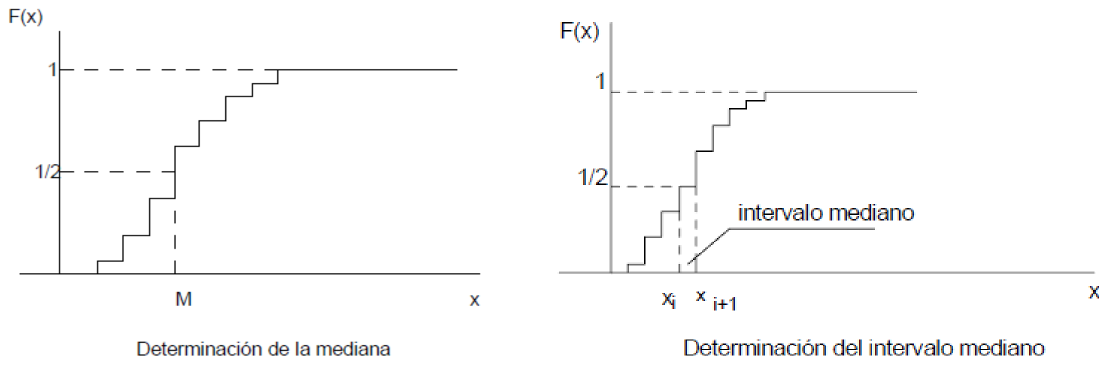

Figura 1. Cálculo de la mediana a partir del disgrama acumulativo de una variable discreta, con número par e impar de valores. Fuente: Calot (1988, p. 57)

- *Si se trata de una distribución de frecuencias con datos agrupados en intervalos,* además del procedimiento anterior, generalmente hay que interpolar para encontrar el valor exacto de la mediana y, en consecuencia, se obtiene un valor aproximado. Siguiendo a Batanero y Diaz-Batanero (2008, p. 56) describimos el procedimiento de cálculo en la siguiente Figura 2.

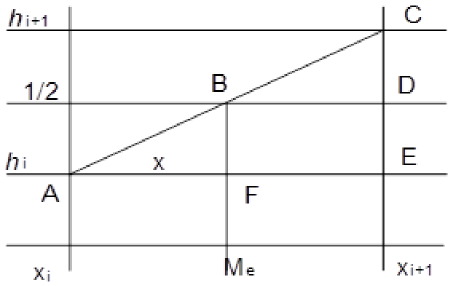

Figura 2. Cálculo de la mediana a partir de datos agrupados. Fuente: Batanero y Díaz (2008, p. 56)

El cálculo se realiza a partir del polígono de frecuencias acumuladas, buscando el valor cuya frecuencia relativa acumulada sea igual a 1/2. Dicha frecuencia corresponderá a un punto del diagrama acumulativo (B en la gráfica), que, ocasionalmente puede coincidir con uno de los extremos del intervalo. Para el caso general, se debe buscar la distancia x que hay que sumar al extremo inferior x_i del intervalo, aplicando proporcionalidad, de modo que:

$$\frac{x}{AE} = \frac{BF}{CE}, \text{ o lo que es lo mismo: } \frac{x}{x_{i+1} - x_i} = \frac{\frac{1}{2} - h_i}{h_{i+1} - h_i}$$

El cálculo de la mediana a partir de un gráfico dependerá del tipo de gráfico y puede ser bastante variable, como se describe en Cobo (2003). En todos los casos, el cálculo de la mediana implica tener en cuenta que ésta es el valor que divide a la población en dos partes de igual tamaño.

Para resumir, en la Tabla 1, se muestran las prácticas matemáticas requeridas para el cálculo de la mediana (equivalentemente para el cálculo de un percentil) mediante los diferentes procedimientos descritos. Consideramos también el uso de software para el cálculo que simplifica bastante el procedimiento.

Tabla 1. *Prácticas matemáticas requeridas en el cálculo de la mediana, mediante diferentes procedimientos*

Prácticas matemáticas requeridas	Procedimiento de cálculo					
	N.º impar datos aislados	N.º par datos aislados	Distribución de frecuencias con datos no agrupados	Distribución de frecuencias con datos agrupados	Gráfico	Software u hoja cálculo
Ordenar los datos	x	x				
Calcular la media		x	x (si F(x)=1/2 se alcanza entre dos valores de la variable)			
Buscar el valor x, tal que F(x)=1/2			x	x	x	
Representar gráficamente la distribución acumulada F(x)			x	x		
Leer e interpretar un gráfico			x	x	x	
Tomar el valor x, donde F(x) pasa de menor a mayor de 1/2			x			
Interpolar				x		
Utilizar la función mediana del software						x

Tabla 2. *Objetos matemáticos asociados al trabajo con diferentes procedimientos*

Objetos matemáticos asociados	Procedimiento de cálculo					
	N.º impar datos aislados	N.º par datos aislados	Distribución de frecuencias con datos no agrupados	Distribución de frecuencias con datos agrupados	Gráfico	Software u hoja cálculo
Variable, valor, orden, desigualdades, porcentaje	x	x	x	x	x	x
Número par o impar de datos	x	x	x	x		
Intervalo de clase, extremos				x	x (dependiendo del gráfico)	
Media		x	x (si F(x)=1/2 se alcanza entre dos valores de la variable)			

Frecuencia, Frecuencia acumulada, función de distribución F(x)			x	x	x	
Ecuación F(x)=1/2			x	x	x	
Función por saltos			x			
Proporcionalidad				x	x	
Interpolación, teorema Thales				x		
Representar una gráfica			x	x	x	
Interpretar una gráfica			x	x	x	
Función mediana del software						x

En la Tabla 2 se presentan los distintos objetos matemáticos elementales requeridos en dichos procedimientos.

Observamos en dichas tablas, que el estudiante debe aprender procedimientos diferenciados, dependiendo del tipo de dato y del modo en que se le proporcionan (agrupados o no). Además, la complejidad de los procedimientos es creciente, comenzando desde el cálculo con datos aislados, tablas de variables discretas o no agrupadas y tablas agrupadas. El cálculo con software es también mucho más simple, pues únicamente se debe conocer la función apropiada del software que produce la mediana o percentil.

Llamamos la atención al hecho de que, para el trabajo con datos presentados en tablas de frecuencia no agrupadas, la función de distribución que representa las frecuencias acumuladas es una función por salto, a la que los estudiantes no están acostumbrados en la educación secundaria (Estepa, 1994). Por otro lado, la interpolación requerida para encontrar la mediana o los percentiles en las tablas de datos agrupados en intervalos se basa en la proporcionalidad, no siempre manejada correctamente por los estudiantes.

Finalmente, habría que añadir los *argumentos* utilizados para el estudio de los estadísticos de orden. Estos argumentos a veces son formales y deductivos, utilizando lenguaje algebraico, por ejemplo, para deducir la fórmula de cálculo de un estadístico. Pero la mayoría de las veces, en los niveles educativos considerados, se utilizan argumentos informales tales como mostrar un ejemplo para analizar una propiedad o para introducir una definición o apoyarse en un gráfico, como hemos mostrado en la presentación de los procedimientos de cálculo. Estos argumentos se utilizan también en otros temas matemáticos, por tanto no son específicos de los estadísticos de orden, por lo que no los analizamos con mayor detalle en este trabajo.

En las siguientes secciones detallamos más específicamente las dificultades que los estudiantes encuentran en el trabajo con cada uno de los tipos de objetos matemáticos elementales asociados a los estadísticos de orden que se han descrito en estos apartados.

SIGNIFICADOS PERSONALES DE LOS ESTADÍSTICOS DE ORDEN

Al enseñar los estadísticos de orden, el profesor trata de ayudar a sus estudiantes a adquirir los diferentes elementos de significado (campos de problemas, lenguaje, definiciones, propiedades, procedimientos y argumentos) considerados adecuados desde el punto de vista institucional. Para ello, plantea algunos de los problemas descritos en la sección anterior, esperando que, al resolverlos, el estudiante vaya adquiriendo una comprensión gradual del tema. Sin embargo, veremos a continuación, que algunos estudiantes dotan de un significado diferente al institucional (personal) a algunos de estos elementos.

Comprensión de los campos de problemas

El primer punto en el aprendizaje es reconocer los problemas para los que un objeto matemático (en este caso, los estadísticos de orden) es adecuado, lo que no siempre ocurre.

Por ejemplo, en una investigación llevada a cabo por Mayén, Díaz y Batanero (2009), las autoras propusieron a 518 estudiantes mexicanos de cuarto curso de la ESO y segundo de Bachillerato un cuestionario con diferentes problemas referidos a la mediana. Todos los estudiantes habían estudiado la mediana siguiendo las directrices curriculares. En uno de los problemas del cuestionario se pidió calcular la mediana para un pequeño conjunto de datos impar y al que luego se añade un valor atípico. Se preguntó también a los sujetos del estudio cuál era el mejor representante del conjunto de datos (media, mediana o moda), (campo de problemas CP1). Únicamente el 30% de la muestra supo reconocer que la mediana era mejor representante debido al valor atípico, muy apartado del resto de valores. La explicación es que algunos estudiantes argumentan que la media es siempre el mejor representante de un conjunto de datos, lo que atribuimos al uso extendido de la media en la prensa y medios de comunicación, mientras que el empleo de la mediana es más limitado.

Este error se repite en una investigación de Delson y Mugabe (2013) con 25 estudiantes universitarios de primer curso de estadística, por lo que la dificultad continúa en estos estudiantes. Por su parte, Boaventura y Fernándes (2004), en una investigación con 181 estudiantes de segundo curso de Bachillerato, encontraron que prácticamente todos los estudiantes eligieron la media y la moda como representantes de un conjunto de datos, sin tener en cuenta la mediana en los casos en que era preferible al pedir identificar la mejor medida de posición central para un conjunto de datos.

Mayén y Díaz (2010) analizan en la misma muestra del estudio anterior otro problema en que se pide comparar dos conjuntos de datos ordinales e indican que únicamente el 34% resolvió correctamente el problema, utilizando la mediana para comparar los dos conjuntos de datos. Algunos estudiantes compararon las frecuencias de los datos con el valor máximo o mínimo, mientras que otros transformaron los datos a valores numéricos y compararon las medias, modas o máximos de los datos transformados, lo cual es una solución incorrecta. El mismo problema fue planteado por Cobo (2003) a estudiantes de primer y cuarto curso de la ESO, obteniendo sólo un 13% de soluciones correctas.

Los problemas anteriores aparecen en ocasiones en profesores o futuros profesores. Por ejemplo, Groth y Bergner (2006) analizaron la comprensión de 36 futuros profesores de educación primaria sobre las medidas de posición central, diferenciando varios niveles de comprensión. Sólo tres de los participantes fueron clasificados en el nivel superior, donde se comprende la diferencia de la mediana con la media y moda y se es capaz de decidir cuándo una medida es más adecuada para un conjunto de datos que otra. Por su parte, Estrada, Batanero y Fortuny (2004) encuentran que el 42% en una muestra de 367 futuros profesores de educación primaria no detecta el efecto de los valores atípicos en el cálculo de la media o mediana, y otro 14% no fue capaz de discernir cuándo un valor es atípico para un contexto dado.

En otro trabajo con 65 futuros profesores de matemáticas de ESO y Bachillerato, Gea, Batanero, Fernández y Arteaga (2016) pidieron a los participantes decidir la mejor medida de posición central para la distribución de la variable esperanza de vida al nacer, que es claramente asimétrica, proporcionándoles el histograma, diagrama acumulativo y diagrama de caja de dicha distribución, así como los valores de media, mediana y moda. El 12% de los participantes eligen la media indicando que es más representativa al intervenir en su cálculo todos los valores de los datos. Una minoría fue incapaz de decidirse por una de las medidas de posición central.

Comprensión del lenguaje

El empleo de una variedad de lenguajes por parte del estudiante es uno de los principales desafíos en la enseñanza de la matemática (Schleppegrell, 2007) y este hecho se refleja en el aprendizaje del lenguaje asociado a los estadísticos de orden. Se ha detectado que, en muchas ocasiones, los alumnos confunden la terminología de media y mediana. Esto provoca que calculen la media cuando realmente se les está pidiendo que calculen la mediana, como encuentra Cobo (2003) en un estudio con 168 estudiantes de primer curso de la ESO y 144 de cuarto curso de la ESO. Del mismo modo, en el estudio de Mayén, Díaz y Batanero (2009), algunos estudiantes confundieron la representación simbólica de la mediana con la de la media o la moda.

En otros casos se observa dificultad en la interpretación del lenguaje gráfico asociado a los estadísticos de orden. Así, Gea, Arteaga y Cañadas (2017) piden a 65 futuros profesores de matemáticas en la ESO y Bachillerato, licenciados o ingenieros, que se preparaban como profesores en un curso de Máster, interpretar un diagrama acumulativo y un gráfico de caja de la distribución de la esperanza de vida al nacer en un conjunto de países. Algunos estudiantes se limitan a dar una descripción del contenido del gráfico, por ejemplo, el 58% se limita a indicar que el diagrama acumulativo representa las frecuencias acumuladas de la variable. Además, se encuentran interpretaciones incorrectas, pues un 18% de los participantes confunden frecuencias acumuladas y no acumuladas o bien frecuencias y valor de la variable.

La interpretación del diagrama de caja fue generalmente correcta, aunque el 29,3% se limita a describir el gráfico. Por otro lado, un 7,7% de los casos piensa que el porcentaje de casos comprendidos entre cada dos de los cinco puntos que definen el gráfico (mínimo, primer cuartil, mediana, tercer cuartil y máximo) no es exactamente el mismo y uno de los futuros profesores indicó que el gráfico de la caja representa sólo información sobre la esperanza de vida de la cuarta parte de los países de la muestra.

Comprensión de definiciones y propiedades

La mayor parte de las investigaciones relacionadas con la comprensión de definiciones y propiedades se centran en la mediana, aunque, presumiblemente, muchos de los errores que describimos a continuación se deben extender a los estadísticos de orden.

Son varios los investigadores que sugieren que la definición de mediana no es clara para los estudiantes y la confunden con la media (Carvalho, 2001; Mayén, Díaz y Batanero, 2009), pues, de acuerdo a Barros (2003), es la medida de posición central más compleja de interpretar para los estudiantes. Barr (1980) fue uno de los primeros investigadores en analizar su comprensión, por parte de 95 estudiantes de entre 17 y 21 años, e indica que, aunque la mayoría de ellos conciben la mediana como centro, muchos no saben qué tipo de centro es. A este respecto, Mayén, Díaz y Batanero (2009), por su parte, informan que algunos estudiantes entienden la mediana como centro del conjunto de datos desordenado o como centro geométrico de la distribución (punto medio entre el máximo y el mínimo), interpretación también encontrada por Barros (2003) y Sousa (2002). Todos estos resultados apoyan la dificultad en dar una definición correcta de mediana encontrada tanto por Cobo (2003), como por Mayén, Cobo, Batanero y Balderas (2007).

Las propiedades aritméticas de la mediana no parecen implicar dificultad para los estudiantes. Respecto a las propiedades algebraicas, Mayén, Díaz y Batanero (2009) indican que los estudiantes generalizan propiedades de operaciones aritméticas atribuyéndoselas a la mediana; por ello, suponen que la mediana es una operación interna en el conjunto de datos (de modo que piensan que un conjunto de datos enteros ha de tener una mediana entera).

En relación con las propiedades estadísticas, se sitúa la mediana siempre en el centro geométrico de la distribución, aunque esto sólo ocurre si la distribución es simétrica. No se entiende la propiedad de robustez, porque no se comprende el efecto del valor atípico sobre la media o mediana o no detectan los valores atípicos en una distribución. Como ya hemos indicado al analizar los campos de problemas, muchos estudiantes no comprenden que la mediana puede ser mejor representante de un conjunto de datos que la media, en algunos casos.

Algunas de estas dificultades se presentan en futuros profesores. Por ejemplo, al pedirles interpretar los resúmenes estadísticos de la distribución de la esperanza de vida al nacer, Gea et al. (2016) informan que algunos estudiantes interpretaron que todos los países por debajo del primer cuartil tienen la misma esperanza de vida (el valor del cuartil). Otros pensaban que la mediana debía encontrarse dentro del intervalo modal o que la media era preferible a la mediana en una distribución asimétrica.

Procedimientos de cálculo

Los procedimientos de cálculo son una de las mayores fuentes de dificultad en el trabajo con los estadísticos de orden, debido a la existencia de los varios algoritmos de cálculo descritos al estudiar el significado institucional de estos estadísticos. Como indica Schuyten (2001), incluso los alumnos universitarios encuentran difícil aceptar que se puedan emplear dos algoritmos diferentes de cálculo para un mismo resumen estadístico o que puedan obtenerse valores distintos para el estadístico, al variar la amplitud de los intervalos de clase. Aunque se comprenda la definición del estadístico de orden, para llegar a su cálculo, hay que seguir una serie de pasos y tomar decisiones que el estudiante aprende de memoria sin comprender realmente. Otros estudiantes son capaces de calcular la mediana o un percentil de un conjunto pequeño de datos, pero no cuando se les dan organizados en una tabla de frecuencias (Schuyten, 2001).

También la principal parte de la investigación relacionada con los procedimientos de cálculo se ha limitado al trabajo con la mediana, pero, aun así, se encuentra dificultad incluso con los algoritmos más sencillos. Así, Zawojewski and Shaughnessy (2000) encontraron que un tercio de los estudiantes de segundo curso de Bachillerato (grado 12) eran incapaces de calcular correctamente la mediana de un conjunto de datos no ordenado, error también encontrado por Barr (1980), Cobo (2003) y Mayén, Díaz y Batanero (2009).

Estepa (1994) encontró estudiantes con dificultades para calcular la mediana, a partir de las representaciones gráficas de las frecuencias acumuladas. Para las variables agrupadas en intervalos, una vez construida la gráfica de frecuencias acumuladas, el estudiante tiene que interpolar para hallar el valor de la mediana y los fallos se producen por falta de razonamiento proporcional e insuficiente dominio del manejo de desigualdades. Si se trata de una variable discreta, la gráfica de frecuencias acumulada toma forma de escalera, pero los estudiantes no están acostumbrados a las funciones discontinuas a saltos, por lo que dudan si el valor de la mediana corresponde a un valor de la variable o al siguiente.

Carvalho (2001) observa que los estudiantes de 13-14 años no ordenan los datos para calcular la mediana, error también encontrado en Cobo (2003) y Mayén (2009). En algunos casos, calculan el dato central de las frecuencias absolutas ordenadas de forma creciente; es decir, confunden la frecuencia con el valor de la variable y, en otros, calculan la moda en lugar de la mediana. Barros (2003), Mayén et al., (2007) y Sousa (2002) señalan que algunos estudiantes utilizan solo los valores de la variable, pero no su frecuencia, al calcular la mediana. Por su parte Cobo (2003) encontró que lo más difícil para los estudiantes de su muestra fue calcular la mediana a partir de un gráfico, posiblemente por falta de competencia en la lectura del mismo.

En el caso de cálculo a partir de una tabla de frecuencias nos encontramos con el problema de que la ecuación F(Me)=1/2, en general, no tiene solución puesto que la función F(x) para variables discretas es una función por saltos. Esto da lugar a varios procedimientos de cálculo. Si el número de datos es impar, el valor n/2 corta a la gráfica en el salto del diagrama acumulativo para uno de los valores de la variable que coincide con el valor de la mediana. Si el número de datos es par, la mediana está indeterminada entre los valores x_i y x_{i+1} y suele tomarse como mediana la media aritmética de estos dos valores.

Finalmente, Cobo (2003) nos informa que sólo el 4% de los estudiantes de su muestra fue capaz de calcular correctamente la mediana a partir de datos dados en un diagrama de barras, posiblemente por falta de facilidad para traducir el gráfico a un listado o tabla de frecuencias.

Estas dificultades aparecen también en los profesores y futuros profesores de educación primaria. Así Friel y Bright (1998) indican que los profesores de educación primaria pueden tener dificultades al calcular la mediana de un conjunto de datos presentado gráficamente. En otro ejemplo, Jacobbe (2012) entrevista a tres profesores con experiencia en la enseñanza de estadística, incluidas las medidas de posición central en educación primaria, e indica que estos no comprenden la razón del diferente algoritmo de cálculo de la mediana ni saben explicar la diferencia conceptual entre media y mediana.

CONCLUSIONES

Los estadísticos de orden tienen una definición muy sencilla, al menos en la primera definición descrita en el trabajo. Pero la comprensión de la definición no ayuda al estudiante a realizar un cálculo correcto, debido a los numerosos algoritmos diferentes, que dependen del tipo de datos y de cómo se presentan. Por otro lado, no siempre se sabe cuándo se debe aplicar un estadístico de orden, por ejemplo, la mediana, quizás porque sea más frecuente pedir al estudiante calcular dicho estadístico que proponerle problemas donde deba elegir qué estadístico aplicar. Son también confusas para los estudiantes las notaciones y las propiedades asociadas. Las mencionadas dificultades no parecen resolverse con la enseñanza actual, ya que aparecen incluso en Bachillerato y futuros profesores.

Estos resultados nos sugieren la necesidad de pensar en la enseñanza actual y si todos los contenidos incluidos deben ser presentados a los estudiantes. Puesto que el software estadístico, tal como CODAP (https://codap.concord.org/) o incluso la misma hoja de cálculo, permiten calcular correctamente los estadísticos, podría plantearse eliminar la enseñanza de los diferentes métodos de cálculo y concentrarse en la comprensión de la definición, propiedades y campos de aplicación.

Una sugerencia es incorporar el trabajo con proyectos estadísticos, usando datos reales, que con frecuencia no siguen la distribución normal o tienen valores atípicos, que pueden reforzar el conocimiento de propiedades y campos de problemas de estos estadísticos. Los diferentes recursos manipulativos disponibles en internet pueden completar esta comprensión.

Finalmente, coincidimos con Jacobbe (2012) en que las carencias de conocimiento en los futuros profesores sugieren la necesidad de una formación específica en el tema. Es preocupante que las pocas investigaciones relacionadas con el tema muestren la existencia de dificultades similares a las que tienen los estudiantes. En este sentido, es necesario mejorar la formación del profesorado en estadística y asimismo en didáctica de la estadística. El contenido de este trabajo puede ser utilizado como lectura y fuente de discusión y reflexión para mejorar la formación del profesorado.

AGRADECIMIENTOS

Proyecto PID2019-105601GB-I00 (AEI) y Grupo FQM126 (Junta de Andalucía).

REFERENCIAS

Barr, G. V. (1980). Some student ideas on the median and the mode. *Teaching Statistics, 2*(2), 38-41.

Barros, P. (2003). *Os futuros professores do 2.º ciclo e a estocástica – Dificuldades sentidas e o ensino do tema.* Lisboa: Associação de Professores de Matemática.

Barwell, R. (2005). Ambiguity in the mathematics classroom. *Language and Education 19*(2), 118–126. DOI: https://doi.org/10.1080/09500780508668667

Batanero, C. y Díaz, C. (2008). *Análisis de datos con Statgraphics*. Departamento de Didáctica de la Matemática, Universidad de Granada.

Betanzos, F. G. y López, J. K. C. (2017). *Estadística aplicada en psicología y ciencias de la salud.* Madrid: Manual Moderno.

Boaventura, M. G. y Fernandes, J. (2004). Dificuldades de alunos do 12.º ano nas medidas de tendência central: O contributo dos manuais escolares. Em J. A. Fernandes (Ed.). *Actas do I Encontro de Probabilidades e Estatística na Escola* (pp. 103-126). Braga: Universidade da Minho.

Calot G. (1988). *Curso de estadística descriptiva.* Madrid: Paraninfo.

Carvalho, C. (2001). *Interaçao entre pares. Contributos para a promoçao do desenvolvimiento lógico e do desempenho estatístico no 7º ano de escolaridade*. Tesis Doctoral. Universidad de Lisboa.

Chambers, J. M. (2018). *Graphical methods for data analysis,* 2º ed. London: Taylor and Francis.

Cobo, B. (2003). *Significado de las medidas de posición central para los estudiantes de secundaria*. Tesis doctoral. Universidad de Granada.

Cobo, B. y Batanero, C. (2000). La mediana en la educación secundaria obligatoria: ¿un concepto sencillo? *UNO, 23,* 85-96.

Delson, A. y Mugabe, D. (2013). O conceito da mediana na perspectiva dos estudantes principiantes. *International Journal of Scientific & Technology Research, 2*(9), 202-206.

Engel, J. (2019). Cultura estadística y sociedad. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Disponible en www.ugr.es/local/fqm126/civeest.html.

Estepa, A. (1994). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores.* Tesis doctoral. Departamento de Didáctica de la Matemática. Universidad de Granada.

Estrada, A., Batanero, C. y Fortuny, J. M. (2004). Un estudio sobre conocimientos de estadística elemental de profesores en formación. *Educación Matemática, 16*(1), 89-111.

Friel, S. N. y Bright, G.W. (1998). Teach-Stat: A model for professional development in data analysis and statistics for teachers K–6. En S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in Grades K–12* (pp. 89–117). Mahwah, NJ: Lawrence Erlbaum.

Gea, M.M., Arteaga, P. y Cañadas, G.R. (2017). Interpretación de gráficos estadísticos por futuros profesores de Educación Secundaria. *Avances de Investigación en Educación Matemática, 12*, 19-37.

Gea, M. M., Batanero, C., Fernández, J. A. y Arteaga, P. (2016). Interpretación de resúmenes estadísticos por futuros profesores de educación secundaria. *Journal of Research in Mathematics Education, 5*(2), 135-157.

Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactiques des Mathematiques, 22*(2/3), 237-284.

Godino, J. D. y Batanero, C. (1998). Clarifying the meaning of mathematical objectsas a priority area of research in mathematics education. En A. Sierpinska y J. Kilpatrick (Eds.), *Mathematics education as a research domain: A search for identity* (pp. 177-195). Dordrecht: Kluwer.

Godino, J.D., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education, 39*(1-2), 127-135.

Godino, J. D., Batanero, C. y Font, V. (2019). The onto-semiotic approach: implications for the prescriptive character of didactics. *For the Learning of Mathematics, 39*(1), 38-43.

Groth, R. E. y Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning, 8,* 37–63.

Hoaglin, D. C., Mosteller, F. y Tukey, J. W. (1983*). Understanding robust and exploratory data analysis.* New York: Wiley.

Jacobbe, T. (2012). Elementary school teachers'understanding of the mean and median. *International Journal of Science and Mathematics Education, 10*(5), 1143-1161.

Johnson, R. y Kuby, P. (2008). *Estadistica elemental: Lo esencial* (10ª ed.). Londres: Cengage Learning Editores.

Mayén, S. (2009). *Comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria y bachillerato.* Tesis doctoral. Universidad de Granada.

Mayén, S., Cobo, B., Batanero, C. y Balderas, P. (2007). Comprensión de las medidas de posición central en estudiantes mexicanos de bachillerato. *Unión, 9*(1), 187-201.

Mayén, S. y Díaz, C. (2010). Is median an easy concept? Semiotic analysis of an open-ended task. In K. Makar (Ed.), *Proceedings the Eighth International Conference on Teaching Statistics*. Voorburg, The Netherlands: International Statistical Institute. Disponible en: http://iase-web. org/documents/papers/icots8/ICOTS8_C265_MAYEN. Pdf

Mayén, S., Díaz, C. y Batanero, C. (2009). Students´ semiotic conflicts in the concept of median. *Statistics Education Research Journal, 8*(2), 74-93.

MECD (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria y del Bachillerato.* Madrid: MECD.

Nortes Checa, A. (1993). *Estadística teórica y aplicada.* Barcelona: PPU.

Ortiz, J. J., Batanero, C. y Serrano, L. (2001). El lenguaje probabilístico en los libros de texto. *Suma, 38,* 5-14.

Pallauta, J.D., Gea, M.M. y Batanero, C. (2020) Un análisis semiótico del objeto tabla estadística en los libros de texto chilenos. *Zetetiké*, *28*, 1-19-e020001. DOI: https://doi.org/10.20396/zet.v28i0.8656257.

Ridgway, J., Nicholson, J. y McCusker, S. (2011). Statistical literacy, globalisation, and the internet. Trabajo presentado en el *58th World Statistical Congress*. Dublin: International Statistical Institute.

Schleppegrell, M. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading and Writing Quarterly, 23,* 139-159.

Schuyten, G. (2001). Research skills: A closely connected triplet of research area, research methodology and statistics. En C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 227-230). Granada: IASE.

Sousa, O. (2002). Investigações estatísticas no 6.º ano. Em Grupo de Trabalho de Investigação (Org.), Reflectir e investigar sobre a prática profissional (pp. 75-97). Lisboa: APM.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Velleman, P. F. y Hoaglin, D. C. (1981). *Applications, basics, and computing of exploratory data analysis*. London: Duxbury Press.

Verma J.P. (2019) Non-parametric tests for psychological data. En J. P. Verma (Ed.), *Statistics and research methods in psychology with Excel* (pp. 477- 521). Singapore: Springer,

Zar, J. H. (2005). Spearman rank correlation. En T. Colton y P. Armitage (Eds.), *Encyclopedia of Biostatistics, 7*. DOI: https://doi.org/10.1002/0470011815.b2a15150

Zawojewski, J. S. y Shaughnessy, J. S. (2000). Data and chance. En E. A. Silver y P. A. Kenney (Eds.), *Results from the seventh mathematics assessment of the national assessment of educational progress* (pp. 235–268). Reston, VA: National Council of Teachers of Mathematics.

Carmen Batanero
Universidad de Granada, España
batanero@ugr.es

Silvia M. Valenzuela-Ruiz
Universidad de Granada, España
svalenzuela@ugr.es

María Magdalena Gea
Universidad de Granada, España
mmgea@ugr.es