MEDIA, MEDIANA Y MODA ¿QUÉ SIGNIFICA ESTO PARA LOS ESTUDIANTES DE SECUNDARIA?

Belén Cobo y Carmen Díaz

27 Congreso Nacional de Estadística e Investigación Operativa
Lleida, 8-11 de abril de 2003

RESUMEN

En este trabajo presentamos algunos resultados obtenidos del análisis de un cuestionario pasado a una muestra de estudiantes de secundaria, con el que pretendíamos estudiar el significado personal que éstos asignan a los promedios. El cuestionario está compuesto por 9 preguntas abiertas (con 26 apartados en total) que los estudiantes debían contestar, razonando sus respuestas. Mediante un análisis multivariante se comparan los resultados de las dos muestras, 168 estudiantes de 13 años y 168 de 16 años, así como del conjunto formado por el total de los participantes en el estudio.

Palabras y frases clave: Educación estadística, investigación
Clasificación AMS: 97C99

1.Introducción

Cuando queremos reflexionar sobre la dificultad que el aprendizaje de ciertos conceptos tiene para los alumnos, es necesario comenzar por hacer un análisis epistemológico de su significado. Como indica Godino (1996), "*el problema de la comprensión está íntimamente ligado a cómo se concibe el propio conocimiento matemático. Esta explicitación requiere responder a preguntas tales como: ¿Cuál es la estructura del objeto a comprender? ¿Qué formas o modos posibles de comprensión existen para cada concepto? ¿Qué aspectos o componentes de los conceptos matemáticos es posible y deseable que aprendan los estudiantes en un momento y circunstancias dadas? ¿Cómo se desarrollan estos componentes?* (pg. 418).
El *propósito* de nuestro trabajo es continuar nuestras investigaciones sobre las medidas de posición central (Batanero, 2000; Cobo, 1988; 2001; Cobo y Batanero, 2000, Navas, Batanero y Godino, 1997). Nos centraremos en la idea de media aritmética, que, aunque aparentemente simple, tienen un *significado* complejo. Siguiendo a Batanero y Godino (2001; 2002) vamos a considerar las siguientes entidades primarias como constituyentes del significado de la media:

1. *Problemas y situaciones* que inducen actividades matemáticas y definen el campo de problemas de donde surge el objeto. Un ejemplo sería repartir una cantidad en forma equitativa, que es el tipo de problema que hemos planteado en este trabajo.

2. *Procedimientos, algoritmos, operaciones*. Cuando un sujeto se enfrenta a un problema y trata de resolverlo, realiza distintos tipos de *prácticas*, que llega a convertir en rutinas con el tiempo. Prácticas características en la solución de problemas de promedios serían sumar una serie de valores y dividir por el número de sumandos.

3. Representaciones materiales utilizadas en la actividad matemática (términos, expresiones, símbolos, tablas, gráficos).

4. *Conceptos y proposiciones*. Las definiciones y propiedades características y sus relaciones con otros conceptos.

5. *Demostraciones* que empleamos para probar las propiedades del concepto y que llegan a formar parte de su significado.

2. Investigaciones previas

A pesar de ser uno de los principales conceptos estadísticos, este concepto no es bien comprendido por los estudiantes. Por ejemplo, Pollatsek, Lima y Well (1981) encontraron que las situaciones en las cuales se debe calcular una media ponderada y la selección de los correspondientes pesos no son fácilmente identificados por los estudiantes. Li y Shen (1992) indican que cuando los datos se agrupan en intervalos, los estudiantes olvidan con frecuencia que cada uno de estos grupos debería ponderarse de modo distinto al calcular la media, error que también es encontrado por Carvalho (1998). Cai (1995) encontró que mientras la mayoría de alumnos de 12-13 años son capaces de aplicar adecuadamente el algoritmo para calcular la media, sólo algunos saben determinar un valor desconocido en un conjunto pequeño de datos para obtener un valor medio dado. Mevarech (1983) sugiere que los estudiantes suelen creer que un conjunto de números, junto con la operación media aritmética satisface los cuatro axiomas de clausura, asociatividad, elemento neutro y elemento inverso. Batanero, Godino y Navas (1997), observaron que los profesores de primaria en formación, encuentran dificultades en el tratamiento de los ceros y valores atípicos en el cálculo de promedios, en identificar las posiciones relativas de media, mediana y moda en distribuciones asimétricas. Watson y Moritz (2000), analizan el significado intuitivo dado por los niños al término "promedio" y hallan un gran número de niños para los cuales el promedio es simplemente un valor en el centro de la distribución (es una idea próxima al concepto de mediana). Estas investigaciones se han centrado en puntos aislados de la comprensión, por ejemplo el cálculo o la comprensión de propiedades. En nuestro trabajo estamos considerando los cinco tipos de comprensión de nuestro modelo teórico.

3. El estudio

Nos basamos en los resultados de un cuestionario pasado a una muestra formada de alumnos y alumnas de primer curso de Enseñanza Secundaria Obligatoria, cuyas respuestas se clasificaron teniendo en cuenta los elementos de significado de la media que usaron los alumnos, Una vez recogidos los datos, se realizó un análisis de contenido de las respuestas dadas por los alumnos a cada una de las preguntas planteadas en los diferentes ítems de la prueba. Puesto que las respuestas eran abiertas, estas respuestas no se limitan a dos opciones (correctas/incorrectas) sino que se pidió a los alumnos, al justificarlas tenían libertad para emplear los diferentes elementos de significado previstos en el análisis a priori de los items.

	Item							
	1	2	3	4	5	6	7	8
Definición de media	X	X	X	X	X	X	X	X
De Definición de mediana								X
Definición de moda								X
La media no conserva el conjunto numérico	X							
La media puede no coincidir con ningún dato	X						X	
Todos los datos son relevantes al calcular la media					X	X		
La media no es una operación interna	X						X	
Los ceros afectan a la media						X		
Media de la suma		X						
La media es conmutativa						X		
La media no es asociacitva		X						
La media es un valor representativo	X						X	
Suma de desviaciones a la media			X	X				
La media y mediana sólo coinciden en distribuciones simétricas					X			
La mediana es robusta, la media no					X			
Hallar un valor representativo			X			X		
Hacer un reparto equitativo	X		X			X		
Adivinar un valor probable	X							
Estimar un valor desconocido a partir de varias medidas								
Calcular la media de datos aislados	X	X		X	X	X	X	X
Calcular medias ponderadas		X						
Estimar la media de un gráfico								X
Estimar la moda de un gráfico								X
Invertir el algoritmo de la media	X			X		X		
Construir unos datos de media dada						X		
Calcular la mediana de datos aislados					X			
Estimar la mediana de un gráfico								X

Tabla 1. Significado evaluado en el cuestionario

La parte que presentamos del cuestionario consta de los 8 ítems que fueron comunes a los dos grupos y que se presentan como apéndice. Se eligieron en base a nuestro análisis teórico previo del concepto y del estudio de los libros de texto de secundaria (Cobo, 1998, 2001). Los items 1 y 2 se adaptaron de Watson y Moritz (2000); los 3, 4 de Tormo (1993); el 6 de Gattusso (1996); el 7 de Konold and Garfield (1992) y el 8 de Zawojeswki (1986); el item 5 fue nuevo en esta investigación. Los estudiantes respondieron individualmente al cuestionario y se les animó a que diesen explicaciones detalladas de sus respuestas. Presentaremos aquí el análisis cuantitativo, dejando el análisis cualitativo para trabajos posteriores.

4. Resultados y discusión

En la Tabla 2 presentamos los resiultados por grupos. Obtuvimos un coeficientes de fiabilidad Alfa=0.78, que coincide con el coeficiente de generalizabilidad por ítem (G_I=0.78) y un coeficiente de generalizabilidad para estudiantes G_S=0.98. Esto sugiere que nuestros resultados son más generalizables a otros estudiantes (con los mismos items) que a otros items (con los mismos estudiantes).

	Edad 13 (n=168)		Edad 16 (n=144)	
Item	Porcentaje	Desviación típica	Porcentaje	Desviación típica
P1.a	.63	.48	.69	.46
P1.b	.27	.44	.37	.48
P2.a	.14	.34	.34	.48
P2.b	.12	.32	.38	.49
P2.c	.36	.48	.33	.47
P3	.46	.50	.49	.50
P4	.45	.50	.66	.48
P5.a	.38	.49	.38	.49
P5.b	.22	.42	.32	.47
P5.c	.09	.29	.33	.47
P6.a	.48	.50	.67	.47
P6.b	.51	.50	.68	.47
P6.c	.53	.50	.61	.49
P7	.39	.49	.67	.47
P8.a	.50	.50	.67	.47
P8.b	.21	.41	.26	.44
P8.c	.04	.21	.20	.40

Tabla 2. Porcentaje de respuestas correctas y desviaciones típicas por curso

Una prueba de Anova multivariantes (variable dependiente: vector de respuestas en los ítems) fue estadísticamente significativa (Lambda de Wilk's = .731; F=6.258; p-value <.001) respecto a los factores curso y centro (factor aleatorio; Wilk's Lambda= .510; F=3.127; p-value <.001) pero no respecto al género o las interacciones. Esto sugiere que la comprensión de los promedios crece con la edad e instrucción, aunque los resultados no son homogéneos en los diferentes centros o con respecto a los diferentes ítems.

La mayor diferencia se produjo en el item 7, que se refiere a propiedades abstractas de los promedios (La media puede no coincidir con ningún valor de los datos; la media no es una operación interna; la media es un valor representativo), Item 4 (Suma de desviaciones a la media; encontrar un reparto equitativo, invertir el algoritmo de la media), Item 6.a (invertir el algoritmo de la media; construir una distribución de media dada) y 6.b (La media es commutativa), Item 8.a (Estimar la media a partir de un gráfico), para el que un gran porcentaje de alumnos dio respuestas correctas después de la enseñanza. Hubo también una mejora significativa en los ítems 2.b y 2.a (La media no es asociativa; cálculo de medias ponderadas), Item 5.c (La mediana es robusta; la media no) 5.b (Cálculo de la media a partir de datos aislados), y 8.c (Estimar la moda de un gráfico), aunque estos ítems todavía son difíciles después de la instrucción.

No hubo diferencias en 2.c (Media de la suma de dos variables), 5.a (Hallar el mejor valor representante), e Item 3 (Hallar el mejor valor representante; Hallar un reparto equitativo; Suma de deviaciones a la media), aunque en estos dos ítems más de la mitad de los estudiantes de ambos grupos dieron respuestas correctas. Los resultadfos fueron bastante pobres en los items 6.c (Al calcular la media todos los datos son relevantes; un valor cero afecta a la media), Item 8.b (Estimar la media de un gráfico) que no mejoraron con la instrucción.

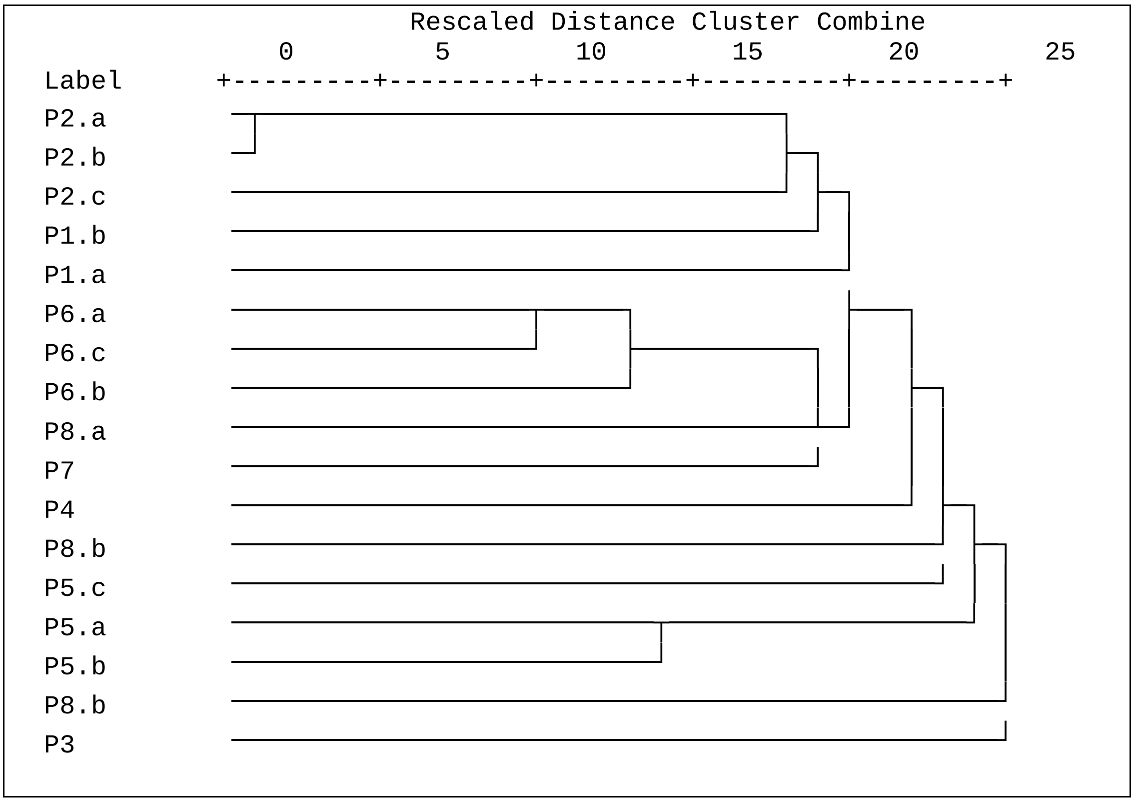


Figura 1. Dendograma en el análisis cluster

Hicimos un análisis cluster de las respuestas a los items (correcta=1, incorrecta=0), usando el coeficiente de correlación como medida de similitude y el método del vecino más próximo como método de aglomeración (ver la Figura 2). Los resultados sugieren una estructura multidimensional, que coincide con lo esperado en nuestro modelo teórico, según el cual los estudiantes podrían comprender un componente del significado de los promedios y no otros. Hay varios clusters, generalmente sólo se agrupan los subítems de un mismo ítem (e.g. 2.a, 2.b, 2.c; 1.a, 1.b; 6.a, 6.b, 6.c, 5.a, 5.b). Otros items permanecen aislados, incluso en caso de subítems del mismo ítem (e.g. en el ítem 7).

5. Conclusiones

Algunas investigaciones recientes tratan de describir niveles de comprensión de los conceptos estadístico, en particular los promedios (Watson & Moritz, 2000). Esta investigación se apoya en los marcos neopiagetiamos, como el de Biggs and Collis (1982, 1991), que usan las respuestas de los estudiantes para clasificarlos en estados discretos, dentro de un continuo unidimensional. Aunque reconocemos la importancia de este tipo de investigación, que usa muestras muy amplias, y el interés que tienen al proporcional al profesor criterios para organizar la instrucción de un tema (como los promedios) en los diversos niveles curriculares, hacemos notar que los ítems usados por Watson y Moritz no se basaron en un estudio epistemológico previo y no evaluaron el significado completo de los promedios.

Nuestros resultados se basan en un tipo más variado de items que los usados en dicha investigación y sugieren una estructura más compleja y no lineal en la comprensión de los estudiantes. Incluso con una muestra moderada, son bastante generalizables a otros alumnos y el análisis multivariantes (no usado por otros investigadores) sugiere la posibilidad de una estructura multifactorial en la comprensión. Estos resultados se confirmaron también en el análisis factorial y en el análisis cualitativo de las respuestas, que muestra cómo diferentes alumnos usan una variedad de elementos correcta e incorrectamente para resolver la misma tarea. En nuestro modelo, comprender un concepto es un proceso constructivo continuo en que el estudiante adquiere progresivamente los diferentes elementos de significado del concepto y los pone en relación. Esta comprensión emerge de las prácticas significativas del estudiante en la resolución reiterada de problemas específicos del concepto. Finalmente, nuestros resultados sugieren el interés de continuar la investigación sobre el significado y la comprensión de los conceptos estadísticos con muestras mayores de estudiante y con diferentes ítems que tengan en cuenta la naturaleza compleja de la estadística y de la actividad estadística.

ANEXO. CUESTIONARIO

Ítem 1. Un periódico dice que el número medio de hijos por familia en Andalucía es 1.2 hijos por familia. a)Explícanos qué significa para ti esta frase. b)Se han elegido 10 familias andaluzas y el número medio de hijos entre las 10 familias es 1.2 hijos por familia. Los García tienen 4 hijos y los Pérez tienen 1 hijo, ¿cuántos hijos podrían tener las otras 8 familias para que la media de hijos en las diez familias sea 1.2? Justifica tu respuesta.

Ítem 2. Maria y Pedro dedican una media de 8 horas cada fin de semana a hacer deporte. Otros 8 estudiantes dedican cada semana una media de 4 horas a hacer deporte. a)¿Cuál es el número medio de horas que hacen deporte cada fin de semana los 10 estudiantes? b) María y Pedro dedican además 1 hora cada fin de semana a escuchar música y los otros 8 estudiantes, 3 horas. ¿Cuál es el número medio de horas que escuchan música los 10 estudiantes? c)¿Cuál sería el número medio de horas que estos 10 estudiantes dedican, cada fin de semana, entre las dos actividades?

Ítem 3. Cuatro amigos se reúnen para preparar una cena. Cada uno de ellos trajo harina para hacer la masa de las pizzas. Como querían hacer cuatro pizzas del mismo tamaño, los que habían traído más harina regalaron a los que llevaban menos. ¿La cantidad de harina regalada por los que habían traído mucha fue mayor, menor o igual a la recibida por los que habían traído poca? ¿Por qué piensas eso?

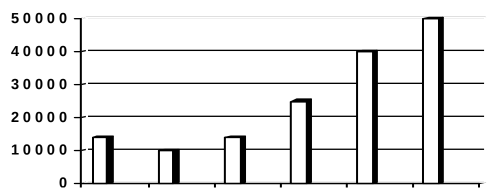
Ítem 4. Tenemos seis números y el más grande es el 5. Sumamos estos números y dividimos la suma por seis. El resultado es 4. ¿Te parece posible? ¿Por qué?

Ítem 5. El peso en kilos de 9 niños es 15, 25, 17, 19, 16, 26, 18, 19, 24. a) ¿Cuál es el peso del niño mediano? b) ¿Cuál es la mediana si incluimos el peso de otro niño que pesa 43 Kg.? c) En este caso, ¿sería la media aritmética un buen representante de los 10 datos? Razona la respuesta.

Ítem 6. Lucía, Juan y Pablo van a una fiesta. Cada uno lleva un cierto número de caramelos. Entre todos llevan una media de 11 caramelos por persona. a) ¿Cuántos caramelos ha llevado cada uno? Lucía _ Juan _ Pablo_ b) ¿Es la única posibilidad? Explica cómo has obtenido tus resultados, c) Un cuarto chico llega a la fiesta y no lleva ningún caramelo. ¿Cuál es ahora la media de caramelos por chico? Explica tu resultado.

Ítem 7. Nueve estudiantes han pesado un objeto en la clase de ciencias, usando la misma escala. Los pesos registrados por cada estudiante (en gramos) se muestran a continuación: 6.2 6.3 6.0 15.2 6.2 6.1 6.5 6.2 6.1 6.2. Los estudiantes quieren determinar con la mayor precisión posible el peso real del objeto. ¿ Qué harías para calcularlo?

Ítem 8. Observa este diagrama de barras que muestra las ventas de bocadillos de la empresa *bocatta* durante los últimos 6 meses del año pasado: a) Da un valor aproximado del número



medio de bocadillos que se vende al mes; b) Da un valor aproximado de la moda; c) Da un valor aproximado de la mediana.

REFERENCIAS

Batanero, C. (2000). Significado y comprensión de las medidas de posición central. *Uno. Revista de Didáctica de las Matemáticas,* 25, 41-58.

Biggs, J. B. & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonom*y. Academic Press: New York.

Cai, J. (1995). Beyond the computational algorithm. Students'unsderstanding of the arithmetic average concept. In L. Meira (Ed.). *Proceedinsg of the 19th PME Conference* (v.3, pp. 144-151). Universidade Federal de Pernambuco, Recife, Brazil.

Carvalho, C. (2001). Interaçao entre pares. *Contributos para a promoçao do desenvolvimiento lógico e do desempenho estatístico no 7º ano de escolaridade.* PhD. University of Lisbon.

Cobo, B. (1998). *Estadísticos de orden en la enseñanza secundaria.* Master's Thesis. University of Granada.

Cobo, B. (2001). Problemas y algoritmos relacionados con la media en los libros de texto de secundaria. *Jornadas Europeas de Enseñanza y Difusión de la Estadística.* Palma de Mallorca: Instituto Balear de Estadística.

Cobo, B. y Batanero, C. (2000). La mediana en la educación secundaria obligatoria:¿ un concepto sencillo. *Uno. Revista de Didáctica de las Matemáticas. Barcelona. Editorial Graó*, (23), 85-96.

Gal, I (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review,* 70(1), 1-25.

Godino, J. D. & Batanero, C. (1994). Significado personal e institucional de los objetos matemáticos (Institutional and personal meaning of mathematical objects). *Recherches en Didactiques des Mathématiques*, 14(3), 325-355.

Godino, J. D. & Batanero, C. (1997) Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. In A. Sierpinska, & J. Kilpatrick (Eds.), *Mathematics Education as a Research Domain: A Search for Identity* (pp. 177-195). Dordrecht: Kluwer.

Batanero, C. y Godino, J. D. (2002). *Estocástica y su didáctica para maestros. Departamento de Didáctica de la Matemática.* Universidad de Granada.

M.E.C. (1992). *Decretos de Enseñanza Secundaria Obligatoria*. Madrid: Ministerio de Educación y Ciencia.

Mevarech, Z. R. (1983). A deep structure model of students'statistical misconceptions. Educational Studies in Mathematics, 14, 415-429.

Navas, F., Batanero, C. y Godino, J. D. (1997). Evaluación de concepciones sobre la noción de promedio en maestros de primaria en formación. Implicaciones para la formación estadística de los futuros profesores. *Actas VII Jornadas LOGSE: Evaluación Educativa*, 301-304.

N.C.T.M. (2000). *Principles and standards for school mathematics.* Reston, VA; N.C.T.M. http://standards.nctm.org/

Pollatsek, A., Lima, S. & Well, A.D. (1981). Concept or Computation: Students'understanding of the mean. *Educational Studies in Mathematics*, 12, 191-204.

Reading, C. (2002). Profiles for statistics understanding. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics.* Cape Town: IASE. CD ROM.

Strauss, S. & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 19 (1), 64-80.

Tormo, C. (1993). *Estudio sobre cuatro propiedades de la media aritmética en alumnos de 12 a 15 años*. Master's thesis. University of Valencia.

Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, v1 (2/3), 11-50.