CONCEPCIONES DE MAESTROS DE PRIMARIA EN FORMACIÓN SOBRE LOS PROMEDIOS

Carmen Batanero, Juan D. Godino y Francisco Navas

Versión ampliada del trabajo publicado en H. Salmerón (Ed.), *VII Jornadas LOGSE: Evaluación Educativa* (pp. 310-304), 1997.

Universidad de Granada.

RESUMEN

En este trabajo analizamos las respuestas de 273 estudiantes de la Facultad de Ciencias de la Educación a un cuestionario escrito que incluye cuatro ítems de opciones múltiples sobre distintos aspectos interpretativos de la media aritmética. La dificultad de las cuestiones ha sido alta -similar a la encontrada en estudiantes de niveles inferiores por otras investigaciones-, particularmente las referidas al tratamiento de los valores atípicos y las relaciones entre la media, la moda y la mediana cuando las distribuciones no son simétricas.Estas dificultades fueron confirmadas en entrevistas posteriores a una muestra reducida de alumnos. Se concluye la necesidad de intensificar el estudio de estas nociones en los planes de formación de profesores, dado el nuevo énfasis sobre el tratamiento de información y el razonamiento estadístico en la reforma de la enseñanza de las matemáticas. Los ítems utilizados podrían también servir para evaluar la comprensión de los promedios en los estudiantes de primaria y secundaria.

1. INTRODUCCIÓN

La reforma curricular promovida por la LOGSE, así como en otros currículos recientes de los paises de nuestro entorno supone un importante reto al sistema educativo, no sólo en los niveles de enseñanza primaria y secundaria, sino también para la formación inicial y permanente de los profesores de las distintas áreas curriculares. En el caso de la formación de los profesores de primaria es preciso contemplar la preparación matemática y didáctica en los nuevos contenidos cuya enseñanza se propone o potencia en la reforma, como es el razonamiento estadístico y el tratamiento de la información (M.E.C., 1992; N.C.T.M., 1989).

Con el fin de poder orientar adecuadamente la enseñanza de este contenido, presentamos en este trabajo los resultados de un estudio sobre las concepciones que los profesores de primaria en formación tienen sobre los promedios. El análisis de las respuestas a un cuestionario escrito aplicado a una muestra de 273 estudiantes de la Facultad de Ciencias de la Educación, junto con entrevistas a algunos de estos alumnos, nos permite mostrar que, a pesar de la simplicidad aparente de estos conceptos, su comprensión por estos estudiantes presenta dificultades similares a las encontradas en sus futuros alumnos. Estos resultados son preocupantes, por el papel que juegan los promedios no sólo dentro de la estadística, sino como fundamento de la formación de los mismos profesores en otras asignaturas de tipo metodológico o incluso de ciencias sociales o humanas, en que con frecuencia se usan estos conceptos.

2. ALGUNOS RESULTADOS PREVIOS

Además de ser conceptos estadísticos básicos, los promedios son imprescindibles en el análisis exploratorio de datos, cuya enseñanza se recomienda en los nuevos currículos de primaria y secundaria. El concepto de media aritmética es aparentemente simple, pero en Batanero y cols (1994) y Tormo (1995) se presenta una síntesis de investigaciones que muestran dificultades procedimentales y conceptuales relacionadas con la media, en alumnos de diferentes edades. Respecto a los errores de cálculo, Pollatsek y cols. (1981) sugieren que un porcentaje importante de estudiantes universitarios calculan incorrectamente la media ponderada y la media simple a partir de una tabla de frecuencias. Asimismo, Mevarech (1983) encontró estudiantes universitarios que atríbuían propiedades inexistentes a la operación de hallar la media aritmética, como la asociatividad, existencia de elemento neutro y elemento inverso. Recientemente Cai (1995) indica que muchos estudiantes de primaria, que son capaces de calcular correctamente la media, no comprenden su algoritmo de cálculo y lo aplican de forma mecánica.

Respecto a la comprensión de los aspectos conceptuales, Russel y Mokros (1990) exploraron las concepciones de los alumnos de primaria sobre los promedios. Sus resultados indican que los niños tienen una idea razonable de la representatividad de la media aunque a veces confunden este concepto con el de moda. Strauss y Bichler (1988) encontraron que los alumnos de primaria tienen dificultades en la comprensión de propiedades como las siguientes:

- La media está situada entre los valores extremos
- La media puede no coincidir con ninguno de los datos
- El valor medio es influenciado por los valores de cada uno de los datos.
- Hay que tener en cuenta los valores nulos en el cálculo de la media.

Como se sabe, la media es un valor "típico " o "representativo" de los datos. Campbell (1974) observa que, debido a ello, se tiende a situar la media en el centro del recorrido de la distribución, propiedad que es cierta para distribuciones simétricas. Pero cuando la distribución es muy asimétrica la media se desplaza hacia uno de los extremos y la moda o la mediana serían un valor más representativo del conjunto de datos.

Gattuso (1996) encontró que la dificultad de algunos problemas, como es el cálculo de la media ponderada y el error consistente en suprimir los valores nulos en el cálculo de la media, continúan en la universidad. Aunque, con la edad e instrucción mejora el empleo del algoritmo de cálculo y los alumnos usan con mayor frecuencia notación y métodos algebraicos, los problemas que requirieron un conocimiento conceptual más profundo fueron resuelto por un porcentaje mayor de alumnos de secundaria, que por universitarios, debido a que los primeros emplearon estrategias intuitivas y los segundos se concentraron en fórmulas algebraicas o definiciones formales.

3. DESCRIPCION DE LA INVESTIGACION

Nuestra investigación tuvo como objetivo explorar los conocimientos y dificultades de comprensión de nuestros alumnos en la Facultad de Educación (futuros profesores de primaria y estudiantes de primeros cursos de pedagogía) sobre los contenidos estocásticos básicos que se incluyen en los nuevos diseños curriculares, con el fin de detectar los puntos en que sería necesario completar su formación. Se realizó en el marco de un proyecto de colaboración conjunto con la Universidad de Minnesota para realizar un estudio comparativo de las concepciones previas de estos estudiantes en diferentes paises.

El cuestionario (elaborado por Konold y Garfield, 1993) se componía de una escala de actitudes y 25 ítems de evaluación de los conocimientos sobre nociones estadísticas y probabilísticas elementales. En este artículo presentamos los resultados de los ítems referidos a la comprensión de los promedios.

Muestra

Los datos se tomaron de dos grupos de alumnos de la Facultad de Ciencias de la Educación de la Universidad de Granada: El primero de ellos, con un total de 132 alumnos de primer curso de Magisterio (especialidad de primaria). La prueba realizada al primer grupo fue previa a la enseñanza del tema de "Introducción a la Estadística", al cual se dedica un crédito (10 h.) en la Facultad de Ciencias de la Educación de Granada. Sin embargo, estos alumnos habían estudiado estas nociones, al menos en 7º curso de E.G.B. y 1º de Bachillerato.

La segunda muestra estuvo formada por un grupo de 141 alumnos de 2º curso de Pedagogía, que habían realizado un curso de metodología de investigación, en el que se incluían algunos contenidos de estadística descriptiva. Algunos de estos alumnos (26) habían elegido voluntariamente un curso de libre configuración sobre *Análisis de Datos y su Didáctica* impartido por nosotros mismos, por lo que la evaluación efectuada también tenía como objeto evaluar sus conocimientos previos sobre el tema. En total participaron tres grupos de alumnos en la primera muestra y dos en la segunda.

Análisis 'a priori' de los ítems

El coeficiente de fiabilidad del cuestionario completo, dentro de nuestra muestra fue α=0.83.

A continuación analizamos los ítems referidos a promedios para informar de la validez del instrumento, respecto a este contenido.

Item 1: Un objeto pequeño se pesó con un mismo instrumento, separadamente por nueve estudiantes en una clase de ciencias. Los pesos obtenidos por cada estudiante (en gramos) se muestran a continuación:

6.2 6.0 6.0 15.3 6.1 6.3 6.2 6.15 6.2

Los estudiantes quieren determinar con la mayor precisión posible el peso real del objeto. ¿Cuál de los siguientes métodos les recomendarías usar?

___ a) Usar el número más común, que es 6.2

___ b) Usar 6.15, puesto que es el peso más preciso

___ c) Sumar los 9 números y dividir la suma por 9

___ d) Desechar el valor 15.3; sumar los otros 8 números y dividir por 8.

Se pide a los estudiantes la mejor estimación del verdadero valor de la medida de una cantidad, en este caso el peso de un objeto. Aunque, en el caso general, la mejor estimación viene dada por la media aritmética, su adecuación a un problema particular debe ser valorada por el alumno. En este ítem se introduce un valor atípico que es preciso descartar antes de proceder al cálculo de la media aritmética, pues ésta es muy sensible a los valores extremos. El alumno que elige la respuesta correcta d), debe tener en cuenta estas dos propiedades. Se trata de discriminar entre el simple conocimiento algorítmico de la fórmula de cálculo, que llevaría al distractor c), de la comprensión relacional del concepto. El distractor a) sugiere como respuesta la moda, que no tiene la propiedad requerida (mejor estimador). El distractor c) se refiere a otro uso diferente en matemáticas de la palabra "precisión", que indica el número de cifras decimales de los datos.

Item 2: Una profesora quiere cambiar la disposición de los asientos en su clase, en la esperanza de que ello incremente el número de preguntas que hacen sus alumnos. Primero, decide ver cuántas preguntas hicieron los estudiantes con la colocación actual de los asientos. Un registro del número de preguntas hechas por sus 8 estudiantes durante una clase se muestra a continuación

Iniciales del alumno								
	A.A	R.F.	A.G.	J.G.	C.K.	N.K.	J.L	A.W
Nº preguntas	0	5	3	22	3	2	1	2

La profesora quiere resumir estos datos, calculando el número típico de preguntas hechas ese dia. ¿Cuál de los siguientes métodos le recomendarias que usara?

___ a) Usar el número más común, que es el 2.

___ b) Sumar los 8 números y dividir por 8.

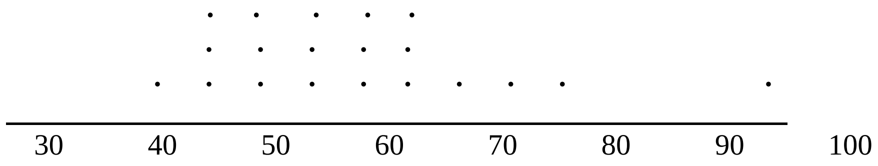
___ c) Descartar el 22, sumar los otros 7 números y dividir por 7.

___ d) Descatar el 0, sumar los otros 7 números y dividir por 7.

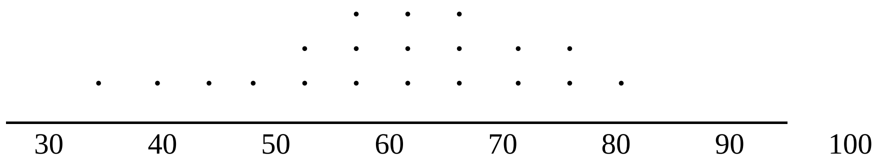
El problema planteado es similar al del ítem 1, si bien la variable es discreta. Se incorpora un valor nulo en los datos, para tener en cuenta la creencia de algunos estudiantes en que el cero no cambia el valor de la media. Así mismo, se incluye un valor atípico, que en este caso no debe descartarse, porque el contexto es diferente. En lugar de pedirse la mejor estimación de un valor, se pide por un resumen de los datos, que debe incluir el valor atípico. La distribución, además, tiene una gran variabilidad, mientras que en el ítem 1 es mucho más homogénea.

El distractor a) introduce la moda como preferible a la media para representar estos datos; el c) propone descartar el valor atípico y el d) descartar el valor nulo.

Item 3: Veinte estudiantes universitarios participaron en un estudio sobre el efecto del sueño sobre las puntuaciones en los exámenes. Diez de estudiantes voluntariamente estuvieron despiertos estudiante toda la noche anterior al examen (grupo que no durmió). Los otros 10 estudiantes (el grupo contro) se acostaron a las 11 la nocha anterior al examen. Las puntuaciones en el examen se muestran en los gráficos siguientes. Cada punto representa la puntuación de un estudiante particular. Por ejemplo, las dos equis encima del número 80 en el gráfico inferior indican que dos estudiantes en el grupo control tuvieron una puntuación de 80 en el examen.



Puntuaciones en el examen del grupo que no durmió



Puntuaciones en el examen del grupo que durmió

Examina los dos gráficos con cuidado. Luego escoge entre las 6 posibles conclusiones que se listan a continuación aquella con la que estés más de acuerdo.

___ a) El grupo que no durmió lo hizo mejor porque ninguno de estos estudianets puntuó por debajo de 40 y la máxima puntuación fue obtenida por un estudiante de ese grupo

___ b) El grupo que no durmió lo hizo mejor porque su promedio parece ser un poco más alto que el promedio del grupo que no durmió.

___ c) No hay diferencia entre los dos grupos, porque hay un solapamiento considerable en las puntuaciones de los dos grupos.

___ d) No hay diferencia entre los dos grupos, porque la diferencia entre sus promedios es pequeña, comparada con la cantidad de variación de sus puntuaciones.

___ e) El grupo de control lo hizo mejor porque hubo en ese grupo más estudiantes que puntuaron 80 o por encima.

___ f) El grupo de control lo hizo mejor, porque su promedio parece ser un poco mayor que el promedio del grupo control.

Se estudia la capacidad de estimar un valor central (promedio) a partir de la representación gráfica, para comparar dos distribuciones de frecuencias (puntuaciones en un examen), con valores atípicos. La interpretación del gráfico corresponde, en la terminología de Curcio (1987), al nivel de "leer más allá de los datos", ya que exige la lectura directa de los mismos, su comparación, estimación de los promedios (que no están directamente representados) y obtención de conclusiones. La comparación de los dos grupos debe tener en cuenta, no sólo la media, sino también la dispersión de los datos.

La respuesta correcta es la f) porque, si se descartan los valores atípicos, el promedio de la segunda muestra es superior al de la primera. Los distractores incluídos son los siguientes:

a) Valorar las diferencias basándose sólo en los máximos y mínimos de las distribuciones.
b) Valorar las diferencias basándose sólo en los promedios, sin descartar los valores atípicos. Puesto que éstos afectan bastante al valor de las medias se obtendría una conclusión errónea.
c) Tener en cuenta sólo la dispersión de los datos (en particular el recorrido que es parecido) y no los promedios.

d) Este distractor es parcialmente cierto, aunque como la dispersión es similar en los dos conjuntos de datos existen diferencias de promedios, aunque sea pequeña.
e) Juzgar la diferencia entre los grupos usando sólo parte de la distribución.

Item 4: El comité escolar de una pequeña ciudad quiere determinar la media del número de niños por familia en su ciudad. Dividen el número total de niños de la ciudad por 50, que es el número de familias. ¿Cuál de las siguientes frases debe ser cierta si la media del número de niños por familia es 2.2?
___ a) La mitad de las familias de la ciudad tienen más de 2 niños.
___ b) En la ciudad más familias tienen 3 niños que 2 niños.
___ c) Hay un total de 110 niños en la ciudad.
___ d) Hay 2.2 niños por adulto en la ciudad.
___ e) El número más común de niños en una familia es 2.

Dada la media de la variable 'número de niños por familia' en un pueblo y conociendo el número de familias, se presentan una serie de afirmaciones sobre los estadísticos de la distribución, cuya forma no se especifica. Sin embargo, por la experiencia, sabemos que está inferiormente acotada por el cero y que es asimétrica. Mientras que en las distribuciones simétricas los valores de la media, mediana y moda coinciden, esto no tiene porqué ocurrir en las distribuciones asimétricas. Se requiere que el alumno conozca la relación entre estas tres características en distribuciones simétricas y asimétricas. El ítem evalúa también la comprensión del algoritmo de cálculo, ya que la solución correcta c) requiere recordar y comprender la fórmula de cálculo de la media y operar con los datos proporcionados.

- El distractor a) sugiere que, para este conjunto de datos, el valor de la mediana es muy próximo a la media.
- El distractor b) sugiere que la mediana está por encima de 3, lo cual es falso, porque en las distribuciones con asímetría a la derecha, la mediana es inferior a la media.
- El distractor d) confunde la variable "número de familias" y "número de adultos".
- El distractor e) sugiere que, para este conjunto de datos, el valor de la moda es muy próximo a la media. Tanto este distractor como el a) permiten poner de manifiesto una idea o creencia de los estudiantes sobre la forma "acampanada" de las distribuciones de frecuencias.

4. RESULTADOS Y DISCUSION

Puesto que los resultados de los dos grupos de alumnos mostraron una alta similaridad, hemos optado por presentar conjuntamente los resultados de las dos muestras. En la tabla 1 presentamos los resultados, en los que podemos ver que la dificultad ha sido alta, aunque variable. Mientras que en los dos primeros ítems más de la mitad de los alumnos dan una respuesta correcta, esta no llega al 25% en el último, lo que muestra, la falta de comprensión del algoritmo de cálculo de la media, el desconocimiento de los alumnos de la relación entre media, mediana y moda en distribuciones no simétricas o la creencia en que todas las distribuciones son simétricas.

Tabla 1: Frecuencias y porcentajes de respuestas correctas a los ítems

Item	Correctos	Incorrecto	Principal Distractor	
1	141 (51.5)	132 (48.5)	c: No desechar el valor atípico	107 (39.2)
2	61 (22.3)	212 (77.6)	b: No descartar el valor atípico	159 (58.2)
3	139 (50.9)	134 (49.1)	a/b: No descartar los valores atípicos	47 (17.2)
			c/d: Sobrevalorar la dispersión	61 (22.3)
4	72 (26.4)	210 (73.4)	e: Suponer una distribución simétrica	141 (51.6)

No obstante, el porcentaje de respuestas incorrectas es alarmante en todos los ítems, especialmente en los futuros profesores que posiblemente deban enseñar estos temas, y teniendo en cuenta el escaso tiempo dedicado a la formación estadística en los planes de Magisterio.

Análisis de distractores

En todos los ítems ha habido un distractor dominante, aunque también otros distractores han tenido un porcentaje importante de respuestas. A continuación realizamos una clasificación de estos distratores, según las dificultades de comprensión que podrían estar subyacentes en los alumnos que los han elegido:

Tratamiento de valores atípicos:

En los tres primeros ítems el distractor dominante se refiere a los valores atípicos y a la discriminación de los casos en los que hay o no que incluirlos en el análisis de los datos. En concreto, los alumnos han elegido los siguientes distractores:

34.1% de respuestas al distractor c) del ítem 1, que sugiere calcular la media sin desechar el valor atípico. Sin embargo, en este contexto, el valor atípico apunta a un valor de medición y perturba notablemente la estimación del valor verdadero del peso del objeto.

29.5 al distractor c) del ítem 2, que consiste en calcular la media, desechando el valor atípico 39. Al contrario que en el caso anterior, debido a la variabilidad de los datos y al efecto del valor atípico sobre la media, debe incluírse en el análisis, pues su supresión afectaría notablemente al valor "representativo" del conjunto de datos.

22.6% al distractor b) del ítem 3, que sugiere no tener en cuenta los valores atípicos en la comparación de dos distribuciones. Puesto que, en este caso, los valore atípicos están situados de tal forma que hacen subir el valor de la media y bajarlo en la otra, desvirtúan por completo las diferencias existentes entre los dos conjuntos de datos y es necesario suprimirlos del análisis.

Como vemos de los resultados de estos tres ítems, el tratamiento de los valores ´tipicos no ha sido fácil para nuestros estudiantes. Ello apunta a la descontextualización de la enseñanza de la estadística recibida en sus estudios previos y la falta de un conocimiento funcional de lo aprendido.

Posiciones relativas de media, mediana y moda:

Otro punto en el que los alumnos han tenido dificultad alta ha sido el conocimiento de las posiciones relativas entre media, mediana y moda en distribuciones no simétricas, que se manifiesta en un 57.4% de respuestas al distractor e) del ítem 4. Esta relación no es fácil de comprender, porque no se deduce claramente del algorímo de cálculo. En la respuesta a este distractor pueden intervenir también la creencia de los alumnos en que todas las distribuciones deben ser simétricas debido, posiblemente a la falta de contextos realistas en la enseñanza.

La media como mejor representante de los datos cuantitativos, frente a la moda o mediana:

La discriminación de cúando media, mediana o moda son preferibles como mejor representante de un conjunto de datos ha tenido también algunas dificultades, ya que un 9.1%de alumnos elige el distractor a) del Item 1, que sugiere usar la moda, la cual no tiene la propiedad requerida (mejor estimador de un conjunto de datos).

Tratamiento de valores nulos en el cálculo de la media:

En unos pocos casos, hemos detectado el error consistente en calcular la media, desechando el 0 (10.9% de respuestas al distractor d) del ítem 2), descrito por Strauss y Bichler (1988).

Comparación inadecuada de distribuciones:

La comparación de dos distribuciones es una de las tareas básicas en el enfoque exploratorio del análisis de datos, recomendado en los diseños curriculares de primaria y secundaria. Sin embargo, los alumnos han mostrado con frecuencia estrategias incorrectas al realizar esta comparación, lo que sugiere la necesidad de completar su formación en técnicas elementales de análisis exploratorio de datos. A continuación presentamos una lista de las estrategias incorrectas detectadas:

- No valorar adecuadamente las diferencias de promedios, con respecto a la dispersión (8.6% de respuestas al distractor d) del ítem 3). A este respecto Reading y Pegg (1996) encontraron que los alumnos tienden a reducir los datos usando con mayor frecuencia las medidas de valor central que las de dispersión y presentan más dificultades en este segundo concepto, lo que puede explicar estas respuestas.
- Juzgar la diferencia entre los grupos usando sólo parte de la distribución (12.9% de respuestas al distractor d) del ítem 3). Esta misma estrategia incorrecta es encontrada en Estepa y Sánchez (1996) en el estudio de las estrategias de los alumnos en la comparación de muestras, quienes la interpretan como indicación de que los alumnos tienen una concepción *local* de la asociación entre variables, consistente en creer que ésta puede deducirse con analizar sólo una parte de los datos.

5. ANALISIS DE ENTREVISTAS

Con objeto de complementar nuestros datos, se realizaron entrevistas en profundidad a seis de los alumnos participantes, elegidos porque sus respuestas al cuestionario indicaban la falta de comprensión de algunos de las propiedades que hemos analizado sobre los promedios y que colaboraron con gran interés en nuestra investigación. Cada uno de ellos fué entrevistado individualmente sobre sus respuestas al cuestionario, así como sobre las propiedades que atribuían a media, mediana y moda, y sobre las diferencias entre estos conceptos. Aunque se preparó previamanete un guión orientativo sobre la entrevista, esta no fue totalmente estructurada y el entrevistador varió ocasionalmente las preguntas, para dejar mayor libertad de expresar sus ideas al alumno. Agradecemos a estos alumnos, así como al resto de alumnos y profesores participantes su valiosa colaboración, que nos ha permitido profundizar en el conocimiento de las dificultades de los estudiantes sobre los promedios.

A continuación analizamos algunos resultados de estas entrevistas:

Influencia de los valores atípicos en el cálculo de la media

Los alumnos entrevistados no son siempre conscientes del efecto de los valores atípicos sobre la media, o al menos no son capaces de aplicar este conocimiento a la resolución de los problemas propuestos, como se muestra en el siguiente ejemplo:

> *Investigador: En la pregunta 1 tu das como opción correcta la c) que sería escoger los nueve números y dividirlos por nueve ¿no? ...¿Qué lectura haces tú de esos nueve datos que tenemos?*
> *Alumno 1. No sé, yo veo que todos son más o menos parecidos ¿no?, que sólo hay uno que sobresale es lo que veo yo ahí.*
> *Investigador: De acuerdo, y... ¿Crees que podría ser un error del alumno?*
> *Alumno: Yo creo que podría ser un error; porque los otros números están entre 6 y 6,2. !Es que sobresale más del doble!*
> *Investigador: Y, entonces, la opción d (*descartar el valor atípico), *¿Porqué no la cogistes?*
> *Alumno: Porque era el valor que más sobresalía, que era diferente a los demás*

Tanto en este problema como en el siguiente, el alumno considera que debe tener en cuenta el valor atípico en el cálculo de la media, precisamente por ser diferente a los demás. No es consciente que este valor hará subir la media global, y que esta ya no será, en consecuencia, un bue representante del conjunto de datos ya que sesgará la estimación pedida en el problema. Razonamientos semejantes fueron dados por otros alumnos en las entrevistas.

Otra concepción curiosa es la mostrada por otro alumno, quien, en el problema 3, muestra que comprende el efecto de los valores atípicos sobre los promedios, pero aún así los tiene en cuenta en el cálculo. En dicho problema cree que puede compensar los valores atípicos de uno y otro grupo, siendo así que el efecto de los mismos es precisamente encubrir las diferencias existentes puesto que en el primer grupo el valor atípico sube la media y en el segundo la baja. Sin embargo, cuando le proponemos un valor atípico excesivamente alejados de los datos, sugiere rechazarlo del cálculo de los promedios.

> *Investigador: Antes has comentado que el valor atípico del primer grupo que es cercano al cien y del del segundo que es cercano al cuarenta se compensan uno con el otro. Yo te haría la*

siguiente pregunta, ¿cómo afectaría tu respuesta si el valor atípico del grupo b lo acercamos a veinte por ejemplo? ¿Captas la idea?
Alumno: Si, claro. Entonces diría que están igualados los dos grupos... Porque claro está que el que está en el veinte no lo voy a poder anular con el que está entre 90 y 100 en el grupo A. Y esa pequeña diferencia quizás taparía un poco del grupo a los defectos de promedio que tiene aquí.
Investigador: Y si lo bajamos un poco más, lo acercamos al diez o incluso al cero
Alumno: Entonces quitaría ese dato.

Comparaciones basadas sólo en parte de la distribución

En algunas ocasiones los alumnos sólo usan una parte de los datos para comparar las dos distribuciones en el problema 3. Esta misma dificultad ha sido señalada por Konold y cols. (1996), quienes consideran que ello se debe a que los estudiantes no son conscientes de que el objeto de estudio de la estadística no son los datos aislados sino las distribuciones de ciertas variables en los datos:

Investigador: Entonces has cogido la opción e) en el problema 3, que comenta que el grupo que lo hizo mejor fué el de control. ¿Por qué crees que es la respuesta correcta?
Alumno: Porque estos alumnos son los que sacan más nota, que han estado durmiendoo por la noche y han sacado notas más altas

Y más adelante, contunúa efectuándo comparaciones usando sólo una parte de la distribución:

Alumno: La diferencia de aprobados, es decir, por encima de cuncuenta es mayor... Porque de sesenta a setenta hay cuatro, y aquí hay menos, Por lo tanto sacan notas más altas, aunque haya suspendido uno.

Posiciones relativas de media, mediana y moda:

De las entrevistas realizadas se deduce que estos alumnos consideran que los valores de media mediana y moda deberían ser próximos entre sí, es decir, no son sensibles al efecto de la asimetría sobre los promedios, como se pone de manifiesto en las respuestas de varios de los estudiantes:

Investigador: Tú en el ítem 4 elegiste la opción e) ¿por qué?
Alumno: No sé; pues hice cálculos y me salía a más o menos dos hijos por familia,.. bueno, que era el más común para que saliera esa media de 2,2.

Este es un tipo de respuesta dado también por otros alumnos que piensan que, puesto que la media es 2,2 el valor más frecuente debe ser el entero más próximo a este valor. Vemos este tipo de razonamienteo que se mantiene estable en otro alumno, cuando al preguntarle cual sería el número más común de niños por familia si variamos la media de la distribución nos da la siguiente respuesta:

Investigador: Si en un contexto similar, sin embargo, nos hubiera dado que la media era de tres hijos por familia..¿cuál sería el número más común de niños por familia?.
Alumno: Pues hubiese puesto que tres.

El razonamiento de la siguiente alumna de magisterio también muestra este error. Esta alumna había elegido la opción b) en el ítem 4, pero cambia su respuesta al preguntarlo el motivo de su elección, como vemos a continuación:

Investigador: Aquí has tomado la opción b ¿Por qué la has elegido?
Alumno: Porque al ver que la media es 2,2 y que una familia no puede tener 0,2 niños,... O sea, las familias tienen 2 y 3 niños.... !Está mál! Me he equivocado, porque al ser la media 2,2 es al revés, hay más familias que tiene 2 niños que familias que tienen 3 niños.

Definiciones de media mediana y moda

Finalmente y, aunque algunos alumnos dieron las definiciones correctas de estos tres promedios, también hemos observado que en otros casos no se recuerdan estas definiciones o bien se confunden los conceptos entre sí. El siguiente ejemplo es un caso prototipico:

Investigador: Ya, para acabar, me gustaría que nos dieras la definición de media, moda y mediana:
Alumno: Pues la media es,... el promedio,.... por ejemplo, de un valor... siempre es la mitad.
Investigador: ¿Y la mediana?
Alumno: No se´, es que yo le veo un poco de relación entre media y mediana. Mediana es la mitad de algo, y la media puede ser... Es que prácticamente lo encuentro lo mismo.
Investigador ¿Y la moda?
Alumno: La moda... La mediana sería el valor que más se repite,... la media sería la suma de todos, dividido por dos. La moda,... !Ya me estoy haciendo un lio!

6. CONCLUSIONES

Nuestros resultados muestran la existencia de errores conceptuales y dificultad de aplicación práctica de los conocimientos sobre los promedios en nuestros estudiantes., es decur, apuntan a la necesidad de potenciar los contenidos estadísticos en la formación de profesores. Como explicación del alto porcentaje de errores en los ítems analizados podemos indicar, en primer lugar, el escaso o nulo tratamiento que se hace en la enseñanza primaria y secundaria del tratamiento adecuado de los valores atípicos. La enseñanza de los promedios se centra habitualmente en la presentación de los algoritmos y fórmulas y su aplicación a casos estereotipados. Esta aproximación no permite que los alumnos comprendan el significado integral del concepto. Por el contrario, la interpretación de los resultados y la reflexión sobre las condiciones de aplicación de los procedimientos estadísticos requieren una atención preferente.

Hay que señalar, por otro lado, que la drástica reducción de las asignaturas de matemáticas y su didáctica en los planes de formación de las diferentes especialidades de nuestra facultad hace inviable en la práctica el poder hacer un estudio profundo de estos contenidos, con el consiguiente perjuicio para la posterior comprensión de otras materias en que son utilizados como instrumentos indispensables. Esperamos que la revisión de los planes de enseñanza potencie la preparación estadística de los profesores. Una solución complementaria es motivar a los alumnos a completar su formación estadística mediante cursos optativos o de libre configuración que incluyan esta materia.

REFERENCIAS

Barr, G. (1980). Some student's ideas on the median and mode. *Teaching Statistics,* 2, 38-81.

Cai, J. (1995). Beyond the computational algorithm. Students' understanding of the arithmetic average concept. En L. Meira y D. Carraher (Eds.), *Proceedings of the XIX Conference on the Psychology of Mathematics Education* (v. 3, pp. 144-151). Universidad de Pernambuco.

Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R. y Holmes, P. (1994). Errors and difficulties in understanding statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.

Campbell, S. K. (1974). *Flaws and fallacies in statistical thinking*. New Jersey: Prentice-Hall.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education,* 18(5), 382-393.

Estepa, A. y Sánchez, F. T. (1996). Association judgments in the comparison of two samples. En L, Puig y A. Gutiérrez (Eds.), *Proceedings of the XX Conference on the Psychology of Mathematics Education* (v.2, pp. 345-352). Universidad de Valencia.

Gattuso, L., y Mari, C. (1996). Development of concepts of the arithmetic average from high school to university. En L, Puig y A. Gutiérrez (Eds.), *Proceedings of the XX Conference on the Psychology of Mathematics Education* (v2, pp. 410-408). Universidad de Valencia.

Konold, C. y Garfield, J (1993). Statistical reasoning assesment, Part 1: Intuitive thinking. University of Massachusetts: Scientific Reasoning Institute.

Konold, C., Pollatsek, A., Well, A. y Gagnon, A. (1996). Students analysing data: Research of critical barriers. En J. B. Garfield y G. Burrill (Ed.), *Research on the Role of Technology in Teaching and Learning Statistics* (pp. 163-182). IASE.

M.E.C.(1992). *Matemática 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.

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

Reading, C. y Pegg, J. (1996). Exploring understanding of data reduction. En L. Puig y A. Gutiérrez (Eds.), *Proceedings of the XX Conference on the Psychology of Mathematics Education* (v4, pp. 187-194). Universidad de Valencia.

Russell, S. J. y Mokros, J. R. (1991). What's typical?: Children's ideas about average. En D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 307-313). Voorburg, The Netherlands: International Statistical Institute.

Strauss, S. y 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. (1995). Dificultades del alumnado respecto a la media aritmética. *UNO*, 5, 29-36.