SIGNIFICADO Y COMPRENSIÓN DE LAS MEDIDAS DE POSICIÓN CENTRAL

Carmen Batanero
Departamento de Didáctica de la Matemática, Universidad de Granada
UNO, 2000, 25, 41-58

Los nuevos diseños curriculares incorporan la enseñanza de la estadística en la escuela primaria y secundaria enfatizando el enfoque exploratorio y el trabajo de los alumnos con proyectos interdisciplinares abiertos. Para afrontar con éxito esta propuesta, el profesor debe ser consciente de la complejidad de los conceptos estadísticos, incluso los "elementales" cuyo significado debe construirse progresivamente. Como ejemplo, analizamos los componentes del significado de las medidas de posición central y describimos las dificultades en su comprensión, que, respecto a estos componentes se han puesto de manifiesto en las investigaciones en educación estadística.

1. Introducción

En la actualidad la estadística se ha incorporado, de forma generalizada, al currículo de matemáticas de la enseñanza primaria y secundaria, debido al uso frecuente de datos y conceptos estadísticos en la vida cotidiana, así como en otras disciplinas que debe cursar el alumno, a la necesidad de un conocimiento básico de estadística en muchas profesiones y a su papel en el desarrollo de un razonamiento crítico. El reconocimiento de estas razones y la consecuente incorporación de la estadística al currículo escolar se debe, en gran medida, al trabajo desarrollado desde el *ISI* (International Statistical Institute), primero por el Comité de Educación y desde 1991 por *IASE*, la International Association for Statistical Education, que a lo largo de tres décadas han promovido congresos y publicaciones específicas orientadas a la introducción de la estadística en la escuela (Batanero, 2000).

Ayudar a los niños y jóvenes a comprender progresivamente las ideas estocásticas fundamentales no es una tarea sencilla, puesto que es necesario adaptar estas ideas a sus capacidades cognitivas y diseñar situaciones didácticas que propicien el aprendizaje significativo. La estadística es enseñada, tradicionalmente, como parte de la asignatura de matemáticas por el profesor de esta materia. Nos encontramos con la paradoja de pedir a estos profesores que impartan un nuevo contenido, para el que no todos han tenido una formación didáctica específica, porque la didáctica de la estadística no está aún suficientemente desarrollada. Mientras que la estadística como ciencia, está en un periodo de notable expansión, el número de investigaciones sobre la enseñanza de la estadística es aún escaso, y sólo estamos comenzando a conocer las principales dificultades de los alumnos en los conceptos más importantes. Es también preciso experimentar y evaluar métodos de enseñanza adaptados a la naturaleza específica de la estadística, a la que no siempre se pueden transferir los principios generales de la enseñanza de las matemáticas.

Un problema particular es que la investigación sobre educación estadística se está llevando a cabo en áreas muy diversas (estadística, psicología, educación matemática y profesores de estadística en diferentes áreas de conocimiento), con diferentes tipos de alumnos, metodología y marcos teóricos. Es preciso, por tanto, realizar una labor de síntesis de estos trabajos, y de posterior difusión entre los profesores, que son los que tienen finalmente la posibilidad y responsabilidad de la formación estadística de los escolares. En este trabajo tratamos de contribuir a esta labor, analizando las dificultades que los alumnos pueden encontrar en las medidas de tendencia central. Nos apoyamos en un modelo teórico que permite sintetizar estas investigaciones, explicar las dificultades y proporcionar al profesor criterios en la organización de la enseñanza del tema. En Batanero y cols. (1994) se describen errores y dificultades de los estudiantes en otros conceptos estadísticos elementales.

2. Significado de las medidas de posición central

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, por consiguiente, intimamente ligado a cómo se concibe el propio conocimiento matemático. Los términos y expresiones matemáticas denotan entidades abstractas cuya naturaleza y origen tenemos que explicitar para poder elaborar una teoría útil y efectiva sobre qué entendemos por comprender tales objetos. 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).

En Godino y Batanero (1994) proponemos un marco teórico sobre el significado de un objeto matemático, partiendo como noción primitiva de la situación-problemática y resaltando la génesis personal e institucional del conocimiento matemático. En lo que sigue, analizaremos el campo de problemas y de actividades del que emerge progresivamente el objeto matemático designado con el término "media", y posteriormente otras medidas de tendencia central. Consideremos el siguiente problema:

> *P1.* Un objeto pequeño se pesa con un mismo instrumento por ocho estudiantes de una clase, obteniéndose los siguientes valores en gramos: 6'2, 6'0, 6'0, 6'3, 6'1, 6'23, 6'15, 6'2 ¿Cuál sería la mejor estimación del peso real del objeto?

Si planteamos este problema a nuestros alumnos, la mayoría sumará los valores y dividirá por ocho para obtener el valor 6'1475. Este es un ejemplo particular de una clase de problemas: *estimación de una cantidad desconocida, en presencia de errores de medida*. En muchas situaciones necesitamos medir una cantidad X desconocida de una cierta magnitud. Pero debido a la imperfección de nuestros instrumentos, en mediciones sucesivas obtenemos distintos números como medidas de X. No tenemos ninguna razón para pensar que el verdadero valor esté más cercano a uno u otro de los datos obtenidos. ¿Cómo determinar, a partir de un conjunto de medidas $x_1, x_2, \ldots, x_n$ la mejor estimación posible del verdadero valor X desconocido? Según Plackett (1970), los astrónomos de Babilonia resolvieron el problema calculando la suma total de las observaciones y dividiendo por el número de datos y esta práctica se ha conservado hasta nuestros días.

Para enunciar y resolver el problema P1 necesitamos representaciones simbólicas de los objetos matemáticos abstractos (números, operaciones, ...). Por ejemplo, actualmente usamos la expresión (1) para resolver el problema P1 en su enunciado general, donde los distintos símbolos representan el número de datos, los valores obtenidos en las distintas mediciones, su suma, la división y el resultado obtenido:

$$\overline{x} = (x_1 + x_2 + \ldots + x_n)/n \qquad (1)$$

Es característico de la actividad matemática extender las soluciones a otros ejemplos, diferentes de la situación concreta particular. Nosotros podríamos generalizar la expresión (1) para un valor arbitrario *n*, posteriormente a un número infinito de valores, en variables discretas o continuas:

(2) $$E(X) = \sum x_i p_i = \mu$$

(3) $$E(X) = \sum xf(x)dx$$

Sumar un conjunto dado de valores y dividir por el número de valores o escribir las expresiones anteriores son ejemplos de *prácticas* matemáticas, es decir de acciones llevadas a cabo para resolver problemas matemáticos, comunicar a otros la solución, mostrar que la solución es correcta y generalizarla a otros contextos y problemas. Aunque en cada problema concreto de estimación de la magnitud de interés, el instrumento de medición, el número de medidas tomadas y los valores concretos obtenidos varían, la expresión (1) es aplicable de forma general para el cálculo de la mejor estimación del valor desconocido. Esta práctica es el germen de la emergencia progresiva del concepto que hoy conocemos como "media aritmética", primeramente como útil implícito en la solución de problemas prácticos, más tarde como objeto de estudio en sí mismo. El estudio y caracterización de sus propiedades llevó progresivamente a la aplicación del concepto en la solución de otras situaciones problemáticas como las siguientes:

P2. Unos niños llevan a clase caramelos. Andrés lleva 5, María 8, José 6, Carmen 1 y Daniel no lleva ninguno. ¿Cómo repartir los caramelos de forma equitativa?

Cuando se necesita obtener una cantidad equitativa a repartir para conseguir una distribución uniforme, como en el ejemplo, se toma la media aritmética. Problemas semejantes serían obtener la "renta per cápita", la velocidad media durante un viaje o la calificación final en un examen compuesto de varios exámenes parciales.

P3. Al medir la altura en cm. que pueden saltar un grupo de escolares, antes y después de haber efectuado un cierto entrenamiento deportivo, se obtuvieron los valores siguientes. ¿Piensas que el entrenamiento es efectivo?

	Altura saltada en cm.									
Alumno	Ana	Bea	Carol	Diana	Elena	Fanny	Gia	Hilda	Ines	Juana
Antes del entrenamiento	115	112	107	119	115	138	126	105	104	115
Después del entrenamiento	128	115	106	128	122	145	132	109	102	117

Otra aplicación de la media es servir de elemento representativo de un conjunto de valores dados x_i, cuya distribución es aproximadamente simétrica. En el ejemplo P3 usaríamos la altura media saltada antes y después del entrenamiento para ver si éste ha producido algún efecto. Para representar un conjunto de datos se toma la media por sus propiedades de localización central, por ser "centro de gravedad" del espacio de valores muestrales o poblacionales. Si la distribución es muy asimétrica, el valor más frecuente (Moda) o el valor central en el conjunto de datos ordenados (Mediana) podría ser más representativo. Vemos que cuando añadimos condiciones a un campo de problemas surgen conceptos relacionados con el de interés con el cual guardan diferencias y semejanzas, que es necesario investigar. De los primitivos problemas extramatemáticos, pasamos posteriormente a problemas internos a la misma matemática, como estudiar las diferentes propiedades de las medidas de posición central.

P4. *La altura media de los alumnos de un colegio es 1'40. Si extraemos una muestra aleatoria de 5 estudiantes y resulta que la altura de los 4 primeros es de 1'38, 1'42, 1'60,* *1'40. ¿Cuál sería la altura más probable del quinto estudiante?*

En otras ocasiones se necesita conocer el valor que se obtendrá con mayor probabilidad al tomar un elemento al azar de una población. Por ejemplo, al predecir la esperanza de vida o el beneficio esperado en una inversión en bolsa, se toma la media de la variable en la población como predicción, como valor esperado, por sus propiedades muestrales derivadas del teorema central del límite. Del concepto de valor esperado se derivan muchos modelos de predicción, como los distintos tipos de regresión. Así, cuando predecimos el peso de una persona en función de su altura, usamos el peso promedio de todas las personas que en la población tienen la altura dada.

Problemas como los P1 a P4 y otros problemas, primero prácticos más tarde teóricos, han llevado a la definición del concepto de media, a la identificación de sus propiedades, más tarde a la definición de otras medidas de posición central, como la mediana o moda, que son preferibles a la media en algunas situaciones concretas. Además, ha sido necesario "probar" o "demostrar" la validez de estas soluciones y propiedades, para aceptarlas como parte del conocimiento matemático.

Por tanto, cuando nos preguntamos por el *significado* de la media o de las medidas de posición central, observamos que este significado tiene un carácter complejo y podemos identificar en el mismo los siguientes tipos de elementos:

- *Elementos extensivos*: El campo de problemas de donde surge el objeto. Ejemplos, para el caso de la media serían los problemas tipo P1 a P4 y sus generalizaciones.
- *Elementos actuativos*: Las prácticas empleadas en la solución de problemas, como sumar una serie de valores y dividir por el número de sumandos, encontrar el valor más frecuente en una tabla de frecuencias, calcular las frecuencias acumuladas y hallar el valor al que corresponde la mitad del número total de datos, o integrar el producto de la variable por la función de densidad en un cierto dominio.
- *Elementos ostensivos*: Las notaciones, gráficos, palabras y en general todas las representaciones del objeto abstracto; como los términos "media", "valor medio", "promedio", $E(X)$, $\sum x_i p_i$ μ, $\sum xf(x)dx$, que podemos usar para referirnos al concepto.
- *Elementos intensivos*: Las definiciones y propiedades características y sus relaciones con otros conceptos. Por ejemplo, en la investigación de Strauss y Bichler (1988) encuentra una proporción importante de niños de entre 8 y 12 años eran capaces de comprender y aplicar adecuadamente las propiedades a) c) y d) siguientes de la media, mientras que el resto de ella resultaron demasiado abstractas:
    a. La media es un valor comprendido entre los extremos de la distribución;
    b. La suma de las desviaciones de cada valor a la media es igual a cero;
    c. El valor medio es influenciado por los valores de cada uno de los datos;
    d. La media no tiene por qué ser igual a uno de los valores de los datos;
    e. El valor obtenido de la media de números enteros puede ser una fracción, que no tenga sentido en el contexto de los datos;
    f. Hay que tener en cuenta los valores nulos en el cálculo de la media;
    g. El valor medio es representativo de los valores promediados.

- *Elementos validativos*: Las demostraciones que empleamos para probar las propiedades del concepto y que llegan a formar parte de su significado y los argumentos que empleamos para mostrar a otras personas la solución de los problemas.

3. Dimensiones institucional y personal del conocimiento

En general los problemas no aparecen de forma aislada, sino que los mismos problemas son compartidos dentro de cada institución, y las soluciones encontradas dependen de los instrumentos y prácticas sociales disponibles. Así, problemas similares a P1 de estimación de una cantidad desconocida son compartidos en instituciones de investigación experimental, como la astronomía o la agronomía y también en las instituciones escolares, pero los instrumentos disponibles son muy diferentes en uno y otro caso, de modo que el significado de un concepto matemático, entendido de una forma sistémica, como hemos descrito, varía según la institución considerada.

Los matemáticos y estadísticos profesionales constituyen una institución interesada en resolver problemas de promedios, pero existen otras instituciones diferentes que también podrían estar interesadas en la media, aunque podría atribuirle un significado más restringido al que recibe dentro de la matemática, por ejemplo:

(I1) En la escuela primaria los curricula proponen que se enseñe a los alumnos:

- la definición de la media, mediana y moda en el caso más simple, empleando una notación sencilla (se evita el sumatorio y la ponderación);
- algunos ejemplos de aplicación, limitando el cálculo de las medidas de tendencia central a conjuntos sencillos de datos, y haciéndolo manualmente o con calculadora.
- discriminación respecto de otras medidas de tendencia central (mediana, moda).

(I2) En la escuela secundaria (y en la universidad) se amplia la definición de la media, trabajándose primero con medias ponderadas y luego con medias de variables aleatorias discretas y continuas. Se enuncian y demuestran algunas propiedades de los promedios y se presentan aplicaciones a situaciones problemáticas más realistas y complejas. Por ejemplo, en la universidad se introduce la noción de media o esperanza matemática de una distribución de probabilidad y se muestra que la media es un parámetro que define algunas distribuciones de probabilidad, como la normal; al iniciar el estudio de la inferencia, distinguimos varias medias: media de la muestra, media de la población, media de la media muestral en todas las muestras de tamaño dado.

(I3) En la "vida diaria" encontramos la media en los medios de comunicación y el trabajo profesional, por ejemplo, cuando analizamos los números índices de la evolución de la bolsa, precios, producción, empleo y otros indicadores económicos.

Por otro lado, el conocimiento sobre cada objeto matemático (como la media) no ha sido siempre igual al actual, sino que se ha desarrollado lentamente a lo largo del tiempo, ya que a medida que se han ido resolviendo problemas progresivamente diferentes y más complejos, el objeto se desarrolla y completa en su significado. Consideremos, por ejemplo, el siguiente enunciado:

> *P5. Hay 10 personas en un ascensor, 4 mujeres y 6 hombres. El peso medio de las mujeres es de 60 kilos y el de los hombres de 80. ¿Cuál es el peso medio de las 10 personas del ascensor?*

No podemos ahora resolver este problema por medio de la media aritmética simple (60+80)/2=70, sino que necesitaríamos ampliar el concepto al de media ponderada: (60x4+70x6)=72. Como cualquier otro concepto, la media y otras medidas de tendencia central han tenido un lento desarrollo dentro de la matemática hasta el momento en que fueron reconocidos como conceptos matemáticos e incluidos en la enseñanza. Durante este desarrollo ha sufrido transformaciones progresivas según se ha ido ampliando el campo de problemas asociado.

La didáctica de la matemática ha puesto de manifiesto cómo el aprendizaje del sujeto es también un proceso lento y progresivo que con frecuencia se asemeja a la construcción de los objetos en la ciencia. Para las medidas de posición central no hay todavía un estudio comprehensivo del desarrollo a diversas edades, aunque el trabajo de Watson y Moritz (en prensa) es un primer paso en este estudio. Como veremos en nuestra exposición, en realidad estos conceptos son bastante elaborados, de modo que el conocimiento que un sujeto puede adquirir fuera del ámbito escolar es necesariamente muy limitado y restringido. Ello posiblemente haya influido en la falta de interés por el desarrollo de estos conceptos por parte de la psicología.

Al considerar una cierta institución escolar, como la escuela primaria, el significado construido por un alumno particular, en un momento del proceso de aprendizaje puede no corresponder exactamente al significado del objeto en la institución dada, por lo que conviene distinguir entre *significado institucional* y *significado personal* de un objeto matemático. En la clase de matemáticas el profesor sigue las directrices curriculares, los libros de texto y materiales didácticos -que marcan un significado particular restringido para la media y las medidas de posición central. Al realizar la evaluación, el profesor considera que el alumno "conoce" o "comprende" las medidas de tendencia central si hay un ajuste entre el significado institucional y el personal construido por el sujeto. Si no hay acuerdo entre estos dos significados consideramos que existen dificultades o que el tema es difícil para el alumno. Por esto, también en la comprensión debemos diferenciar una dimensión personal e institucional. En las instituciones escolares se organizan procesos educativos para determinados alumnos, y se asigna al profesor la tarea de ayudar a los estudiantes a adquirir unas propiedades y relaciones culturalmente aceptadas para los términos y expresiones matemáticas. Godino (1996) indica que la comprensión deja de ser meramente un proceso mental y se convierte en un proceso social y que podemos considerar que un alumno "comprende" suficientemente los promedios desde el punto de vista de la enseñanza secundaria y que no lo comprende desde el punto de vista de unos estudios universitarios.

4. Algunas dificultades en la comprensión de las medidas de tendencia central

Al planificar la enseñanza del tema o al tratar de evaluar el aprendizaje de los alumnos, debemos tener en cuenta los cinco tipos de elementos que constituyen el significado sistémico de un objeto matemático y que hemos descrito anteriormente. La comprensión de un concepto no puede reducirse a conocer las definiciones y propiedades (elementos intensivos), sino a reconocer los problemas donde debe emplearse el concepto (elementos extensivos), las notaciones y palabras con que lo denotamos y en general todas sus representaciones (elementos ostensivos), habilidad operatoria en los diferentes algoritmos y procedimientos relacionados con el concepto (elementos actuativos) y capacidad de argumentar y justificar propiedades, relaciones y soluciones de problemas (elementos validativos). A continuación describimos los resultados de investigaciones que han resaltado dificultades en cada unos de estos puntos en relación a las medidas de tendencia central.

Elementos actuativos

El cálculo de la media parece sencillo. Sin embargo Pollatsek, Lima y Well (1981) encontraron que incluso alumnos universitarios no ponderan adecuadamente los valores al resolver el problema P5 y en ocasiones usan la media simple, en lugar de la media ponderada. Li y Shen (1992) indican que cuando se pide a los estudiantes calcular la media a partir de una tabla de frecuencias donde los datos se agrupan en intervalos, los estudiantes olvidan con frecuencia que cada uno de estos grupos debe ponderarse de modo distinto al calcular la media.

En otros casos el algoritmo se aplica de forma mecánica sin comprender su significado. Cai (1995) encontró que mientras la mayoría de alumnos de 12-13 años en su investigación eran capaces de aplicar adecuadamente el algoritmo para calcular la media, sólo algunos alumnos eran capaces de determinar un valor desconocido en un conjunto pequeño de datos para obtener un valor medio dado. Incluso encontrando el valor desconocido, fueron pocos los que lo hicieron a partir de un uso comprensivo del algoritmo, multiplicando el valor medio por el número de valores para hallar la suma total y de ahí el valor faltante, sino que la mayoría simplemente usó el ensayo y error.

Otros errores de cálculo en media, mediana y moda descritos por Carvalho (1998) al analizar las producciones escritas de los alumnos al resolver tareas estadísticas son los siguientes:

- Moda: Tomar la mayor frecuencia absoluta;
- Mediana: No ordenar los datos, para calcular la mediana; calcular el dato central de las frecuencias absolutas ordenadas de forma creciente; calcular la moda en vez de la mediana; equivocarse al calcular el valor central;
- Media: Hallar la media de los valores de las frecuencias; no tener en cuenta la frecuencia absoluta de cada valor en el cálculo de la media.

En realidad, el cálculo de la mediana es complejo, porque el algoritmo de cálculo es diferente, según tengamos un número par o impar de datos, y según los datos se presenten en tablas de valores agrupados o sin agrupar (Cobo y Batanero, en prensa) y también el valor obtenido es diferente, según se aplique uno u otro algoritmo. Esto puede resultar difícil para los alumnos que están acostumbrados a un único método de cálculo y una única solución para los problemas matemáticos.

Gattuso y Mary (1998) analizan la evolución de la comprensión del algoritmo de cálculo de la media ponderada de los alumnos durante la enseñanza secundaria, usando problemas con diferentes contextos y forma de representación. Las tareas presentadas fueron: cálculo de medias ponderadas, efecto que el cambio de un dato produce sobre la media y hallar un valor faltante en un conjunto de datos para obtener un promedio dado. Identifican las siguientes variables didácticas que afectan a la dificultad de las tareas: formato (tabla, serie de números, gráfico), si los valores de las variables son o no mucho mayores que los de las frecuencias (lo que influye en que el niño discrimine los dos conceptos); si una de las frecuencias es mucho mayor que las otras (de modo que se fuerce al niño a tener en cuenta las frecuencias). Observaron el efecto de estas variables y también la mejora con la instrucción, aunque no fue muy persistente en el tiempo.

Elementos intensivos

Cuando los alumnos comienzan a estudiar la media, mediana y moda por primera vez ya conocen ciertas operaciones aritméticas como la suma y multiplicación, e inconscientemente aplican a la operación de "promediar" algunas propiedades de las anteriores operaciones que no se cumplen en el caso de los promedios. Mevarech (1983) observa que incluso los estudiantes universitarios piensan que la media tiene la propiedad asociativa y cuando tienen que hallar la media de un conjunto grande de números, lo dividen en partes hallando primero la media de cada parte y luego promediando el resultado obtenido. Podemos comprobar que esta propiedad, en general, no es cierta, si hallamos primero la media de tres números diferentes y luego promediamos los dos primeros y hacemos la media del valor obtenido con el último elemento. En otros casos no se tiene en cuenta el cero en el cálculo de la media, como si fuese un elemento neutro o bien se piensa que la media debe ser un elemento del mismo conjunto numérico del que se toman los datos.

Strauss y Bichler (1988) analizan la comprensión de los niños de 8 a 14 años de las propiedades a) a e) que hemos listado al describir los elementos intensivos en el punto 2. Aunque una proporción importante de niños parecieron usar espontáneamente estas propiedades, algunos niños no tenían en cuenta el cero para calcular la media, o bien suponían que la media podría estar fuera del rango de variación de la variable, o que debería coincidir con uno de los valores de los datos.

León y Zawojeski (1991) realizan entrevistas a niños entre 8 y 14 años y analizan el efecto de la edad sobre la comprensión de estas propiedades. Además de encontrar una importante influencia de la edad sobre la comprensión de la media, también observaron que la contextualización de las tareas facilita mucho su resolución. Sin embargo, propiedades tales como que la suma de desviaciones respecto a la media es cero, que la media es un valor representativo de los valores promediados o que hay que tener en cuenta los valores nulos en el cálculo de la media continuaron siendo demasiado abstractas para una proporción importante de alumnos de 14 años.

La idea de representante de un conjunto de datos es importante en las aplicaciones prácticas, por ejemplo, al comparar dos conjuntos de datos respecto a una misma variable de interés. Como indican Mokros y Russell (1995) hasta que los niños no conciben el conjunto de datos como un todo, y no como un agregado de valores, no podrán comprender las ideas de resumen de los datos o representante de los datos, que se refiere al conjunto global y no a ninguno de sus valores aislados.

Por otro lado, 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. Esto no es siempre comprendido por algunos alumnos quienes invariablemente eligen la media como mejor representante de los datos sin tener en cuenta la simetría de la distribución o la existencia de valores atípicos, como hemos observado en nuestra propia experiencia.

Respecto a la comprensión de la mediana Barr, (1980) indica que los alumnos entienden que la mediana es el centro de "algo" pero no siempre comprenden a que se refiere ese "algo" porque no comprenden realmente que una tabla de frecuencia es sólo un resumen de los datos y no son capaces de pasar de la tabla a la lista de valores que es una representación alternativa de los datos. Incluso si se les da los datos en forma de lista no entienden por qué hay que ordenarlos para calcular la mediana, porque no entienden que la mediana es un estadístico que se refiere al conjunto ordenado de datos.

Elementos extensivos

No sirve de nada conocer las definiciones de las medidas de posición central y saber calcularlas si luego no se reconocen los problemas relacionados con estos conceptos. Pollasek y cols. (1981) propusieron a sus alumnos el siguiente problema, que es semejante al P4 descrito anteriormente:

> *La media en fluidez verbal de una clase de un colegio es de 400. Si extraemos una muestra aleatoria de 5 estudiantes y resulta que la puntuación de los 4 primeros es de 380, 420, 600, 400. ¿Cuál sería aproximadamente la puntuación esperada para el quinto estudiante?*

La respuesta correcta a este ítem es 400, el valor esperado en la población. Sin embargo, son pocos los alumnos que dieron una respuesta correcta al problema, en la investigación citada, sino que, generalmente se busca un valor de la puntuación del quinto sujeto tal que, sumada a las cuatro anteriores, dé una media de 400.

Respecto al problema P3 (usar la media como representante de un conjunto de datos), resulta aún más difícil para los alumnos construir un conjunto de datos que tenga un promedio dado. Goodchild (1988) proporcionó a los estudiantes cajas de cerillas en las que se había impreso la frase "contenido medio 35 cerillas" y pidió a sus alumnos construir una distribución hipotética del contenido de 100 cajas. Lo que más le sorprendió fue que las distribuciones construidas por los alumnos, no tenían forma acampanada como la distribución normal. Goodchild sugirió que ello se debe a la falta de comprensión de la media como medida de posición central de la distribución.

En nuestra investigación (Estepa y Batanero, 1994) con alumnos del curso preuniversitario hemos observado casos de alumnos que basan la comparación de dos conjuntos de datos en valores aislados, por ejemplo, en la comparación de los máximos o los mínimos, o bien en la comparación de totales, o la inspección visual de la distribución global.

Elementos ostensivos y validativos

Los términos matemáticos con que designamos los conceptos tiene un significado preciso, pero éste no siempre coincide con el asignado al término en el lenguaje coloquial. Russell y Mokros (1991) clasificaron en cuatro categorías los significados incorrectos atribuidos por los estudiantes a la palabra "media": valor más frecuente (en realidad esto sería una confusión son la palabra "moda"), "valor razonable" (significado coloquial del término), "punto medio" (confusión con la mediana) y "algoritmo" (es un significado restringido, donde la media se ve sólo como el algoritmo de cálculo). Watson y Moritz (en prensa), 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). Pocas veces se relaciona la palabra "promedio" con la moda y menos aún con la media aritmética. Las siguientes definiciones de "promedio" fueron obtenidas en entrevistas a niños realizadas por Watson y Moritz (en prensa): *"Significa igual", "que es normal", "no eres realmente bueno, pero tampoco malo".*

Al preguntar que quiere decir que el número medio de niños por familia es 2'3, obtienen respuestas correctas y otras como las siguientes: *"Que tienen dos niños grandes y otro que no ha crecido todavía", "que en las familias australianas el número más frecuente de niños es 2'3", "el '3 es un niño que tiene que crecer para hacerse mayor. Por ejemplo, tiene 3 años ahora y cuando cumpla 10, contará como 1 y entonces el número promedio de niños será 3".* Para el profesor los enunciados sobre los promedios pueden parecer muy claros, pero estas respuestas indican la necesidad de poner atención al significado que las palabras y valores numéricos tienen para los estudiantes en relación a contextos específicos.

Eisenbach (1994) plantea a estudiantes universitarios en un curso introductorio de estadística el significado de la frase: *"¿Qué quiere decir que el salario medio de un empleado es 3.600 dólares?"* obteniendo respuestas como *"que la mayoría de los empleados gana alrededor de 3.600 dólares",* o que *"es el salario central; los otros trabajadores ganan más o menos de 3600 dólares",* que muestran la confusión terminológica entre las palabras "media", "mediana" y "moda".

Ya hemos indicado que la idea de promedio no se puede comprender hasta tanto se visualice el conjunto de datos como un todo y también que la forma de presentación de los datos (tabla, gráfico, datos sin tabular) incide en la dificultad de las tareas. Reading y Pegg (1996) estudiando la forma en que los niños de grados 7 a 12 reducen los conjuntos de datos observaron que algunos alumnos que eran capaces de dar un resumen de datos presentado en forma numérica, fracasaron en la tarea cuando los datos se presentaban por medio de un gráfico estadístico. También observaron que los niños mostraban dificultad a la hora de dar un argumento o justificar su respuesta de por qué se elegía un cierto promedio, al plantearles el siguiente problema, que es, esencialmente una versión simplificada del problema P3.

Como parte de un proyecto los estudiantes de una clase miden cada uno su número de calzado, obteniéndose los siguientes datos:

26 26 26 27 27 27 27 28 28 28 28 28 28 29
29 29 29 29 30 30 30 30 30 30 30 31 32 32
33

Si te preguntan cuál sería el mejor número para representar este conjunto de datos, ¿Qué número o números elegirías? Explícanos por qué has elegido ese(esos) número(s).

En su investigación clasifica a los alumnos en 8 niveles diferentes de respuesta, pero, incluso los estudiantes de nivel 6 y 7 que son capaces de proporcionar un resumen como la media, son incapaces de argumentar el por qué de su decisión, más allá de dar la definición del concepto. Solo una pequeña parte de los estudiantes de su investigación (nivel 8) fueron capaces de justificar la elección de las medidas de valor central y dispersión relacionándolas con características del conjunto de datos.

Por nuestra parte en Estepa y Batanero (1999) hemos encontrado algunos alumnos del curso preuniversitario que basan sus argumentos en sus teorías previas, en lugar de en los datos al plantearles problemas similares al P3.

5. Algunas implicaciones para la enseñanza

Este trabajo pone de manifiesto que los conceptos estadísticos, incluso los más sencillos como la media, mediana y moda tienen un significado complejo y por tanto será necesario un periodo dilatado de enseñanza a lo largo de la educación primaria y secundaria para lograr el progresivo acoplamiento de los significados personales que construyen los alumnos a los significados institucionales que pretendemos adquieran.

Tradicionalmente, en el aprendizaje de la estadística, se ha dado una gran importancia al cálculo y a los aspectos actuativos, que ahora pierden importancia, debido a las nuevas tecnologías. En lugar de tener que ejercitarse en la realización con lápiz y papel de cálculos y gráficos, el alumno debe aprender el uso de calculadoras gráficas y programas de ordenador, como la hoja de cálculo. Las nuevas tecnologías introducen también nuevos elementos actuativos y ostensivos, ya que el rango de representaciones disponibles es mucho mayor. Permiten también plantear situaciones de aprendizaje en las que el alumno se enfrente a problemas más reales cuya solución requiera el uso y aprendizaje de conceptos estadísticos. Estas situaciones requieren también el trabajo cooperativo, motivan el interés del alumno y le permiten explorar tanto los datos, como los conceptos implicados, reforzando los elementos intensivos y validativos.

Por supuesto la enseñanza de la estadística en la escuela usando ordenadores requiere una planificación cuidadosa. Frecuentemente los datos reales son demasiado complejos y es necesario tomar versiones simplificadas de los conjuntos de datos. Si queremos mostrar una cierta propiedad, será preciso a veces manipular el conjunto de datos, para, por ejemplo, conseguir que el valor de la media, mediana y moda sean marcadamente diferentes.

El trabajo con ordenador debe, además, ser complementado con otras situaciones encaminadas a que el alumno se familiarice con los campos de problemas, las representaciones, tipos de prácticas y propiedades de los promedios, y que ejercite su capacidad de argumentación. Las distintas situaciones problemáticas que hemos presentado como ejemplo a lo largo de este trabajo pueden servir de partida para plantear en clase situaciones didácticas que contribuyan a lograr este aprendizaje.

Referencias

Batanero, C. (2000). Cap on va l'educació estadística?. *Blaix*, 15, 2-13.

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.

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

Cai, J. (1995). Beyod the computational algorithm: Students' understanding of the arithmetic average concept. En L. Meira (Ed.). *Proceedings of the 19th PME Conference* (v.3, pp. 145-151). Universidade Federal de Pernambuco, Recife, Brasil.

Carvalho. C. (1998). Tarefas estatísticas e estratégias de resposta. Comunicación presentada en el *VI Encuentro en Educación Matemática de la Sociedad Portuguesa de Ciencias de la Educación.* Castelo de Vide, Portugal.

Cobo, B. y Batanero, C. (En prensa). La mediana en la educación secundaria ¿Un concepto sencillo? *UNO.*

Estepa, A. y Batanero, C. (1999). Judgments of association in the comparison of two samples: students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education.*

Eisenbach, R. (1994). What does de mean mean? Comunicación presentada en el *Fourth International Conference on Teaching Statistics*. Marrakesh, Marruecos.

Gattuso, L. y Mary, C. (1998). Development of the concept of weighted average among high-school students. En L. Pereira-Mendoza, L. Seu Keu, T. Wee Kee y W. K. Wong, *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 685-691). Singapur: International Association for Statistical Education.

Godino, J. D. (1996). Mathematical concepts, their meanings and understanding]. En, L. Puig y A. Gutiérrez (Eds.), *Proceedings of the 20th PME Conference* (v.2, pp. 417-424). Universidad de Valencia, España.

Godino, J. D. y Batanero, C. (1994). Significado personal e institucional de los objetos matemáticos. *Recherches en Didactique des Mathématiques,* 14 (3), 325-355.

Goodchild (1988). School pupils' understanding of average. *Teaching Statistics,* 10, 77-81.

Leon, M. R., y Zawokeswski, J. S. (1991). Use of the arithmetic mean: An investigation of four properties. Issues and preliminary resuls. En D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 302-306). Voorburg, Holanda: International Statistical Institute.

Li, K. Y. y Shen, S. M. (1992). Students' weaknesses in statistical projects. *Teaching Statistics*, 14 (1), 2-8.

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

Mokros, J. y Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education,* 26, 20-39.

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

Plackett, R.L. (1970), The principle of the arithmetic mean. En E. S. Pearson y M. Kendall (Eds), *Studies in the history of statistics and probability* (v,1, pp. 121-126). London, Charles Griffin.

Reading, C. y Pegg, J. (1996). Exploring understanding of data reduction. En L. Puig y A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International group for the Psychology of Mathematics Education* (v.4, pp. 187-194). Universidad de Valencia, España.

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: 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.

Watson, J. M. y Moritz, J. B. (En prensa). The longitudinal development of understanding of average. *Mathematical Thinking and Learning.*