

SIGNIFICADO Y COMPRENSIÓN DE LAS MEDIDAS DE TENDENCIA CENTRAL. UN ESTUDIO DE CASO EN ESTUDIANTES DE EDUCACIÓN SUPERIOR.

FÉLIX ALCÁNTARA MORENO Y MARÍA DE LOURDES RÍOS YESCAS

Universidad Pedagógica Nacional

Resumen

Se presenta un estudio de caso desde un análisis conceptual y didáctico de las medidas de tendencia central: la mediana y la media, así como un acercamiento al concepto, cálculo y propiedades. El tema de estudio se justifica, dada la importancia que tiene en el análisis de datos, en donde cobra un papel significativo.

Introducción

En la actualidad se vive un proceso de reforma de la enseñanza que ha alcanzado el nivel superior, en el cual se observa que se da una mayor importancia a la formación estocástica de los alumnos. Este fenómeno ocurre en todos los países que han venido preocupándose desde varias décadas de la mejora de sus sistemas de enseñanza.

Los nuevos diseños curriculares recomiendan la introducción de una filosofía de aplicación de la estadística en una cantidad, cada vez mayor, de problemas interdisciplinarios y con datos reales. Asimismo, se recomienda el trabajo con ordenadores o calculadoras científicas en el análisis de datos.

Las posibles dificultades de comprensión conceptual de la mediana, por parte de los alumnos, no han sido tan profusamente estudiadas como lo ha sido la media aritmética (Batanero y cols., 1994 y Tormo, 1995).

En este trabajo presentamos un estudio de caso desde el punto de vista matemático y didáctico de dos de los estadísticos de orden más importantes: la mediana y la media, tanto en su aspecto conceptual como de su uso en la resolución de problemas de análisis exploratorio de datos.

La mediana y la media ¿son conceptos de fácil comprensión?

Suele definirse la mediana como “el punto medio” o el “centro de la distribución”. Desde esta definición, los alumnos parecen comprender que la mediana es el centro de “algo”, pero con frecuencia la identifican como el centro del recorrido de la variable o el valor que ocupa la posición central, incluso aunque el conjunto de datos no esté ordenado. Sin embargo, no es sencillo darles una definición clara y precisa, exenta de ambigüedades, que no les ocasione confusión.

Existen diferentes definiciones equivalentes de la mediana, usadas en los textos de nivel superior, reproducimos algunas de ellas a continuación:

Definición 1 (Moore, 1998)

Para hallar la mediana de una distribución:

1. Ordena todas las observaciones de la mínima a la máxima.

2. Si el número de observaciones n es impar, entonces la mediana M es la observación central de la lista ordenada. Halla la posición de la mediana contando $(n+1)/2$ observaciones desde el comienzo de la lista.
3. Si el número de observaciones n es par, entonces la mediana M es la media de las dos observaciones centrales de la lista ordenada. La posición de la mediana se halla otra vez, contando $(n+1)/2$ observaciones desde el comienzo de la lista.

Definición 2 (Johnson, 2005)

La mediana es el número medio obtenido cuando los datos se ponen en orden de acuerdo con su tamaño.

Si el número de datos n es impar, la posición de la mediana se determina mediante la fórmula

$$\text{Posición de la mediana} = i = \frac{n+1}{2},$$

donde n representa el número de datos e i es la posición que la mediana ocupa en los datos arreglados por rango.

Si n es par, la posición de la mediana será siempre una fracción, compuesta por un entero y una mitad; es decir, la posición de la mediana quedará entre dos enteros; entonces la mediana se calcula tomando la mitad de la suma de los valores correspondientes a dichos enteros. Por ejemplo, para el conjunto de valores ordenados 6, 7, 8, 9, 9 y 10, se tiene que $n = 6$ y $\frac{n+1}{2} = 3.5$; por lo tanto, la mediana $= \frac{8+9}{2} = 8.5$.

Definición 3 (Hernández, 1979)

Suponiendo que los valores de una muestra están arreglados en orden de magnitud, la mediana de una muestra se define como el valor central, o sea el valor que está “en el medio”, si el número de valores de la muestra es impar, o como el promedio de los dos valores centrales si el número de valores es par. Por ejemplo, para la muestra 2, 3, 3, 5, 3, 6, 9, 8 se observa que al ordenarla se obtiene 2, 3, 3, 3, 5, 6, 8, 9; entonces la mediana es $\frac{3+5}{2} = 4$.

Mientras que para la muestra 2, 3, 3, 5, 3, 5, 6, 5, 9, 8, 6, que tiene 11 valores, se puede verificar que la mediana es 5.

Definición

La definición de la media aritmética suele ser fácil de explicar no así su comprensión cuando se aplica en situaciones que requieren un análisis más

concienzudo. En general, la media para un conjunto de datos, digamos n , se define como:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

¿Cómo se interpretan los datos con la mediana y la media?

La principal dificultad que nos encontramos es que nuestros alumnos que estudian la licenciatura en pedagogía no aplican estos conceptos en el análisis de los datos, de manera correcta, en múltiples situaciones. Es decir, deben aprender para qué sirven y cómo se usan tanto la mediana como la media en la interpretación de un conjunto de datos.

Antes de dar un ejemplo que nos acerque al uso de estas medidas de tendencia central en la interpretación de los datos de una muestra, damos las siguientes definiciones.

Una **observación atípica** de un conjunto de datos es la observación individual que no queda descrita por el aspecto general del conjunto de datos y es una **observación típica** el caso contrario.

La mediana adquiere sentido cuando se compara con la media, pues ello permite analizar en qué situaciones la mediana tiene ventajas y en qué otras desventajas con respecto a la media en el análisis de una muestra como

medidas de tendencia central de los datos; lo cual depende de la forma en que están distribuidos los datos del propósito de la información que se obtenga.

Eisenbach (1994) plantea a estudiantes universitarios en un curso introductoria 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 3,600 dólares, que muestran la confusión terminológica entre las palabras “media”, “mediana” y “moda”.

Analicemos el siguiente ejemplo en donde se muestra el uso de la mediana y la media (Fuente: Moore, 1995).

Un estudio realizado en Suiza examinó el número de histerectomías (extirpación del útero) realizadas durante un año por un grupo de médicos. He aquí los datos de una muestra de 15 cirujanos.

27 50 33 25 86 25 85 31 37 44 20 36 59 34 28

El diagrama de tallos muestra que la distribución es asimétrica hacia la derecha y que hay dos observaciones atípicas en el lado de los valores mayores.

2	05578
3	13467
4	4
5	09
6	
7	
8	56

La media de las operaciones es $\bar{x} = 41.3$. Se observa que sólo 5 de los 15 médicos realizaron un número de histerectomías mayor que la media. Esto se debe a las dos observaciones atípicas (85 y 86) hacen aumentar el valor de la media. Se verifica que la media de las restantes 13 observaciones es 34.5. La media es el promedio de los valores, pero no es un número de operaciones realizadas por un cirujano típico. Es decir; es evidente que el resultado no refleja la verdad. Es quizás una situación semejante la que hizo decir a alguien que “existen tres tipos de mentiras: la mentira, la mentira descarada y la estadística”. Este ejemplo es un caso de “engaño estadístico” si se afirmará con base en la media, que el cirujano típico realiza un total de 34.5 histerectomías durante un año.

La mediana de las mismas 15 observaciones es 34, pensamos que en esta situación es una medida de tendencia central que aporta una mejor información sobre el número de operaciones realizadas por un cirujano típico.

Este ejemplo ilustra un hecho importante sobre la media como medida de tendencia central: es sensible a la influencia de unas pocas observaciones extremas, no así la mediana.

La media y la mediana son dos maneras distintas de medir el centro y ambas son útiles. Utilizamos la mediana si queremos el número de histerectomías realizadas por un cirujano típico. Utilizamos la media si nos interesa el número total de operaciones realizadas por todos los cirujanos, pues el número total de operaciones es n veces la media, donde n es el número de cirujanos. El total de operaciones no se puede calcular a partir de la mediana.

Estudio de Caso

Realizamos un estudio de caso con un grupo de 20 estudiantes universitarios de pedagogía pertenecientes a un curso introductorio de estadística. Se les plantearon dos preguntas.

En primer lugar se les plantearon dos preguntas:

1. ¿Qué quiere decir que el número promedio de histerectomías en una muestra de 1000 cirujanos es de 90? Seleccione la respuesta correcta.
 - a) La mayoría de los médicos realizan alrededor de 90 operaciones mensuales.

- b) Es el número de operaciones central que realizan los cirujanos.
- c) Una minoría de los cirujanos realizan más o menos 90 operaciones.

2. 10 estudiantes presentan un examen de biología, 4 mujeres y 6 hombres. El promedio de las calificaciones de las mujeres es de 6 y el de los hombres es de 8. ¿Cuál es el promedio de los 10 estudiantes?

Las frecuencias para las opciones a), b) y c) para la primera de las preguntas se distribuyó como sigue:

- 12 estudiantes seleccionaron la respuesta a)
- 7 estudiantes dieron como verdadera la respuesta b)
- Sólo 3 estudiantes consideraron como verdadera la opción c)

Las respuestas para esta pregunta muestran que existe una confusión de las medidas de tendencia central. Es decir, de manera general no existe una discriminación entre dichas medidas. Nuestros resultados se asemejan a los obtenidos por Eisenbach. Después del estudio, se les indicó que la idea de promedio no se puede comprender hasta tanto no se tenga una visualización de conjunto de los datos.

En la segunda pregunta, 14 estudiantes la resolvieron usando la media aritmética simple $(6+8)/2=7$. Sólo 6 de los estudiantes aplicaron el concepto de media ponderada: $(6 \times 4 + 8 \times 6)/10 = 7.2$. Se observa nuevamente que la media

aritmética como en el caso de la mediana tiene un lento desarrollo en nuestros estudiantes; es decir, requiere tiempo y trabajo para su construcción.

Algunas implicaciones para la enseñanza

Consideramos que es necesario que nuestro estudiante comprenda que la media y la mediana de una distribución simétrica se encuentran muy cerca. Más aún, si la distribución es exactamente simétrica, la media y la mediana son exactamente iguales. En una distribución asimétrica, la media queda desplazada hacia la cola más larga.

Nuestros estudiantes tienden 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; en tal caso la moda o la mediana podrían ser valores más representativos 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.

Necesitamos que nuestros estudiantes se enfrenten a situaciones de aprendizaje significativas, de tal manera que sus habilidades matemáticas en el campo de la estadística se vean enriquecidas por distintas y múltiples experiencias.

Estos son sólo algunos aspectos en los que no se hace mucho hincapié cuando se explican las medidas de tendencia central; es decir, se definen éstas de manera descontextualizada, dando como consecuencia que nuestros alumnos no aprendan a usar los conceptos estadísticos en situaciones reales.

Referencias

- Batanero, C., Rodino, J. D., Vallecillos, A., Green, D. R. y Holmes, P. (1994). Errors and difficulties in understanding statistical concepts. *Internacional Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.
- Eisenbach, R. (1994). What does the mean mean? Comunicación presentada en el Fourth International Conference on Teaching Statistics. Marrakesh, Marruecos.
- Hernández, O. (1979). Elementos de probabilidad y Estadística. Fondo de Cultura Económica. México.
- Johnson, R. (2005). Estadística Elemental. Trillas, Séptima reimpresión, 2005. México.
- Moore, D. (1998). Estadística Aplicada Básica. Antoni Bosch. España.
- Tormo, C. (1995). Dificultades del alumnado respecto a la media aritmética. Uno, 5, 29-36.