

Tema 3. Medidas descriptivas

Introducción

Cuando se describe un conjunto de datos, suelen utilizarse medidas de tendencia central y de variabilidad o dispersión. Las primeras (media, mediana y moda) muestran el punto central del conjunto de datos; por su parte, las segundas (desviación estándar, varianza, valores mínimo y máximo, entre otras) indican la dispersión o variación de los datos.



Explicación

Media, moda y mediana

De acuerdo con Obando (2022), las medidas de tendencia central son estadísticas que resumen la información; entre ellas, se encuentran las siguientes:

- Media. Se obtiene al sumar todos los datos y dividir el resultado entre el total de elementos en la muestra.
- Mediana. Divide el conjunto de datos en dos partes iguales, es decir, una mitad de los valores está por encima y la otra por debajo de un determinado elemento.
- Moda. Es el valor que aparece con mayor frecuencia en el conjunto de datos.

Cada medida de tendencia central tiene sus ventajas y limitaciones, así que elegir la más adecuada dependerá del tipo de datos y del objetivo del análisis.

Observa esta colección de datos: {1, 2, 3, 4, 5, 6, 7, 8, 9}. Para calcular la medida de tendencia central llamada media aritmética o promedio, debes hacer lo siguiente:

1. Suma todos los datos de la colección:

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 45$$

2. Determina cuántos números hay en la serie. En este caso, puedes observar que el total de datos es nueve.

3. Para calcular la media, divide la suma de los datos entre la cantidad de elementos.

$$\mu = \frac{45}{9} = 5$$

4. Por tanto, el promedio o media de esta muestra es 5.

La siguiente medida es la mediana; para calcularla, considera el siguiente conjunto de datos:

$$\{2, 4, 10, 3, 1, 12, 5, 9, 7, 6, 11, 8, 13\}$$

1. Ordena los datos de la muestra de menor a mayor:

$$\{2, 4, 10, 3, 1, 12, 5, 9, 7, 6, 11, 8, 13\} \rightarrow \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$$

2. Determina cuántos datos hay y verifica si es una cifra par o impar; en este caso, hay 13 valores, es decir, se trata de un número impar.
3. Encuentra el dato que queda justo en medio de los valores de la muestra. Como es una cantidad impar, entonces la mediana equivale al elemento que se encuentra en la posición central:

$$\text{Posición central} = (\text{número de datos} + 1) / 2 = (13 + 1) / 2 = 7$$

Como el valor central en la posición siete es precisamente el 7, este número equivale a la mediana, ya que hay una misma cantidad de datos antes y después que él. Sin embargo, el total de datos que contiene una serie también puede ser par y, por tanto, cambia la manera como se calcula esta medida.

En este caso, considera la siguiente colección con una cantidad par de valores: {1, 4, 3, 4, 2, 4, 5, 6, 7, 6, 8, 5}. El cálculo de la mediana se realiza de la siguiente manera:

1. Ordena los datos de menor a mayor:

$$\{1, 4, 3, 4, 2, 4, 5, 8, 7, 6, 8, 5\} \rightarrow \{1, 2, 3, 4, 4, 4, 5, 5, 6, 7, 8, 8\}$$

2. Determina cuántos valores tiene la muestra; en este caso, el total es 12, una cifra par. Para encontrar el valor de la mediana, calcula la media de los dos valores centrales:

$$\text{Mediana} = (4 + 5) / 2 = 4.5$$

Ahora bien, encontrar la moda de un conjunto de datos resulta muy sencillo, solo necesitas esclarecer cuál es el valor que aparece con mayor frecuencia en la muestra.

Como el valor central en la posición siete es precisamente el 7, este número equivale a la mediana, ya que hay una misma cantidad de datos antes y después que él. Sin embargo, el total de datos que contiene una serie también puede ser par y, por tanto, cambia la manera como se calcula esta medida.

En este caso, considera la siguiente colección con una cantidad par de valores: {1, 4, 3, 4, 2, 4, 5, 6, 7, 6, 8, 5}. El cálculo de la mediana se realiza de la siguiente manera:

1. Ordena los datos de menor a mayor:

$$\{1, 4, 3, 4, 2, 4, 5, 8, 7, 6, 8, 5\} \rightarrow \{1, 2, 3, 4, 4, 4, 5, 5, 6, 7, 8, 8\}$$

2. Determina cuántos valores tiene la muestra; en este caso, el total es 12, una cifra par. Para encontrar el valor de la mediana, calcula la media de los dos valores centrales:

$$\text{Mediana} = (4 + 5) / 2 = 4.5$$

Ahora bien, encontrar la moda de un conjunto de datos resulta muy sencillo, solo necesitas esclarecer cuál es el valor que aparece con mayor frecuencia en la muestra.

$$\{1, 4, 3, 4, 2, 4, 5, 6, 7, 6, 8, 5\} \rightarrow \{1, 2, 3, 4, 4, 4, 5, 5, 6, 6, 7, 8\}$$

2. Cuenta las veces que se repite cada valor de la muestra y coloca dicha recurrencia en una lista de pares, la cual debe incluir valor y frecuencia.

$$\{(1, 1), (2, 1), (3, 1), (4, 3), (5, 2), (6, 2), (7, 1), (8, 1)\}$$

3. Analiza cuál de los valores tiene la mayor frecuencia; en este caso, se trata del 4. Por tanto, es el número que representa la moda de la colección.

Así como las medidas de centralización se utilizan para representar los datos con respecto a su posición, existen otras conocidas como medidas de dispersión. De acuerdo con Zapata (2022), estas se utilizan para medir la distancia que se encuentra entre cada uno de los datos con respecto a la media o promedio.

Varianza y desviación estándar

Entre las medidas de dispersión se encuentran las siguientes:

- Rango. Es la diferencia de valores entre el dato mayor (máximo) y el menor (mínimo) en una colección. Por ejemplo, en la serie {1, 2, 3, 4, 5, 6}, el mínimo es 1 y el máximo 6; por tanto, el rango se obtiene a partir de esa diferencia.

$$\text{Rango} = \text{Máximo} - \text{Mínimo} = 6 - 1 = 5$$

- Varianza. Esta medida de dispersión se utiliza para comparar el dato central (media aritmética) con respecto a toda la colección. Esta es su fórmula:

$$\sigma^2 = \frac{\sum(x - \bar{x})^2 \cdot f}{n}$$

Donde:

σ^2 = Varianza.

σ = Desviación estándar.

x = Cada uno de los datos.

\bar{x} = Media aritmética de los datos.

n = Cantidad de datos.

f = Frecuencia (cantidad de veces que se repite el dato).

Observa la siguiente colección de datos:

$$\{8, 9, 11, 13, 15, 10, 8, 7, 8, 9, 10, 12\}$$

1. Para obtener la varianza, primero calcula la media:

$$\sigma^2 = \frac{8 + 9 + 11 + 13 + 15 + 10 + 8 + 7 + 8 + 9}{10} = \frac{98}{10} = 9.8$$

2. Una vez calculada la media, obtienes la varianza:

$$\frac{(8 - 10)^2 + (9 - 10)^2 + (11 - 10)^2 + (13 - 10)^2 + (15 - 10)^2 + (10 - 10)^2 + (8 - 10)^2 + (7 - 10)^2 + (8 - 10)^2 + (9 - 10)^2}{10} =$$

$$\sigma^2 = \frac{(-2)^2 + (-1)^2 + (1)^2 + (3)^2 + (5)^2 + (0)^2 + (-2)^2 + (-3)^2 + (-2)^2 + (-1)^2}{10} =$$

$$\sigma^2 = \frac{4 + 1 + 1 + 9 + 25 + 1 + 4 + 9 + 7 + 1}{10} = \frac{62}{10} = 6.2$$

- Desviación estándar. Es otra medida de dispersión y se caracteriza por ser la raíz cuadrada de la varianza. Este valor indica la dispersión o distancia entre el dato central (media) con respecto al conjunto de datos.

Para el ejemplo anterior, la desviación estándar se obtiene de esta forma:

$$\sigma = \sqrt{\sigma^2} = \sqrt{6.2} = 2.49$$

Cierre



Es indispensable representar los datos en su totalidad, ya que esto facilita su correcta descripción y estudio. Entre las maneras de analizar un conjunto de valores, se encuentran las medidas de posición o centralización, como la media aritmética o promedio, que resume el total de elementos; el mejor ejemplo se da con las notas escolares, ya que con esta medida se conoce la calificación general. Por otro lado, las medidas de dispersión evidencian que tan cerca o lejos se encuentran todos los datos en relación con el valor central. En conjunto, ambos tipos de medidas permiten tener información de los datos estudiados.

Checkpoint



Asegúrate de:

- Identificar las medidas de centralización y dispersión para aplicarlas en la resolución de ejercicios.
- Calcular de manera adecuada cada una de las medidas de centralización y dispersión para resolver correctamente casos de estudio.

Bibliografía



- Obando, I. (2022). *¿Qué son las medidas de tendencia central y para qué sirven?* Recuperado de <https://easyfreeconomics.com/que-son-las-medidas-de-tendencia-central-y-para-que-sirven/>
- Zapata, F. (2022). *Medidas de dispersión.* Recuperado de <https://www.lifeder.com/medidas-de-dispersion/>

La obra presentada es propiedad de ENSEÑANZA E INVESTIGACIÓN SUPERIOR A.C. (UNIVERSIDAD TECMILENIO), protegida por la Ley Federal de Derechos de Autor; la alteración o deformación de una obra, así como su reproducción, exhibición o ejecución pública sin el consentimiento de su autor y titular de los derechos correspondientes es constitutivo de un delito tipificado en la Ley Federal de Derechos de Autor, así como en las Leyes Internacionales de Derecho de Autor.

El uso de imágenes, fragmentos de videos, fragmentos de eventos culturales, programas y demás material que sea objeto de protección de los derechos de autor, es exclusivamente para fines educativos e informativos, y cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por UNIVERSIDAD TECMILENIO.

Queda prohibido copiar, reproducir, distribuir, publicar, transmitir, difundir, o en cualquier modo explotar cualquier parte de esta obra sin la autorización previa por escrito de UNIVERSIDAD TECMILENIO. Sin embargo, usted podrá bajar material a su computadora personal para uso exclusivamente personal o educacional y no comercial limitado a una copia por página. No se podrá remover o alterar de la copia ninguna leyenda de Derechos de Autor o la que manifieste la autoría del material.