Medidas de Variabilidad

La *dispersión* o *variación* de los datos es el grado en que los datos numéricos tienden a esparcirse alrededor de un valor promedio. Existen diversas modas de dispersión o variación, siendo las más comunes el rango, la desviación media, el rango semiintercuartilar, el rango percentilar 10-90 y la desviación estándar.

Las medidas de variación son una herramienta para el análisis de los datos ya que son importantes para indicarnos qué tan alejados están los datos o bien cuál es el valor del rango para ver la amplitud o la desviación estándar que nos sirve para poder tener una visión más general a la hora de tomar decisiones.

Rango

El *rango* de un conjunto números es la diferencia entre el número mayor y el menor del conjunto.

Ejemplo:

El rango del conjunto 2, 3, 3, 5, 5, 5, 8, 10, 12 es 12-2= 10. Algunas veces el rango se obtiene simplemente señalando el número más pequeño y el más grande; en el conjunto anterior, por ejemplo, el rango podría indicarse como de 2 a 12 o 2-12.

Desviación media

La *desviación media o desviación promedio* de un conjunto de *N* números X_1, X_2.... X_N se abrevia DM y se define como:

$$(\mathrm{DM}) = \frac{\sum_{j=1}^{N} |X_j - \bar{X}|}{N} = \frac{\sum |X - \bar{X}|}{N} = \overline{|X - \bar{X}|}$$

Donde $\bar{X}$ es la media aritmética de los números y $|X_j - \bar{X}|$es el valor absoluto de la desviación de X_j respecto $\bar{X}$. (El *valor absoluto* de un número es el número sin el signo asociado y se indica con dos líneas verticales colocadas a los lados del número; así |-4|=4, |+3|=3, |6|=6 y |-0.84|= 0.84.)

Ejemplo

Calcule la desviación media del conjunto 2, 3, 6, 8, 11.

Media aritmética ($\bar{X}$) = 2+3+6+8+11/5 = 6

$$\mathrm{DM} = \frac{|2-6| + |3-6| + |6-6| + |8-6| + |11-6|}{5} = \frac{|-4| + |-3| + |0| + |2| + |5|}{5} = \frac{4+3+0+2+5}{5} = 2.8$$

Media aritmética ($\bar{X}$) = 2.8

Rango semiintercuartilar

El *rango semiintercuartilar* o *desviación cuartilar* de un conjunto de datos se denota por Q y se define como

$$Q = \frac{Q_3 - Q_1}{2}$$

Donde Q_1 y Q_3 son el primer y tercer cuartiles de los datos.

Algunas veces se usa el rango intercuatilar Q_3-Q_1, aunque el rango semiintercuartilar es más común como medida de dispersión.

Para profundizar y mayor explicación del tema de cuartiles y rango te invitamos a que visites el siguiente video:

Video: Cuartiles y Rango intercuartilico

https://www.youtube.com/watch?v=w1lyL5WCr-0

Desviación estándar

La *desviación estándar* de un conjunto de N números X_1, X_2, ..., X_N se denota por ***s*** y se define como:

$$= \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

donde x representa las desviaciones de cada uno de los números X respecto de la media. Así que s es la media cuadrática de las desviaciones en relación con la media o, como se le llama en forma común, *desviación de la media cuadrática*.

Ejemplo:

Busque la desviación estándar ***s*** para el siguiente conjunto de números:

9, 3, 8, 8, 9, 8, 9, 18

Primero obtenemos la media aritmética

$\overline{X}$= 9+3+8+8+9+8+9+18/8

$\overline{X}$= 72/8 = 9

Aplicamos la fórmula de la desviación estándar

$$= \sqrt{\frac{\sum (x - \bar{x})^2}{N}}$$

Sustituimos:

$$= \sqrt{\frac{(9-9)^2 + (3-9)^2 + (8-9)^2 + (8-9)^2 + (9-9)^2 + (8-9)^2 + (9-9)^2 + (18-9)^2}{8}}$$

$$= \sqrt{15} = 3.87$$

<u>Varianza:</u>

La varianza de un conjunto de datos se define como el cuadrado de la desviación estándar, por lo tanto, se representa como S^2.

Cuando es necesario distinguir la desviación estándar de una población de la desviación estándar de una muestra obtenida de dicha población, con frecuencia se utiliza el símbolo s para esta última y σ (sigma griega minúscula) para la primera. Por lo tanto S^2 y σ representan la *varianza de una muestra* y la *varianza de una población*, respectivamente.

Coeficiente de variación:

Es una medida de confiabilidad de los datos y es mejor cuando es pequeño. Se representa de la siguiente manera:

$$c.v. = \frac{s}{\bar{x}} \times 100$$

Ejemplo: el coeficiente de variación de los días del cuadro ulceroso es:

S= 7.5023805

$\overline{X}$= 5.5

$$c.v. = \frac{7.5023805}{5.5} \times 100 = 136\%$$

Referencia:

Spiegen M. y Stepens L. (2002). Estadística. México: McGraw-Hill.

Matematicas profe Alex. (2017). Cuartiles y Rango Intercuartílico para datos agrupados en intervalos. Recuperado a partir de: https://www.youtube.com/watch?v=w1IyL5WCr-0