



NOMBRE: ALBERTO BERMUDEZ TRUJILLO

ESTADISTICA

PLATAFORMA

SUPER NOTA

Las medidas de dispersión o variabilidad son herramientas estadísticas que permiten evaluar el grado de dispersión o extensión de los valores de un conjunto de datos respecto a su tendencia central. Estas medidas nos dan información sobre cómo se distribuyen los datos, si están concentrados alrededor de la media o si se dispersan ampliamente. Algunas de las medidas más comunes de dispersión son la desviación estándar, la varianza, el rango, el rango intercuartílico, y la varianza intercuartílica.

Desviación Estándar (σ): Es una medida que nos indica la cantidad promedio en la que los valores de un conjunto de datos se desvían respecto a la media. Cuanto mayor es la desviación estándar, más dispersos están los datos; si la desviación estándar es pequeña, los datos están más agrupados. La fórmula para calcular la desviación estándar es:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$



$\sigma =$

N

1

$i=1$

Σ

N

$(x$

i

$-\mu)$

2



Donde:

x

i

x

i

son los valores de la muestra.

μ

μ es la media de los datos.

N

N es el número total de datos.

Varianza (σ^2): Es el cuadrado de la desviación estándar. La varianza proporciona la misma información que la desviación estándar, pero en términos cuadráticos, lo que puede dificultar su interpretación directa. Se calcula mediante la siguiente fórmula:

σ

2

=

1

N

\sum

i

=

1

N

(

x

i

-

μ

)

2

σ

2

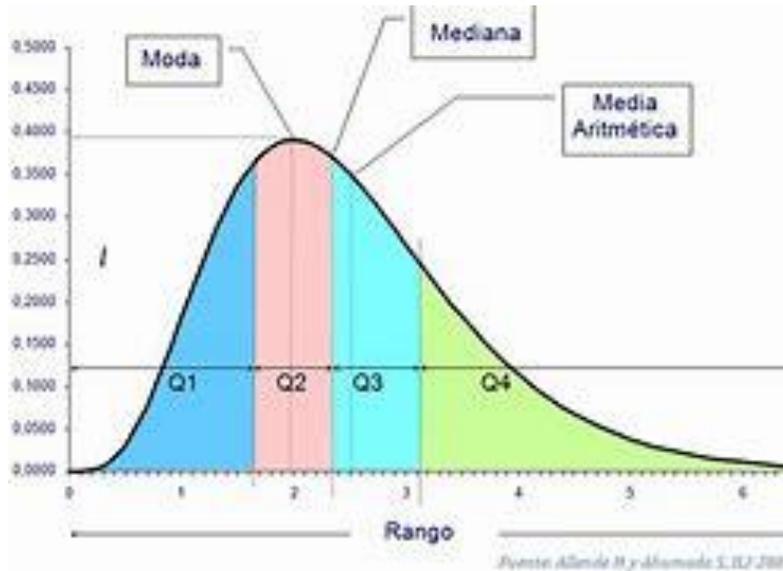
=

N

1

$i=1$

\sum



N

(x

i

-μ)

2

Rango: Es la diferencia entre el valor máximo y el valor mínimo de un conjunto de datos. Aunque es una medida simple, no siempre refleja la dispersión de manera adecuada, ya que solo depende de los valores extremos.

Rango Intercuartílico (IQR): Representa la distancia entre el primer cuartil (Q1) y el tercer cuartil (Q3). Es útil para describir la dispersión de los datos en términos de los valores centrales y está menos influenciado por los valores atípicos.

Medidas de Posición Central

Las medidas de posición central son estadísticas que nos indican el valor central o representativo de un conjunto de datos. Las tres principales son la media, la mediana y la moda. Cada una tiene sus propias ventajas y desventajas, dependiendo de la distribución de los datos:

Media (μ): Es la medida de posición central más común. Se calcula sumando todos los valores y dividiendo el resultado entre el número total de elementos. La media se ve afectada por valores extremos (outliers), por lo que no siempre refleja con precisión el valor central cuando hay mucha variabilidad.

μ

=

\sum

i

=

1

N

x

i

N

$\mu =$

N

Σ

$i=1$

N

x

i

Mediana (M): Es el valor que ocupa la posición central en un conjunto de datos ordenados. Si el número de observaciones es impar, la mediana es el valor central; si es par, es el promedio de los dos valores centrales. A diferencia de la media, la mediana no se ve afectada por valores extremos.

Moda: Es el valor que aparece con mayor frecuencia en un conjunto de datos. Un conjunto de datos puede tener una o más modas (si hay más de una frecuencia máxima). La moda es útil especialmente en variables cualitativas.

Coefficiente de Variación Estándar

El coeficiente de variación (CV) es una medida relativa de la dispersión, que nos ayuda a comparar la variabilidad de diferentes conjuntos de datos, incluso si tienen unidades de medida diferentes. Se calcula como la relación entre la desviación estándar y la media, expresada como un porcentaje. Su fórmula es:

C

V

=

σ

μ

\times

100

CV=

μ

σ

$\times 100$

El coeficiente de variación es útil cuando se desea comparar la variabilidad de datos de diferentes magnitudes. Un CV más alto indica una mayor variabilidad en relación con la media, mientras que un CV bajo sugiere una variabilidad relativamente baja.

En resumen, las medidas de dispersión y variabilidad son fundamentales para entender la extensión y distribución de los datos, mientras que las medidas de posición central nos dan una idea clara del valor promedio o representativo del conjunto. El coeficiente de variación estándar es una herramienta valiosa para comparar la dispersión relativa entre diferentes conjuntos de datos.

