Varianza y Desviación Estándar

Varianza

La varianza mide la dispersión o variabilidad de los datos respecto a la media. Existen dos tipos principales:

1. **Varianza Poblacional** (σ^2):

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

   - x_i: Cada valor de la población.
   - μ: Media de la población.
   - N: Tamaño de la población.

2. **Varianza Muestral** (s^2):

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

   - x_i: Cada valor de la muestra.
   - $\bar{x}$: Media de la muestra.
   - n: Tamaño de la muestra.

Utilidad de la Varianza

La varianza es útil para:

- **Medir la dispersión:** Cuantifica qué tan dispersos están los datos respecto a la media.
- **Comparar variabilidad:** Ayuda a determinar qué conjunto de datos es más variable.
- **Describir incertidumbre:** Es clave en estadísticas inferenciales para entender el comportamiento de los datos.

Ejemplo: Varianza Poblacional

Datos: $160, 165, 170, 175, 180$

1. Calcular la media:

$$\mu = \frac{160 + 165 + 170 + 175 + 180}{5} = 170$$

2. Calcular las desviaciones al cuadrado:
   - $(160 - 170)^2 = 100$
   - $(165 - 170)^2 = 25$
   - $(170 - 170)^2 = 0$
   - $(175 - 170)^2 = 25$
   - $(180 - 170)^2 = 100$

3. Calcular la varianza:

$$\sigma^2 = \frac{100 + 25 + 0 + 25 + 100}{5} = 50$$

Resultado: La varianza poblacional es $50\,\text{cm}^2$.

Ejemplo: Varianza Muestral

Datos: $160, 165, 170, 175$

1. Calcular la media:

$$\bar{x} = \frac{160 + 165 + 170 + 175}{4} = 167.5$$

2. Calcular las desviaciones al cuadrado:
   - $(160 - 167.5)^2 = 56.25$
   - $(165 - 167.5)^2 = 6.25$
   - $(170 - 167.5)^2 = 6.25$
   - $(175 - 167.5)^2 = 56.25$

3. Calcular la varianza:

$$s^2 = \frac{56.25 + 6.25 + 6.25 + 56.25}{4 - 1} = 41.67$$

Resultado: La varianza muestral es $41.67\,\text{cm}^2$.

Desviación Estándar

La desviación estándar mide la dispersión promedio de los datos respecto a la media. Se obtiene como la raíz cuadrada de la varianza:

1. **Poblacional:** $\sigma = \sqrt{\sigma^2}$
2. **Muestral:** $s = \sqrt{s^2}$

Utilidad de la Desviación Estándar

Es útil para:

- **Medir la dispersión:** Indica cuánto varían los datos respecto a la media.
- **Comparar conjuntos de datos:** Evalúa qué conjunto es más consistente.
- **Interpretar distribuciones:** En distribuciones normales, se pueden analizar intervalos de confianza.

Ejemplo: Desviación Estándar Poblacional

Datos: $10, 12, 14, 16, 18$

1. Calcular la media:

$$\mu = \frac{10 + 12 + 14 + 16 + 18}{5} = 14$$

2. Calcular las desviaciones al cuadrado:
   - $(10 - 14)^2 = 16$
   - $(12 - 14)^2 = 4$
   - $(14 - 14)^2 = 0$
   - $(16 - 14)^2 = 4$
   - $(18 - 14)^2 = 16$

3. Calcular la desviación estándar:

$$\sigma = \sqrt{\frac{16 + 4 + 0 + 4 + 16}{5}} = 2.83$$

Conclusión

- La varianza y la desviación estándar son medidas esenciales para entender la variabilidad en los datos.
- Permiten comparar y analizar diferentes conjuntos de datos.
- Son herramientas fundamentales en estadística descriptiva e inferencial.