Estadística I

Capítulo 2. Medidas de posición y dispersión

Carmen Trueba Salas

Lorena Remuzgo Pérez

Vanesa Jordá Gil

José María Sarabia Alegría

DPTO. DE ECONOMÍA

Medidas de posición

Las medidas de posición proporcionan información resumida de la variable objeto de estudio.

Medidas de posición centrales

- Media (aritmética, geométrica y armónica)
- Mediana
- Moda

Medidas de posición no centrales

- Cuantiles (cuartiles, deciles y percentiles)

Media aritmética

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \ldots + x_i n_i + \ldots + x_k n_k}{N} = \frac{1}{N} \sum_{i=1}^{k} x_i n_i$$

❖ Media aritmética ponderada

$$\bar{x}_w = \frac{w_1 x_1 + w_2 x_2 + \ldots + w_k x_k}{w}$$

❖ Media de la composición de poblaciones

$$\bar{x}_p = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2 + \ldots + N_k \bar{x}_k}{N}$$

Media geométrica

$$\bar{x}_g = \sqrt[N]{x_1^{n_1} \cdots x_k^{n_k}} = (x_1^{n_1} \cdots x_k^{n_k})^{1/N}$$

Media armónica

$$\bar{x}_a = \frac{N}{\dfrac{n_1}{x_1} + \dfrac{n_2}{x_2} + \ldots + \dfrac{n_k}{x_k}}$$

Mediana

1. Ordenar los datos de menor a mayor
2. Calcular $N/2$
3. Obtener las frecuencias absolutas acumuladas N_i

4. Datos no agrupados en intervalos

Buscar el primer valor x_i de la variable cuya $N_i \geq N/2$

4.1) $N_i = N/2 \rightarrow Me = \dfrac{x_i + x_{i+1}}{2}$

4.2) $N_i > N/2 \rightarrow Me = x_i$

Mediana

1. Ordenar los datos de menor a mayor
2. Calcular $N/2$
3. Obtener las frecuencias absolutas acumuladas N_i

4. Datos agrupados en intervalos

Buscar el primer intervalo (intervalo mediano) cuya $N_i \geq N/2$

4.1) $N_i = N/2 \quad \rightarrow \quad Me = L_i$

4.2) $N_i > N/2 \quad \rightarrow \quad Me = L_{i-1} + \dfrac{N/2 - N_{i-1}}{n_i} \cdot c_i$

Moda

a. Datos no agrupados en intervalos

La moda es el valor de la variable con mayor frecuencia.

b. Datos agrupados en intervalos

b.1) Misma longitud

Buscar el intervalo (intervalo modal) cuya frecuencia sea mayor,

$$Mo = L_{i-1} + \frac{n_{i+1}}{n_{i-1} + n_{i+1}} \cdot c_i$$

b.2) Distinta longitud

Buscar el intervalo (intervalo modal) cuya densidad sea mayor,

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot c_i$$

Cuantiles

1. Ordenar los datos de menor a mayor
2. Calcular $(r/k)\cdot N$
3. Obtener las frecuencias absolutas acumuladas N_i

4. Datos no agrupados en intervalos

Buscar el primer valor x_i de la variable cuya $N_i \geq \frac{r}{k}\cdot N$

4.1) $N_i = \frac{r}{k}\cdot N \quad \rightarrow \quad C_{r/k} = \frac{x_i + x_{i+1}}{2}$

4.2) $N_i > \frac{r}{k}\cdot N \quad \rightarrow \quad C_{r/k} = x_i$

Cuantiles

1. Ordenar los datos de menor a mayor
2. Calcular $(r/k)\cdot N$
3. Obtener las frecuencias absolutas acumuladas N_i

4. Datos agrupados en intervalos

Buscar el primer intervalo (intervalo cuantílico) cuya $N_i \geq \frac{r}{k}\cdot N$

4.1) $N_i = \frac{r}{k}\cdot N \quad \rightarrow \quad C_{r/k} = L_i$

4.2) $N_i > \frac{r}{k}\cdot N \quad \rightarrow \quad C_{r/k} = L_{i-1} + \frac{(r/k)N - N_{i-1}}{n_i}\cdot c_i$

Medidas de dispersión

Las medidas de dispersión estudian la separación existente entre los valores que toma la variable.

- **Medidas de dispersión absolutas**
  - Rango
  - Recorrido intercuartílico
  - Desviación absoluta media respecto a la media
  - Varianza
  - Desviación típica
- **Medidas de dispersión relativas**
  - Coeficiente de apertura
  - Recorrido relativo
  - Recorrido semi-intercuartílico
  - Coeficiente de variación
  - Variable tipificada

Rango

$$R = \text{máx}\{x_1, \ldots, x_k\} - \text{mín}\{x_1, \ldots, x_k\}$$

Cuanto más pequeño sea → menor dispersión.

Recorrido intercuartílico

$$R_I = Q_3 - Q_1$$

Cuanto más pequeño sea → menor dispersión.

Desviación absoluta media

$$D_{\bar{x}} = \frac{1}{N} \sum_{i=1}^{k} | x_i - \bar{x} | n_i$$

Si es pequeña (en relación a la media) → los valores de la variable se distribuyen en torno a ella.

Varianza

$$S^2 = \frac{1}{N}\sum_{i=1}^{k}(x_i - \overline{x})^2 n_i = \frac{1}{N}\sum_{i=1}^{k} x_i^2 n_i - \overline{x}^2$$

Si es pequeña (dispersión pequeña) → la media es una buena representación de las observaciones de la variable

Desviación típica

$$S = +\sqrt{S^2}$$

Coeficiente de apertura

$$A = \frac{\text{máx}\{x_1, \ldots, x_k\}}{\text{mín}\{x_1, \ldots, x_k\}}$$

Si es pequeño (en relación al nº de observaciones) → existe poca dispersión.

Recorrido relativo

$$R_R = \frac{\text{máx}\{x_1, \ldots, x_k\} - \text{mín}\{x_1, \ldots, x_k}{\overline{x}}$$

Si es pequeño (en relación al nº de observaciones) → existe poca dispersión.

Recorrido semi-intercuartílico

$$R_S = \frac{Q_3 - Q_1}{Q_1 + Q_3}$$

Si es pequeño (en relación al nº de observaciones) → existe poca dispersión.

Coeficiente de variación

$$CV = \frac{S}{|\bar{x}|}$$

Cuanto menor es el coeficiente de variación → menor es la dispersión relativa.

Variable tipificada

$$Z = \frac{X - \overline{x}}{S}$$

Los valores tipificados permiten comparar características de dos individuos que pertenecen a poblaciones diferentes.