

Tema 3: Medidas de posición

Estadística I

Universidad de Salamanca

Curso 2010/2011

Outline

- 1 Medidas de tendencia central
 - Media
 - Moda
 - Mediana

- 2 Medidas de posición no centrales
 - Cuartiles
 - Deciles
 - Percentiles

Outline

1 Medidas de tendencia central

- Media
- Moda
- Mediana

2 Medidas de posición no centrales

- Cuartiles
- Deciles
- Percentiles

Media aritmética

Definición: $\bar{X}$

$$\bar{X} = \frac{\sum_{i=1}^N x_i}{N} = \frac{\sum_{i=1}^k x_i n_i}{N} = \sum_{i=1}^k x_i f_i$$

Propiedades

- Es única
- No tiene porque ser un valor observado de la variable
- En su calculo intervienen todos los datos
- Sea X con media $\bar{X}$ entonces

$$Y = aX + b \implies \bar{Y} = a\bar{X} + b$$

Media aritmética

Propiedades

- Sea Y y $X_1, X_2, \dots, X_n$ variables donde

$$Y = a_1X_1 + \dots + a_nX_n + b \implies \bar{Y} = a_1\bar{X}_1 + \dots + a_n\bar{X}_n + b$$

- La suma de las desviaciones con respecto a la media es cero. Sea $d_i = x_i - \bar{X}$ la desviación de x_i con respecto de la media $\bar{X}$

$$\sum_{i=1}^N d_i = 0$$

Media aritmética

Propiedades

- Si la población objeto de estudio se divide en H partes: $N^1, N_2, \dots, N^H = N$. La media de una variable en toda la población es

$$\bar{X} = \frac{\bar{X}^1 N^1 + \dots + \bar{X}^H N^H}{N}$$

siendo $\bar{X}^1, \dots, \bar{X}^H$ las medias de esa variable en cada subpoblación

Media ponderada

Definición: $\bar{X}_p$

$$\bar{X}_p = \frac{\sum_{i=1}^k x_i w_i}{\sum_{i=1}^k w_i}$$

Propiedades

- Generalización de la media aritmética

Media geométrica

Definición: G

$$G = \sqrt[n]{x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}}$$

Propiedades

- Es única
- Utiliza todos los elementos
- Sólo se puede aplicar a variables que tomen valores positivos

Media armónica

Definición: H

$$H = \frac{\sum_{i=1}^k n_i}{\sum_{i=1}^k \frac{n_i}{x_i}} = \frac{N}{\sum_{i=1}^k \frac{n_i}{x_i}}$$

Propiedades

- Utiliza todos los elementos
- Sólo se puede aplicar a variables que tomen valores positivos

Outline

1 Medidas de tendencia central

- Media
- **Moda**
- Mediana

2 Medidas de posición no centrales

- Cuartiles
- Deciles
- Percentiles

Moda

Definición: Mo

Es aquel valor de la variable que más veces se repite o que la frecuencia sea máxima

Propiedades

- No es única
- Siempre es un valor observado de la variable
- En su calculo no intervienen todos los datos
- Sea X con media $\bar{X}$ entonces

$$Y = aX + b \implies Mo_Y = a Mo_X + b$$

Moda

Cálculo para datos sin agrupar

x_i	n_i
x_1	n_1
$\vdots$	$\vdots$
x_k	n_k
Total	N

Mo_X

La moda de la variable X es el x_i con mayor n_i

Moda

Cálculo para datos agrupados

l_i	n_i	c_i	h_i
$L_0 - L_1$	n_1	c_1	h_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$L_{k-1} - L_k$	n_k	c_k	h_k

Pasos para intervalos

- Seleccionamos el intervalo con máxima h_i , el **intervalo modal**: $l_i = L_{i-1} - L_i$

$$Mo_x = L_{i-1} + c_i \left(\frac{h_{i+1}}{h_{i-1} + h_{i+1}} \right)$$

- Cuando el intervalo de mayor altura es el primero, $h_{i-1} = 0$
- Cuando el intervalo de mayor altura es el último, $h_{i+1} = 0$

Outline

1 Medidas de tendencia central

- Media
- Moda
- **Mediana**

2 Medidas de posición no centrales

- Cuartiles
- Deciles
- Percentiles

Mediana

Definición: *Me*

Es aquel valor de la variable que divide a la distribución de frecuencias en dos partes iguales, ordenando previamente los datos

Propiedades

- Es única
- En su calculo no intervienen todos los datos
- Siempre es un valor observado de la variable
- Sea X con media $\bar{X}$ entonces

$$Y = aX + b \implies Me_Y = a Me_X + b$$

Mediana para datos no agrupados

Cálculo para datos no agrupados

x_i	n_i	N_i	F_i
x_1	n_1	N_1	F_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	N_k	1

Pasos

- Ordenamos los datos de menor a mayor
- Calculamos frecuencias absolutas y relativas acumuladas

Mediana para datos no agrupados

Con frecuencias absolutas acumuladas

$$Me_X = \begin{cases} x_i & \text{si } N_{i-1} < \frac{N}{2} < N_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } N_i = \frac{N}{2} \end{cases}$$

Con frecuencias relativas acumuladas

$$Me_X = \begin{cases} x_i & \text{si } F_{i-1} < 0,5 < F_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } F_i = 0,5 \end{cases}$$

Mediana para datos agrupados

Cálculo para datos agrupados

l_i	n_i	c_i	h_i	N_i	F_i
$L_0 - L_1$	n_1	c_1	h_1	N_1	F_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$L_{k-1} - L_k$	n_k	c_k	h_k	N_k	F_k

Pasos

- Ordenamos los intervalos de menor a mayor
- Calculamos frecuencias absolutas y relativas acumuladas

Mediana para datos agrupados

Con frecuencias absolutas acumuladas

$$Me_X = \begin{cases} L_{i-1} + c_i \left(\frac{\frac{N}{2} - N_{i-1}}{n_i} \right) & \text{si } N_{i-1} < \frac{N}{2} \leq N_i \\ L_i & \text{si } N_i = \frac{N}{2} \end{cases}$$

Con frecuencias relativas acumuladas

$$Me_X = \begin{cases} L_{i-1} + c_i \left(\frac{0,5 - F_{i-1}}{f_i} \right) & \text{si } F_{i-1} < 0,5 \leq F_i \\ L_i & \text{si } F_i = 0,5 \end{cases}$$

Outline

1 Medidas de tendencia central

- Media
- Moda
- Mediana

2 Medidas de posición no centrales

- Cuartiles
- Deciles
- Percentiles

Cuartiles

Definición: Q_j , $j = 1, 2, 3$

Son los valores de la variable que dividen la distribución en cuatro partes iguales. Su calculo es análogo al de la mediana

$$Q_2 = Me$$

Pasos

- Ordenar los datos de menor a mayor
- Calculamos frecuencias absolutas y relativas acumuladas

Cuartiles para datos no agrupados

Con frecuencias absolutas acumuladas

$$Q_j = \begin{cases} x_i & \text{si } N_{i-1} < \frac{jN}{4} < N_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } N_i = \frac{jN}{4} \end{cases}$$

Con frecuencias relativas acumuladas

$$Q_j = \begin{cases} x_i & \text{si } F_{i-1} < \frac{j}{4} < F_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } F_i = \frac{j}{4} \end{cases}$$

Cuartiles para datos agrupados

Con frecuencias absolutas acumuladas

$$Q_j = L_{i-1} + c_i \left(\frac{\frac{jN}{4} - N_{i-1}}{n_i} \right) \quad \text{si } N_{i-1} < \frac{jN}{4} \leq N_i$$

Con frecuencias relativas acumuladas

$$Q_j = L_{i-1} + c_i \left(\frac{\frac{j}{4} - F_{i-1}}{f_i} \right) \quad \text{si } F_{i-1} < \frac{j}{4} \leq F_i$$

Outline

1 Medidas de tendencia central

- Media
- Moda
- Mediana

2 Medidas de posición no centrales

- Cuartiles
- **Deciles**
- Percentiles

Deciles

Definición: $D_j, j = 1, \dots, 9$

Son los valores de la variable que dividen en diez partes iguales a la distribución

$$D_5 = Q_2 = Me$$

Pasos

- Ordenar los datos de menor a mayor
- Calculamos frecuencias absolutas y relativas acumuladas

Deciles para datos no agrupados

Con frecuencias absolutas acumuladas

$$D_j = \begin{cases} x_i & \text{si } N_{i-1} < \frac{jN}{10} < N_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } N_i = \frac{jN}{10} \end{cases}$$

Con frecuencias relativas acumuladas

$$D_j = \begin{cases} x_i & \text{si } F_{i-1} < \frac{j}{10} < F_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } F_i = \frac{j}{10} \end{cases}$$

Deciles para datos agrupados

Con frecuencias absolutas acumuladas

$$D_j = L_{i-1} + c_i \left(\frac{\frac{jN}{10} - N_{i-1}}{n_i} \right) \quad \text{si } N_{i-1} < \frac{jN}{10} \leq N_i$$

Con frecuencias relativas acumuladas

$$D_j = L_{i-1} + c_i \left(\frac{\frac{j}{10} - F_{i-1}}{f_i} \right) \quad \text{si } F_{i-1} < \frac{j}{10} \leq F_i$$

Outline

1 Medidas de tendencia central

- Media
- Moda
- Mediana

2 Medidas de posición no centrales

- Cuartiles
- Deciles
- **Percentiles**

Percentiles

Definición: P_j , $j = 1, \dots, 99$

Son los valores de la variable que dividen en cien partes iguales a la distribución

Pasos

- Ordenar los datos de menor a mayor
- Calculamos frecuencias absolutas y relativas acumuladas

Percentiles para datos no agrupados

Con frecuencias absolutas acumuladas

$$P_j = \begin{cases} x_i & \text{si } N_{i-1} < \frac{jN}{100} < N_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } N_i = \frac{jN}{100} \end{cases}$$

Con frecuencias relativas acumuladas

$$P_j = \begin{cases} x_i & \text{si } F_{i-1} < \frac{j}{100} < F_i \\ \frac{x_i + x_{i+1}}{2} & \text{si } F_i = \frac{j}{100} \end{cases}$$

Percentiles para datos agrupados

Con frecuencias absolutas acumuladas

$$P_j = L_{i-1} + c_i \left(\frac{\frac{jN}{100} - N_{i-1}}{n_i} \right) \quad \text{si } N_{i-1} < \frac{jN}{100} \leq N_i$$

Con frecuencias relativas acumuladas

$$P_j = L_{i-1} + c_i \left(\frac{\frac{j}{100} - F_{i-1}}{f_i} \right) \quad \text{si } F_{i-1} < \frac{j}{100} \leq F_i$$