

Tema 2: Estadística descriptiva univariante

Estadística I

Universidad de Salamanca

Curso 2010/2011

Outline

- 1 Tablas de distribución de frecuencias
 - Definición de frecuencias
 - Tablas de frecuencias

- 2 Representaciones gráficas
 - Representaciones gráficas de variables cualitativas
 - Representaciones gráficas de variables cuantitativas discretas
 - Representaciones gráficas de variables cuantitativas continuas

Outline

- 1 **Tablas de distribución de frecuencias**
 - Definición de frecuencias
 - Tablas de frecuencias

- 2 **Representaciones gráficas**
 - Representaciones gráficas de variables cualitativas
 - Representaciones gráficas de variables cuantitativas discretas
 - Representaciones gráficas de variables cuantitativas continuas

Definición de frecuencias

Recordamos que tenemos N individuos descritos según una variable X cuyos valores ordenados de mayor a menor son $x_1, x_2, \dots, x_k$ (número de valores distintos o clases)

Example

- Población: número de vecinos de un portal, $N = 10$
- Variable X : número de hijos
- Datos:

1, 2, 2, 1, 14, 2, 2, 1, 1, 2

- Datos ordenados de mayor a menor:

1, 1, 1, 1, 2, 2, 2, 2, 2, 14

$$x_1 = 1, x_2 = 2, x_3 = 14$$

Frecuencia absoluta: n_i

Frecuencia absoluta del valor x_i

Número de veces que se repite el valor x_i

Propiedades

- $n_i \geq 0, i = 1, \dots, k$
- $n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_i = N$

Frecuencia absoluta: n_i

Example

Valores de x_i	Frecuencia absoluta: n_i
$x_1 = 1$	$n_1 = 4$
$x_2 = 2$	$n_2 = 5$
$x_3 = 14$	$n_3 = 1$
Total	$N = 10$

Frecuencia relativa: f_i

Frecuencia relativa del valor x_i

Proporción de veces que se repite el valor x_i

$$f_i = \frac{n_i}{N}$$

Propiedades

- $0 \leq f_i \leq 1, i = 1, \dots, k$
- $f_1 + f_2 + \dots + f_k = \sum_{i=1}^k f_i = 1$

Frecuencia relativa: f_i

Example

Valores de x_i	F. absoluta: n_i	F. relativa: f_i
$x_1 = 1$	$n_1 = 4$	$f_1 = 0,4$
$x_2 = 2$	$n_2 = 5$	$f_2 = 0,5$
$x_3 = 14$	$n_3 = 1$	$f_3 = 0,1$
Total	$N = 10$	1

Frecuencia absoluta acumulada: N_i

Frecuencia absoluta acumulada del valor x_i

Cantidad de datos que hay menores o iguales a x_i

Propiedades

- $n_1 = N_1$
- $N_2 = n_1 + n_2 = \sum_{j=1}^2 n_j = N_1 + n_2$
- $\vdots$
- $N_i = n_1 + n_2 + \dots + n_i = \sum_{j=1}^i n_j = N_{i-1} + n_i$
- $\vdots$
- $N_k = n_1 + n_2 + \dots + n_k = \sum_{j=1}^k n_j = N_{k-1} + n_k = N$

Frecuencia absoluta acumulada: N_i

Example

x_i	n_i	f_i	N_i
$x_1 = 1$	$n_1 = 4$	$f_1 = 0,4$	$N_1 = 4$
$x_2 = 2$	$n_2 = 5$	$f_2 = 0,5$	$N_2 = 9$
$x_3 = 14$	$n_3 = 1$	$f_3 = 0,1$	$N_3 = 10$
Total	$N = 10$	1	

N_2 : Número de vecinos con hasta 2 hijos

Frecuencia relativa acumulada: F_i

Frecuencia relativa acumulada del valor x_i

Es la proporción de datos que hay menores o iguales a x_i

$$F_i = \frac{N_i}{N}$$

Propiedades

- $F_1 = \frac{N_1}{N} = f_1$
- $F_2 = \frac{N_2}{N} = f_1 + f_2 = \sum_{j=1}^2 f_j = F_1 + f_2$
- $\vdots$
- $F_i = \frac{N_i}{N} = f_1 + f_2 + \dots + f_i = \sum_{j=1}^i f_j = F_{i-1} + f_i$
- $\vdots$
- $F_k = \frac{N_k}{N} = f_1 + f_2 + \dots + f_k = \sum_{j=1}^k f_j = F_{k-1} + f_k = 1$

Frecuencia relativa acumulada: F_i

Example

x_i	n_i	f_i	N_i	F_i
$x_1 = 1$	$n_1 = 4$	$f_1 = 0,4$	$N_1 = 4$	$F_1 = 0,4$
$x_2 = 2$	$n_2 = 5$	$f_2 = 0,5$	$N_2 = 9$	$F_2 = 0,9$
$x_3 = 14$	$n_3 = 1$	$f_3 = 0,1$	$N_3 = 10$	$F_3 = 1$
Total	$N = 10$	1		

F_2 : Proporción de vecinos con hasta 2 hijos

Outline

- 1 **Tablas de distribución de frecuencias**
 - Definición de frecuencias
 - **Tablas de frecuencias**

- 2 **Representaciones gráficas**
 - Representaciones gráficas de variables cualitativas
 - Representaciones gráficas de variables cuantitativas discretas
 - Representaciones gráficas de variables cuantitativas continuas

Tabla de distribución de frecuencias

Definición

Conjunto de valores que toma una variable junto con sus respectivas frecuencias

Tipos

- Tablas de frecuencias para variables discretas y atributos ordinales
- Tablas de frecuencias para variables o atributos nominales
- Tablas de frecuencias para variables continuas

Tabla de distribución de frecuencias

Tablas de frecuencias para variables o atributos nominales

En este caso la variable presenta datos sin agrupar en intervalos y perfectamente especificados pero sólo tiene sentido calcular las frecuencias acumuladas cuando las variables se pueden ordenar. No deben usarse en el caso de variables cualitativas nominales.

Tablas de frecuencias para variables continuas

En este caso interesa agrupar los valores en intervalos (**Intervalos de clase**) para atribuir un significado coherente a los datos o para simplificar el problema

Tabla de distribución de frecuencias para variables continuas

Definiciones

- **Intervalo** i -ésimo: $(L_{i-1}, L_i]$
- **Límite inferior** del intervalo i -ésimo: L_{i-1}
- **Límite superior** del intervalo i -ésimo: L_i
- **Marca de clase**: Punto central o medio de cada intervalo:

$$x_i = \frac{L_{i-1} + L_i}{2}$$

- **Amplitud del intervalo**: Diferencia entre el extremo superior y el inferior del intervalo

$$c_i = L_i - L_{i-1}$$

Tabla de distribución de frecuencias para variables continuas

Definiciones

- **Frecuencia absoluta** del intervalo i -ésimo: n_i
- **Frecuencia relativa** del intervalo i -ésimo: f_i
- **Frecuencia absoluta acumulada** del límite superior del intervalo i -ésimo: N_i
- **Frecuencia relativa acumulada** del límite superior del intervalo i -ésimo: F_i
- **Altura** del intervalo i -ésimo: h_i
 - Absoluta: $h_i = \frac{n_i}{c_i}$
 - Relativa: $h_i = \frac{f_i}{c_i}$
- **Recorrido** de la variable: $R = x_{max} - x_{min}$

Tabla de distribución de frecuencias para variables continuas

Intervalos

- La cantidad de intervalos depende del elemento a estudiar
- Cuanto menos intervalos haya mejor
- Para saber cuantos intervalos hacemos: $\sqrt{N}$
- Si los intervalos tienen que tener la misma amplitud, esta se calcula:

$$c_i = \frac{R}{\sqrt{N}},$$

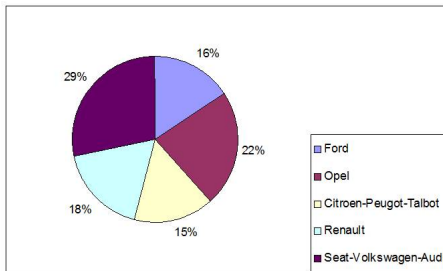
redondeando a la alza

Outline

- 1 Tablas de distribución de frecuencias
 - Definición de frecuencias
 - Tablas de frecuencias
- 2 Representaciones gráficas
 - Representaciones gráficas de variables cualitativas
 - Representaciones gráficas de variables cuantitativas discretas
 - Representaciones gráficas de variables cuantitativas continuas

Diagrama de sectores

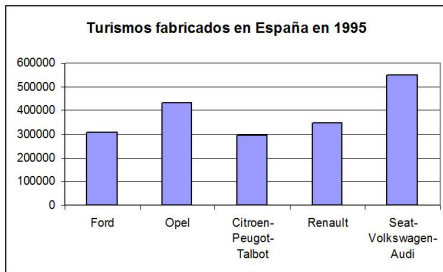
Se realizan dibujando sobre un círculo sectores de tal manera que el ángulo central sea proporcional al valor que represente



Representaciones gráficas de variables cualitativas

Diagrama de barras

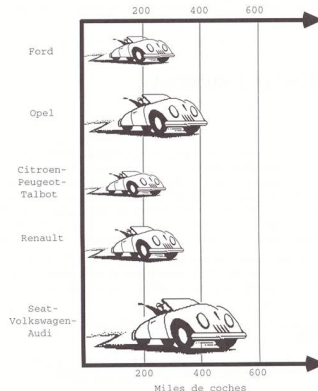
Cada modalidad de la variable está representada por un rectángulo cuya altura corresponde a su frecuencia absoluta o relativa



Representaciones gráficas de variables cualitativas

Pictograma

Cada modalidad de la variable está representada por un dibujo alusivo a su significado



Outline

1 Tablas de distribución de frecuencias

- Definición de frecuencias
- Tablas de frecuencias

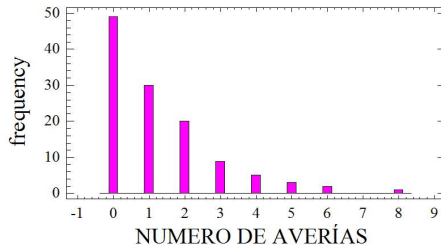
2 Representaciones gráficas

- Representaciones gráficas de variables cualitativas
- **Representaciones gráficas de variables cuantitativas discretas**
- Representaciones gráficas de variables cuantitativas continuas

Representaciones gráficas de variables cuantitativas discretas

Diagrama de barras

Cada modalidad de la variable está representada por un rectángulo cuya altura corresponde a su frecuencia absoluta o relativa. Se deja hueco entre barras para indicar los valores que no son posibles



Outline

- 1 Tablas de distribución de frecuencias
 - Definición de frecuencias
 - Tablas de frecuencias
- 2 Representaciones gráficas
 - Representaciones gráficas de variables cualitativas
 - Representaciones gráficas de variables cuantitativas discretas
 - Representaciones gráficas de variables cuantitativas continuas

Histograma

Es el más adecuado para representar variables cuantitativas agrupadas en intervalos

