

Tema 7: Estadística descriptiva bivalente

Estadística I

Universidad de Salamanca

Curso 2010/2011

Outline

- 1 Introducción
- 2 Tablas estadísticas bivariantes
- 3 Representaciones gráficas
 - Representaciones gráficas de tablas de doble entrada
 - Representaciones gráficas de datos individualizados
- 4 Distribuciones marginales y condicionadas
 - Distribuciones marginales
 - Distribuciones condicionadas

Introducción

En este tema estudiaremos las relaciones existentes entre dos variables

La naturaleza de las características objeto de estudio determinará las herramientas estadísticas más adecuadas para su análisis

En cualquier caso, para realizar el análisis utilizaremos tablas estadísticas, representaciones gráficas y resúmenes numéricos

Tablas estadísticas bivariantes

Frecuencia relativa conjunta: f_{ij}

Proporción de individuos que presentan simultáneamente el valor x_i y el valor y_j

$$f_{ij} = \frac{n_{ij}}{N}$$

Tabla doble de frecuencias relativas

	y_1	y_1	$\cdots$	y_p	$f_{i.}$
x_1	f_{11}	f_{12}	$\cdots$	f_{1p}	$f_{1.}$
x_2	f_{21}	f_{22}	$\cdots$	f_{2p}	$f_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	f_{k1}	f_{k2}	$\cdots$	f_{kp}	$f_{k.}$
$f_{.j}$	$f_{.1}$	$f_{.2}$	$\cdots$	$f_{.p}$	1

Outline

- 1 Introducción
- 2 Tablas estadísticas bivariantes
- 3 Representaciones gráficas**
 - Representaciones gráficas de tablas de doble entrada
 - Representaciones gráficas de datos individualizados
- 4 Distribuciones marginales y condicionadas
 - Distribuciones marginales
 - Distribuciones condicionadas

Representaciones gráficas

Diagrama de barras agrupadas

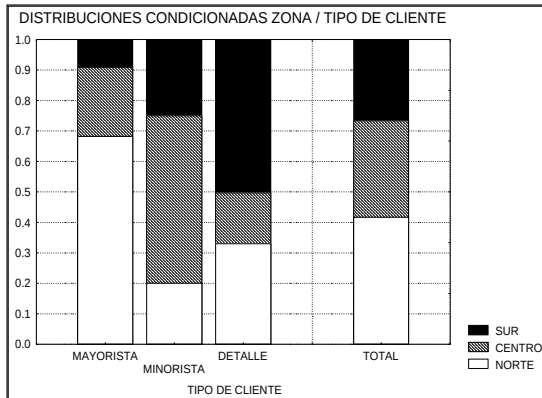
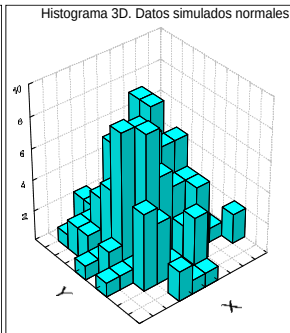
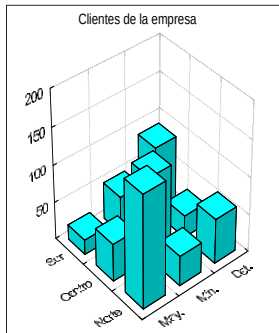


Diagrama de barras agrupadas



Representaciones gráficas

Histograma de frecuencias tridimensional

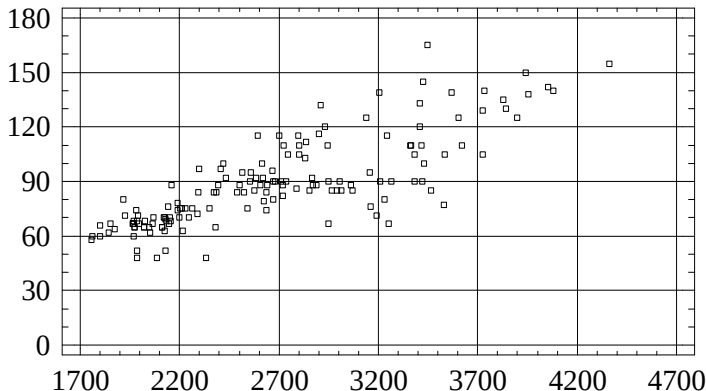


Outline

- 1 Introducción
- 2 Tablas estadísticas bivariantes
- 3 **Representaciones gráficas**
 - Representaciones gráficas de tablas de doble entrada
 - **Representaciones gráficas de datos individualizados**
- 4 Distribuciones marginales y condicionadas
 - Distribuciones marginales
 - Distribuciones condicionadas

Representaciones gráficas

Diagrama de dispersión o nube de puntos



Distribuciones marginales y condicionadas

A partir de **una variable bidimensional** podemos estudiar **dos tipos de variables unidimensionales**:

- **Distribuciones marginales**
- **Distribuciones condicionadas**

Outline

- 1 Introducción
- 2 Tablas estadísticas bivariantes
- 3 Representaciones gráficas
 - Representaciones gráficas de tablas de doble entrada
 - Representaciones gráficas de datos individualizados
- 4 **Distribuciones marginales y condicionadas**
 - **Distribuciones marginales**
 - Distribuciones condicionadas

Distribución marginal de X

Frecuencia absoluta marginal: $n_{i.}$

- $n_{i.} = n_{i1} + n_{i2} + \dots + n_{ip}$

$$n_{1.} = n_{11} + n_{12} + \dots + n_{1p}$$

$$n_{2.} = n_{21} + n_{22} + \dots + n_{2p}$$

$$\dots = \dots\dots\dots$$

$$n_{k.} = n_{k1} + n_{k2} + \dots + n_{kp}$$

- $n_{1.} + n_{2.} + \dots + n_{k.} = N$

Distribución marginal de X

Frecuencia relativa marginal: $f_{i.}$

- $f_{i.} = \frac{n_{i.}}{N}$
- $f_{i.} = f_{i1} + f_{i2} + \dots + f_{ip}$

$$f_{1.} = f_{11} + f_{12} + \dots + f_{1p}$$

$$f_{2.} = f_{21} + f_{22} + \dots + f_{2p}$$

$$\dots = \dots\dots\dots$$

$$f_{k.} = f_{k1} + f_{k2} + \dots + f_{kp}$$

- $f_{1.} + f_{2.} + \dots + f_{k.} = 1$

Distribución marginal de X

Media marginal: $\bar{X}$

$$\bar{X} = \frac{\sum_{i=1}^N x_i}{N} = \frac{\sum_{i=1}^k x_i n_{i.}}{N} = \sum_{i=1}^k x_i f_{i.}$$

Varianza marginal: S_X^2

$$S_X^2 = \frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N} = \frac{\sum_{i=1}^k (x_i - \bar{X})^2 n_{i.}}{N} = \sum_{i=1}^k (x_i - \bar{X})^2 f_{i.}$$

Distribución marginal de Y

Frecuencia absoluta marginal: $n_{.j}$

- $n_{.j} = n_{1j} + n_{2j} + \dots + n_{kj}$

$$n_{.1} = n_{11} + n_{21} + \dots + n_{k1}$$

$$n_{.2} = n_{12} + n_{22} + \dots + n_{k2}$$

$$\dots = \dots\dots\dots$$

$$n_{.p} = n_{1p} + n_{2p} + \dots + n_{kp}$$

- $n_{.1} + n_{.2} + \dots + n_{.k} = N$

Distribución marginal de Y

Frecuencia relativa marginal: $f_{.j}$

- $f_{.j} = \frac{n_{.j}}{N}$
- $f_{.j} = f_{1j} + f_{2j} + \dots + f_{kj}$

$$f_{.1} = f_{11} + f_{21} + \dots + f_{k1}$$

$$f_{.1} = f_{12} + f_{22} + \dots + f_{k2}$$

$$\dots = \dots\dots\dots$$

$$f_{.p} = f_{1p} + f_{2p} + \dots + f_{kp}$$

- $f_{.1} + f_{.1} + \dots + f_{.p} = 1$

Distribución marginal de Y

Media marginal: $\bar{Y}$

$$\bar{Y} = \frac{\sum_{j=1}^N y_j}{N} = \frac{\sum_{j=1}^p y_j n_{.j}}{N} = \sum_{j=1}^p y_j f_{.j}$$

Varianza marginal: S_Y^2

$$S_Y^2 = \frac{\sum_{j=1}^N (y_j - \bar{Y})^2}{N} = \frac{\sum_{j=1}^p (y_j - \bar{Y})^2 n_{.j}}{N} = \sum_{j=1}^p (x_j - \bar{Y})^2 f_{.j}$$

Outline

- 1 Introducción
- 2 Tablas estadísticas bivariantes
- 3 Representaciones gráficas
 - Representaciones gráficas de tablas de doble entrada
 - Representaciones gráficas de datos individualizados
- 4 Distribuciones marginales y condicionadas**
 - Distribuciones marginales
 - Distribuciones condicionadas**

Distribución de X condicionada a que $Y = y_j$

Frecuencia absoluta condicionada: $n_{i/j}$

- $n_{i/j} = n_{ij}$
- $\sum_{i=1}^k n_{i/j} = \sum_{i=1}^k n_{ij} = n_{.j}$

Frecuencia relativa condicionada: $f_{i/j}$

- $f_{i/j} = \frac{n_{ij}}{n_{.j}}$
- $\sum_{i=1}^k f_{i/j} = 1$

Distribución de Y condicionada a que $X = x_i$

Frecuencia absoluta condicionada: $n_{j/i}$

- $n_{j/i} = n_{ij}$
- $\sum_{j=1}^p n_{j/i} = \sum_{j=1}^p n_{ij} = n_i.$

Frecuencia relativa condicionada: $f_{j/i}$

- $f_{j/i} = \frac{n_{ij}}{n_i}.$
- $\sum_{j=1}^p f_{j/i} = 1$