

Aplicacions Estadístiques

Enginyeria Edificació 2009/10.

Antonio E. Teruel

Temari

- ▶ **Estadística Descriptiva**

- Tema 1. **Anàlisi exploratori de dades.**

- Tema 2. Distribucions estadístiques bidimensionals.

- ▶ Probabilitat.

- Tema 3. Teoria de la probabilitat.

- ▶ Estadística Inferencial.

- Tema 4. Variables aleatòries discretes.

- Tema 5. Variables aleatòries contínues.

- Tema 6. Estimació de paràmetres.

- Tema 7. Contrast d'hipotesis.

Representació de dades estadístiques

- ▶ Les dades obtingudes d'un estudi estadístic s'anomenen dades brutes.
- ▶ Aquestes dades s'organitzen en taules de freqüències.

x_i		x_i	n_i	N_i	f_i	F_i	p_i	P_i
x_1		x_1	n_1	N_1	f_1	F_1	p_1	P_1
x_2		x_2	n_2	N_2	f_2	F_2	p_2	P_2
x_3	\Rightarrow	x_3	n_3	N_3	f_3	F_3	p_3	P_3
\vdots		\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_n		x_k	n_k	$N_k = n$	f_k	$F_k = 1$	p_k	$P_k = 100$

Dades brutes

Taula de Freqüències

Representació de dades estadístiques

- ▶ Les taules de freqüència contenen informació de:
 - ▶ els valors de la variable: x_i
 - ▶ nombre de vegades que apareix cada valor (freqüència absoluta): n_i
 - ▶ el nombre total de valors: n
 - ▶ les freqüències absolutes acumulades: $N_i = n_1 + n_2 + \dots + n_i$
 - ▶ les freqüències relatives: $f_i = \frac{n_i}{n}$
 - ▶ les freqüències relatives acumulades: $F_i = \frac{N_i}{n}$
 - ▶ els percentatges absoluts: $p_i = f_i \times 100\%$
 - ▶ els percentatges acumulats: $P_i = F_i \times 100\%$
- ▶ En color **vermell** nomès per a variables quantitatives i ordinals.

Representació de dades estadístiques

Diagram illustrating the representation of statistical data, showing the transformation from raw data to a frequency table.

Raw Data (Dades brutes): A vertical list of n values: $x_i, x_1, x_2, x_3, \dots, x_n$.

Frequency Table (Taula de Freqüències): A table with 7 columns and k rows (plus a final summary row). The columns are labeled with their respective measures:

- freq.abs.acum.** (Cumulative Absolute Frequency): N_i
- freq.rel.acum.** (Cumulative Relative Frequency): F_i
- %abs.acum.** (Cumulative Absolute Percentage): P_i
- freq.abs** (Absolute Frequency): n_i
- freq.rel** (Relative Frequency): f_i
- % abs** (Absolute Percentage): p_i

The table structure is as follows:

x_i	n_i	N_i	f_i	F_i	p_i	P_i
x_1	n_1	N_1	f_1	F_1	p_1	P_1
x_2	n_2	N_2	f_2	F_2	p_2	P_2
x_3	n_3	N_3	f_3	F_3	p_3	P_3
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_k	n_k	$N_k = n$	f_k	$F_k = 1$	p_k	$P_k = 100$

En color **vermell** nomès per a variables quantitatives i ordinals.

Representació de dades estadístiques

- Exemple: nota d'estadística de 10 persones

x_i
7
5
9
7
5
6
7
6
4
7



x_i	n_i	N_i	f_i	F_i	p_i	P_i
4	1	1	0.1	0.1	10	10
5	2	3	0.2	0.3	20	30
6	2	5	0.2	0.5	20	50
7	4	9	0.4	0.9	40	90
9	1	10	0.1	1.0	10	100

$$n = 10 \quad f = 1$$

Dades brutes

Taula de Freqüències

Representació de dades estadístiques amb intervals

- ▶ Representació de dades amb intervals
 - ▶ només variables quantitatives i qualitatives ordinals
 - ▶ s'agrupen en intervals de valors
 - ▶ s'anomena marca de classe (m) al valor representatiu d'un interval i és igual al valor mitjà de l'interval

Representació de dades estadístiques

- ▶ Exemple: nota d'estadística de 20 persones

8.4	9.4
5.5	7.5
9.1	8.1
6.9	5.7
5.3	6.4
6.2	5.8
7.8	4.8
6.1	6.0
4.2	4.7
3.9	7.9

Dades brutes



x_i	m_i	n_i	N_i	f_i	F_i	p_i	P_i
$[0, 4)$	2	1	1	0.05	0.05	5	5
$[4, 5)$	4.5	3	4	0.15	0.20	15	20
$[5, 7)$	6	9	13	0.45	0.65	45	65
$[7, 9)$	8	5	18	0.25	0.90	25	90
$[9, 10)$	9.5	2	20	0.10	1.00	10	100

$n = 20$

$f = 1$

Taula de Freqüències

Representació gràfica de dades estadístiques

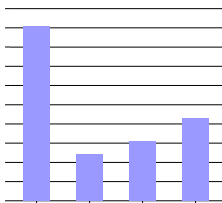


Diagrama de Barres

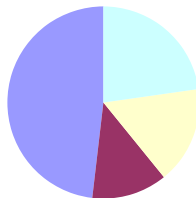
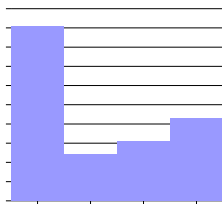
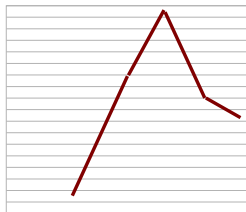


Diagrama de tarta



Histograma



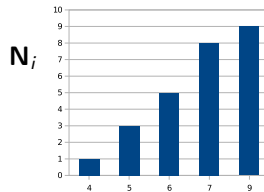
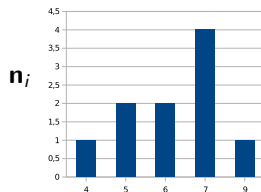
Polígon de Freqüències

En color **vermell** nomès per a variables quantitatives

Representació gràfica de dades estadístiques

- ▶ Diagrama de barres:
 - ▶ Una barra per a cada valor o interval de valors
 - ▶ Alçada de les barres proporcional a la freqüència (absoluta o relativa)
- ▶ Exemple:

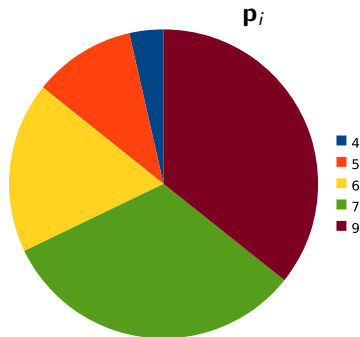
x_i	n_i	N_i	f_i	F_i	p_i	P_i
4	1	1	0.1	0.1	10	10
5	2	3	0.2	0.3	20	30
6	2	5	0.2	0.5	20	50
7	4	9	0.4	0.9	40	90
9	1	10	0.1	1.0	10	100



Representació gràfica de dades estadístiques

- ▶ Diagrama de tarta:
 - ▶ Un sector per a cada valor o interval de valors
 - ▶ Àrea del sector proporcional a la freqüència (absoluta o relativa)
- ▶ Exemple:

x_i	n_i	N_i	f_i	F_i	p_i	P_i
4	1	1	0.1	0.1	10	10
5	2	3	0.2	0.3	20	30
6	2	5	0.2	0.5	20	50
7	4	9	0.4	0.9	40	90
9	1	10	0.1	1.0	10	100



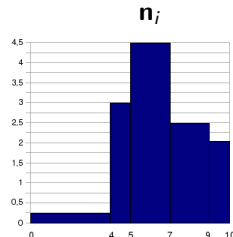
Representació gràfica de dades estadístiques

► Histograma:

- Una barra per a cada interval de valors
- Intervals de valors consecutius i sense espai entre barres (només variables quantitatives contínues)
- Àrea de la barra (no altura) proporcional a la freqüència (absoluta o relativa) (histograma de densitats)

► Exemple:

x_i	m_i	n_i	N_i	f_i	F_i	p_i	P_i
$[0, 4)$	2	1	1	0.05	0.05	5	5
$[4, 5)$	4.5	3	4	0.15	0.20	15	20
$[5, 7)$	6	9	13	0.45	0.65	45	65
$[7, 9)$	8	5	18	0.25	0.90	25	90
$[9, 10)$	9.5	2	20	0.10	1.00	10	100



Representació gràfica de dades estadístiques

- ▶ Polígon de freqüències:
 - ▶ A partir de l'histograma
 - ▶ Línees que uneixen els centres dels intervals
- ▶ Exemple:

x_i	m_i	n_i	N_i	f_i	F_i	p_i	P_i
$[0, 4)$	2	1	1	0.05	0.05	5	5
$[4, 5)$	4.5	3	4	0.15	0.20	15	20
$[5, 7)$	6	9	13	0.45	0.65	45	65
$[7, 9)$	8	5	18	0.25	0.90	25	90
$[9, 10)$	9.5	2	20	0.10	1.00	10	100

