

Representació de dades estadístiques

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

| |
|----------|
| x_1 |
| x_2 |
| x_3 |
| x_4 |
| x_5 |
| x_6 |
| \vdots |
| x_n |

Dades brutes



| 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 |
| x_4 | n_4 | N_4 | f_4 | F_4 | p_4 | P_4 |
| x_5 | n_5 | N_5 | f_5 | F_5 | p_5 | P_5 |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| x_k | n_k | $N_k = n$ | f_k | $F_k = 1$ | p_k | $P_k = 100$ |

Taula de freqüències

Representació de dades estadístiques

- Les taules de freqüència contenen informació de:
 - els valors de la variable: x_i
 - el 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**: $f_i = n_i / n$
 - Les **freqüències relatives acumulades***: $F_i = N_i / n$
 - Els **percentatges absoluts i acumulats***: $p_i = f_i \times 100\%$,
 $P_i = F_i \times 100\%$

*només per a variables quantitatives i ordinals

Representació de dades estadístiques

Dades brutes

n valors
totals

| |
|----------|
| x_1 |
| x_2 |
| x_3 |
| x_4 |
| x_5 |
| x_6 |
| \vdots |
| x_n |

Taula de freqüències

| k valors diferents | freq. absolutes acumulades (*) | | freq. relatives acumulades (*) | | Percentatges absoluts | | Percentatges acumulats(*) |
|-----------------------|-----------------------------------|-----------|-----------------------------------|-----------|--------------------------|-------------|------------------------------|
| | freq. absolutes | | freq. relatives | | | | |
| 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 | |
| x_4 | n_4 | N_4 | f_4 | F_4 | p_4 | P_4 | |
| x_5 | n_5 | N_5 | f_5 | F_5 | p_5 | P_5 | |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | |
| x_k | n_k | $N_k = n$ | f_k | $F_k = 1$ | p_k | $P_k = 100$ | |
| | n | | 1 | | | | |

(*) Només per a variables
quantitatives i ordinals

Representació de dades estadístiques

Exemple: nota d'estadística de 9 persones

Dades brutes

| |
|---|
| 7 |
| 5 |
| 9 |
| 7 |
| 5 |
| 6 |
| 7 |
| 6 |
| 4 |



Taula de freqüències

| x_i | n_i | N_i | f_i | F_i | p_i | P_i |
|-------|-------|-------|-------|-------|-------|-------|
| 4 | 1 | 1 | 0,11 | 0,11 | 11 | 11 |
| 5 | 2 | 3 | 0,22 | 0,33 | 22 | 33 |
| 6 | 2 | 5 | 0,22 | 0,55 | 22 | 55 |
| 7 | 3 | 8 | 0,33 | 0,89 | 33 | 89 |
| 9 | 1 | 9 | 0,11 | 1 | 11 | 100 |

$n=9$

1

Representació de dades estadístiques

Representació de dades amb intervals (variables quantitatives i ordinals)

- Per a simplificar la representació de les dades en ocasions aquestes 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

Dades brutes

| | |
|-----|-----|
| 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 |

Taula de freqüències

| 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 | 10 | 100 |

$n=20$

l

Representació gràfica de dades estadístiques

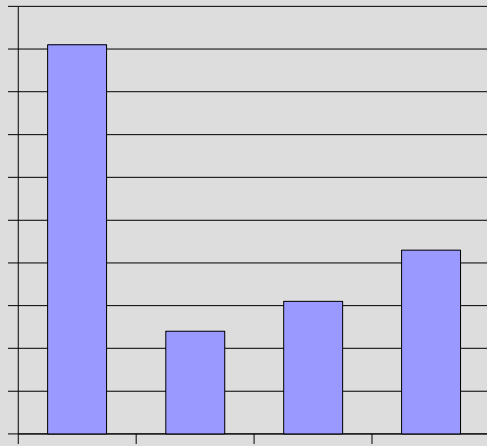


Diagrama de barres

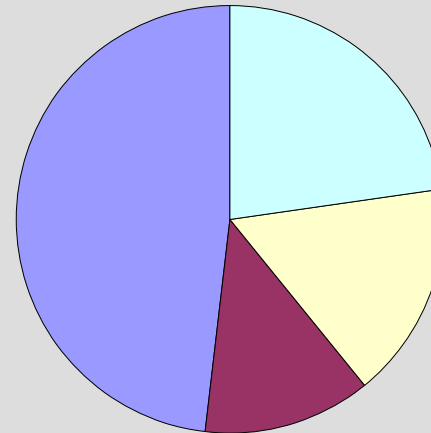
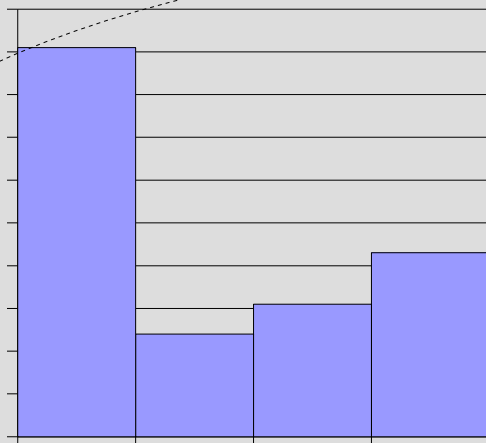
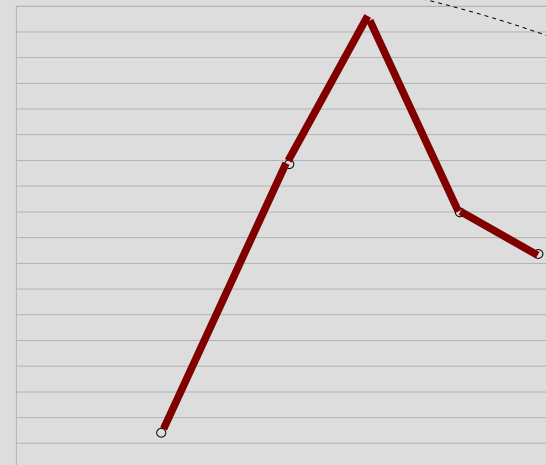


Diagrama de sectors (pastís)



Histograma



Polígon de freqüències

Variables
quantitatives

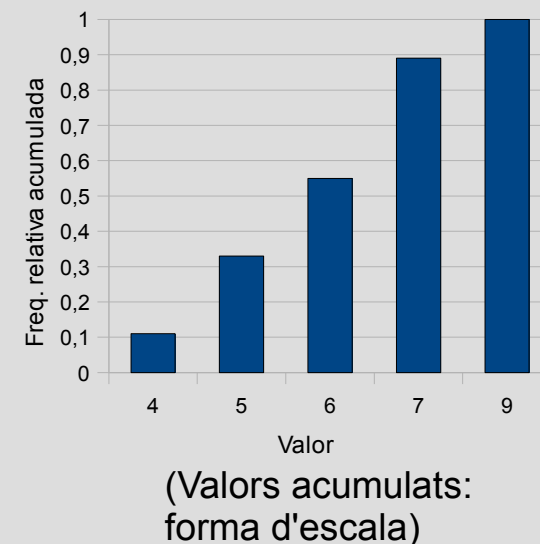
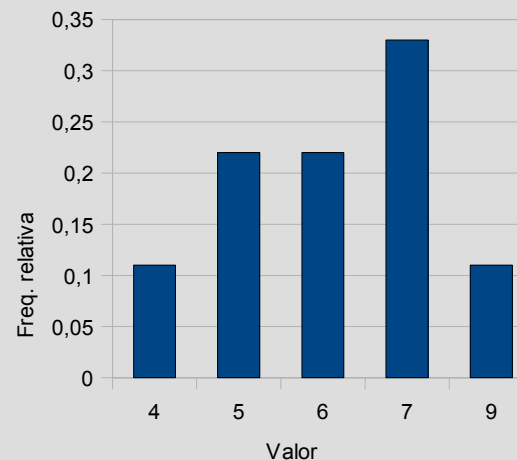
Representació gràfica de dades estadístiques

– 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 | 0,11 | 0,11 | 11 | 11 |
| 5 | 2 | 3 | 0,22 | 0,33 | 22 | 33 |
| 6 | 2 | 5 | 0,22 | 0,55 | 22 | 55 |
| 7 | 3 | 8 | 0,33 | 0,89 | 33 | 89 |
| 9 | 1 | 9 | 0,11 | 1 | 11 | 100 |



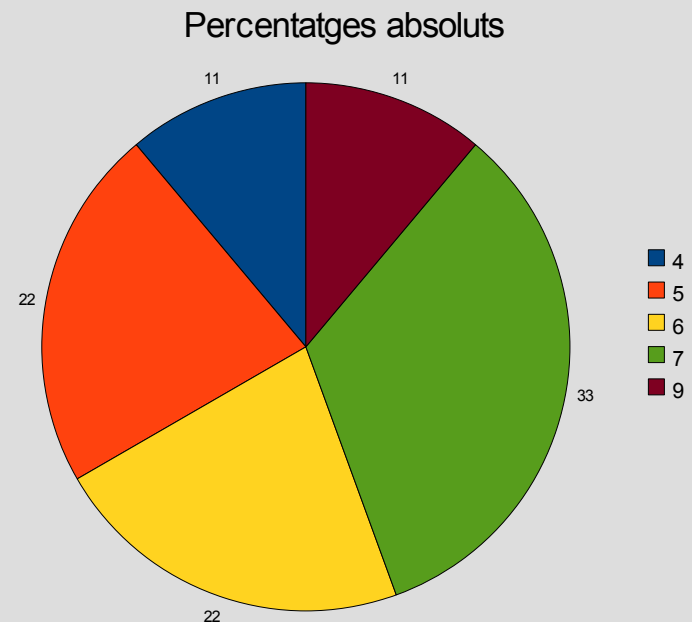
Representació gràfica de dades estadístiques

– Diagrama de sectors:

- 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 | 0,11 | 0,11 | 11 | 11 |
| 5 | 2 | 3 | 0,22 | 0,33 | 22 | 33 |
| 6 | 2 | 5 | 0,22 | 0,55 | 22 | 55 |
| 7 | 3 | 8 | 0,33 | 0,89 | 33 | 89 |
| 9 | 1 | 9 | 0,11 | 1 | 11 | 100 |



Representació gràfica de dades estadístiques

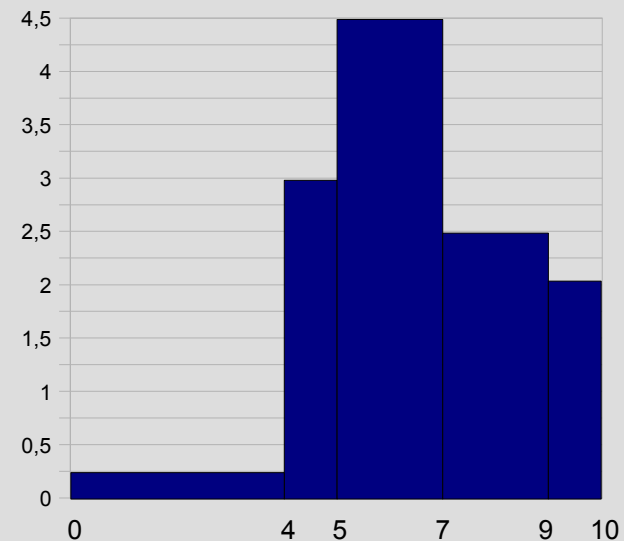
– Histograma:

- Una barra per a cada interval de valors
- Intervals de valors consecutius i sense espai entre barres (només variables quantitatives continues)
- 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 | 10 | 100 |

Histograma de freqüències absolutes



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 | 10 | 100 |

Polígon de freqüències absolutes

