

# 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.

# Mesures de tendència central

- ▶ El conjunt de valors de freqüència associats a una variable estadística rep el nom de **distribució de freqüències** de la variable.
  - ▶ Les característiques principals d'una distribució de freqüències es poden resumir amb uns pocs valors numèrics anomenats **estadístics**.
  - ▶ Els **estadístics de tendència central** resumeixen el comportament global de la distribució.
    - ▶ **Moda**: valor amb màxima freqüència absoluta.
    - ▶ **Mitjana**: mitjana aritmètica dels valors (únicament variables quantitatives)
    - ▶ **Mediana**: valor que ocupa la posició central de les dades ordenades (Únicament variables quantitatives i ordinals)
      - ▶ **Percentil**  $P_p$ : valor  $x$  tal que el  $p\%$  dels valors de la distribució són inferiors o iguals a  $x$ .
      - ▶ **Quartils**:  $Q_1 = P_{25}$ ,  $Q_2 = P_{50}$  i  $Q_3 = P_{75}$ .
- La mitjana és igual al percentil 50 i al 2<sup>on</sup> quartil.

# Mesures de tendència central

► Exemple: **Moda**

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

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

# Mesures de tendència central

► Exemple: **Moda**

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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
<b>7</b>	4	9	0.4	0.9	40	90
9	1	10	0.1	1.0	10	100

$n = 10$        $f = 1$

# Mesures de tendència central

## ► Exemple: Mediana

$$n = \begin{cases} d_{(\frac{n+1}{2})} & n \text{ impar} \\ \frac{d_{(n/2)} + d_{(n/2+1)}}{2} & n \text{ par} \end{cases}$$

$d_i$
7
5
9
7
5
6
7
7
6
4
7

⇒

$d_i$
4
5
5
6
6
7
7
7
7
9

⇒

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

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

# Mesures de tendència central

## ► Exemple: Mediana

$$n = \begin{cases} d_{(\frac{n+1}{2})} & n \text{ impar} \\ \frac{d_{(n/2)} + d_{(n/2+1)}}{2} & n \text{ par} \end{cases}$$

$d_i$		$d_i$	
7		4	
5		5	
9		5	
7		6	
5	$\Rightarrow$	6	$\Rightarrow$
6		7	
7		7	
6		7	
4		7	
7		9	

mediana 6.5

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

$n = 10$      $f = 1$

# Mesures de tendència central

## ► Exemple: Mediana

$$n = \begin{cases} d_{(\frac{n+1}{2})} & n \text{ impar} \\ \frac{d_{(n/2)} + d_{(n/2+1)}}{2} & n \text{ par} \end{cases}$$

$\Rightarrow$

**mediana 6.5**

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

$n = 10 \quad f = 1$



# Mesures de tendència central

- Exemple: **Percentils**  $k = \text{Int} \left( \frac{n \cdot p}{100} \right)$ ,  $d = \frac{n \cdot p}{100} - k$ ,

$$P_p = \begin{cases} d_{k+1} & \text{si } d \neq 0 \\ \frac{d_k + d_{k+1}}{2} & \text{si } d = 0 \end{cases}$$

$d_i$
7
5
9
7
5
6
7
6
4
7

$\Rightarrow$

$d_i$
4
5
5
6
6
7
7
7
7
9

$\Rightarrow$

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

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

# Mesures de tendència central

► Exemple: **Percentils**  $k = \text{Int} \left( \frac{n \cdot p}{100} \right)$ ,  $d = \frac{n \cdot p}{100} - k$ ,

$p = 20 \rightarrow k = 2, d = 0$

$$P_p = \begin{cases} d_{k+1} & \text{si } d \neq 0 \\ \frac{d_k + d_{k+1}}{2} & \text{si } d = 0 \end{cases}$$

$d_i$
7
5
9
7
5
6
7
6
4
7

$P_{20}$

$\Rightarrow$

$d_i$
4
5
5
6
6
7
7
7
7
9

$\Rightarrow$

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

$n = 10$      $f = 1$

# Mesures de tendència central

- Exemple: **Percentils**  $k = \text{Int} \left( \frac{n \cdot p}{100} \right)$ ,  $d = \frac{n \cdot p}{100} - k$ ,

$$p = 20 \rightarrow k = 2, d = 0$$

$$P_p = \begin{cases} d_{k+1} & \text{si } d \neq 0 \\ \frac{d_k + d_{k+1}}{2} & \text{si } d = 0 \end{cases}$$

$d_i$
7
5
9
7
5
6
7
6
4
7

$P_{20}$

$\Rightarrow$

$d_i$
4
5
5
6
6
7
7
7
7
9

$\Rightarrow$

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

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

# Mesures de tendència central

- Exemple: **Percentils**  $k = \text{Int} \left( \frac{n \cdot p}{100} \right)$ ,  $d = \frac{n \cdot p}{100} - k$ ,

$$P_p = \begin{cases} d_{k+1} & \text{si } d \neq 0 \\ \frac{d_k + d_{k+1}}{2} & \text{si } d = 0 \end{cases}$$

$$p = 20 \rightarrow k = 2, d = 0$$

$$p = 65 \rightarrow k = 6, d = 5$$

$d_i$
7
5
9
7
5
6
7
6
4
7

$P_{20}$

$\Rightarrow$

$P_{65}$

$d_i$
4
5
5
6
6
7
7
7
7
9

$\Rightarrow$

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

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

# Mesures de tendència central

► Exemple: **Percentils**  $k = \text{Int} \left( \frac{n \cdot p}{100} \right)$ ,  $d = \frac{n \cdot p}{100} - k$ ,

$$P_p = \begin{cases} d_{k+1} & \text{si } d \neq 0 \\ \frac{d_k + d_{k+1}}{2} & \text{si } d = 0 \end{cases}$$

$p = 20 \rightarrow k = 2, d = 0$

$p = 65 \rightarrow k = 6, d = 5$

$d_i$
7
5
9
7
5
6
7
6
4
7

$P_{20}$

$\Rightarrow$

$P_{65}$

$d_i$
4
5
5
6
6
7
7
7
7
9

$\Rightarrow$

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

$n = 10$      $f = 1$

# Mesures de tendència central

## ► Exemple: Quartils

$$Q_1 = P_{25}, Q_2 = P_{50} \text{ i } Q_3 = P_{75}.$$

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

$n = 10$        $f = 1$

# Mesures de tendència central

► Exemple: **Mitjana**

$$\bar{x} = \frac{d_1 + d_2 + \dots + d_n}{n}$$

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n}$$

$d_i$
4
5
5
6
6
7
7
7
7
9

⇒

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

# Mesures de tendència central

► Exemple: **Mitjana**

$$\bar{x} = \frac{d_1 + d_2 + \dots + d_n}{n}$$

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n}$$

$d_i$
4
5
5
6
6
7
7
7
7
9

⇒

$x_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
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

$$\bar{x} = 6.3$$

$$63 \rightarrow \bar{x} = \frac{63}{10} = 6.3$$



# Mesures de tendència central, tablas con intervalos

- **Moda:** Sigui  $[x_M, x_{M+1}]$  l'interval que conté la freqüència més gran

$$\text{moda} = \frac{x_M + x_{M+1}}{2}$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
$[0, 4)$	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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, tablas con intervalos

- **Moda:** Sigui  $[x_M, x_{M+1}]$  l'interval que conté la freqüència més gran

$$\text{moda} = \frac{x_M + x_{M+1}}{2}$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
$[0, 4)$	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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, tablas con intervalos

- **Moda:** Sigui  $[x_M, x_{M+1}]$  l'interval que conté la freqüència més gran

$$\text{moda} = \frac{x_M + x_{M+1}}{2} = \frac{5 + 7}{2} = 6$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
$[0, 4)$	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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, taules amb intervals

- **Percentil:** Sigui  $[x_k, x_{k+1}]$  l'interval que conté el percentil

$$P_p = x_k + \frac{\frac{p \cdot n}{100} - N_{k-1}}{n_k} (x_{k+1} - x_k);$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
$[0, 4)$	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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, taules amb intervals

- **Percentil:** Sigui  $[x_k, x_{k+1}]$  l'interval que conté el percentil

$$P_p = x_k + \frac{\frac{p \cdot n}{100} - N_{k-1}}{n_k} (x_{k+1} - x_k);$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
[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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, taules amb intervals

- **Percentil:** Sigui  $[x_k, x_{k+1}]$  l'interval que conté el percentil

$$P_p = x_k + \frac{\frac{p \cdot n}{100} - N_{k-1}}{n_k} (x_{k+1} - x_k);$$

$$P_{75} = 7 + \frac{0.75 \cdot 10 - 13}{5} (9 - 7) = 7.8$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
[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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, taules amb intervals

## ► Mitjana:

$$\bar{x} = \frac{m_1 n_1 + m_2 n_2 + \dots + m_k n_k}{n}$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
[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

$$n = 10$$

$$f = 1$$

# Mesures de tendència central, taules amb intervals

## ► Mitjana:

$$\bar{x} = \frac{m_1 n_1 + m_2 n_2 + \dots + m_k n_k}{n}$$

$$\bar{x} = \frac{2*1+4.5*3+6*9+8*5+9.5*2}{10} = 6.425$$

$x_i$	$m_i$	$n_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
$[0, 4)$	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

$$n = 10$$

$$f = 1$$