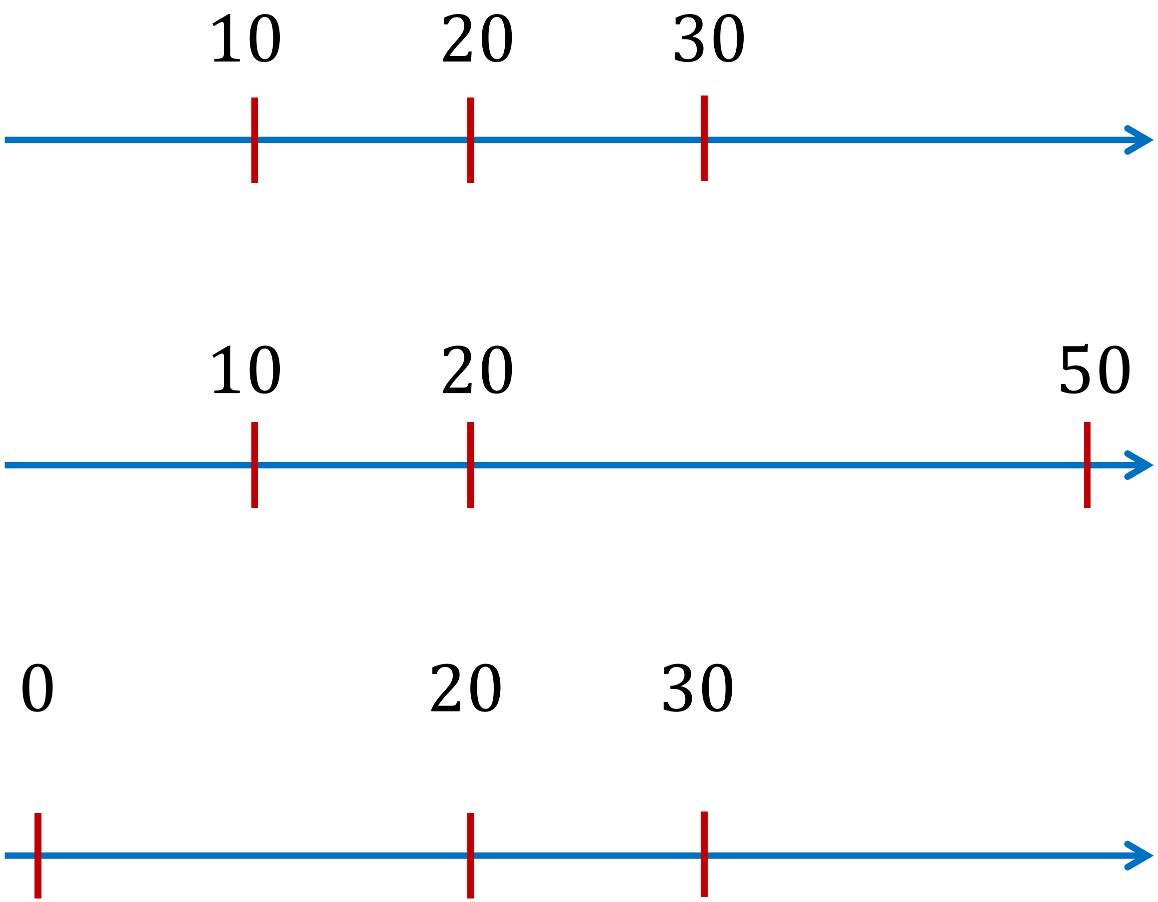
Media e mediana a confronto

$n = 3$



Indici di posizione

dati : $x_1, x_2, \cdots, x_n$

dati **riordinati**:

$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$

$x_{(10)} = 135$

$x_{(11)} = 136$

135.5

147.5

Unità	Peso	Peso o.
1	118	106
2	151	114
3	143	118
4	172	118
5	147	122
6	146	127
7	138	127
8	175	132
9	134	134
10	172	135
11	118	136
12	151	138
13	155	139
14	155	140
15	146	143
16	135	143
17	127	144
18	178	146
19	136	146
20	180	147
21	151	148
22	186	151
23	122	151
24	132	151
25	114	153
26	171	155
27	140	155
28	187	159
29	106	171
30	159	172
31	127	172
32	191	175
33	192	178
34	181	179
35	143	180
36	153	181
37	144	186
38	139	187
39	148	191
40	179	192

La **posizione** del **I° quartile** è $\frac{n+1}{4}$ (10.25)

Il **valore** della **I° quartile** è $\frac{x_{(n/4)} + x_{(n/4+1)}}{2}$

(**IL I° quartile**)

Indici di posizione

dati : $x_1, x_2, \cdots, x_n$

dati **riordinati**:

$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$

La **posizione** del **III° quartile** è $\dfrac{3 \times (n+1)}{4}$ (30.75)

Il **valore** della **III° quartile** è $\dfrac{x_{(3(n+1)/4)} + x_{(3(n+1)/4)}}{2}$

(**IL III° quartile**)

Unità	Peso	Peso o.
1	118	106
2	151	114
3	143	118
4	172	118
5	147	122
6	146	127
7	138	127
8	175	132
9	134	134
10	172	135
11	118	136
12	151	138
13	155	139
14	155	140
15	146	143
16	135	143
17	127	144
18	178	146
19	136	146
20	180	147
21	151	148
22	186	151
23	122	151
24	132	151
25	114	153
26	171	155
27	140	155
28	187	159
29	106	171
30	159	172
31	127	172
32	191	175
33	192	178
34	181	179
35	143	180
36	153	181
37	144	186
38	139	187
39	148	191
40	179	192

135.5

147.5

172

Indici di posizione

dati : $x_1, x_2, \cdots, x_n$

$n = 11$

dati **riordinati**:

$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$

Unità	Peso	**Peso o.**
1	118	118
2	151	118
3	143	**134**
4	172	138
5	147	143
6	146	146
7	138	147
8	175	151
9	134	172
10	172	172
11	118	175

La **posizione** del **I° quartile** è $\frac{n+1}{4}$ (3)

Il **valore** del **I° quartile** è $x_{\left(\frac{n+1}{4}\right)}$ 134

(**IL I° quartile**)

Indici di posizione

dati : $x_1, x_2, \cdots, x_n$

$n = 11$

dati **riordinati**:

$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$

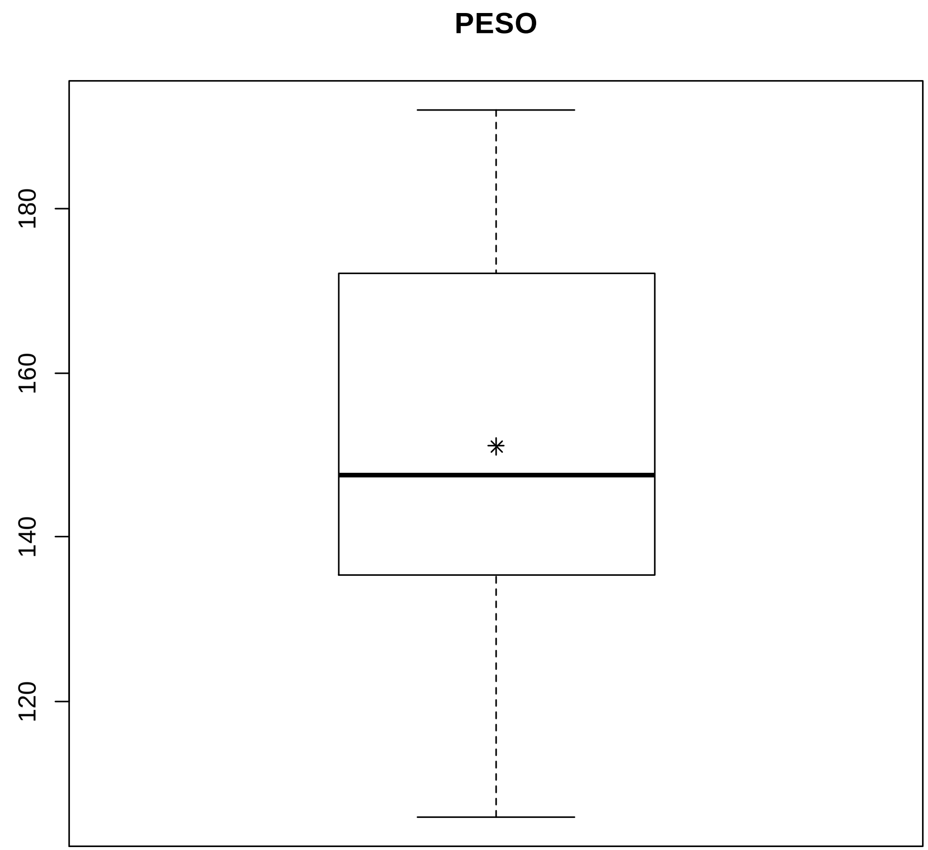
Unità	Peso	Peso o.
1	118	118
2	151	118
3	143	134
4	172	138
5	147	143
6	146	146
7	138	147
8	175	151
9	134	**172**
10	172	172
11	118	175

La **posizione** del **III° quartile** è $\frac{\mathbf{3(n+1)}}{\mathbf{4}}$ (9)

Il **valore** del **III° quartile** è $x_{\left(\frac{3(n+1)}{4}\right)}$ 172

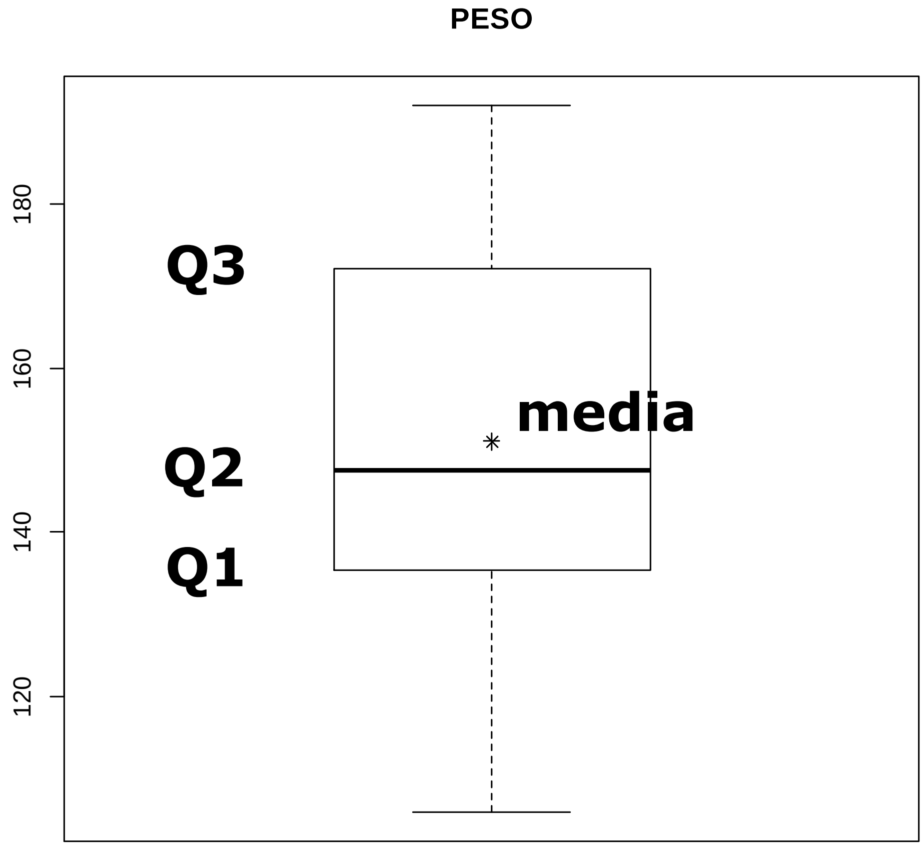
(**IL III° quartile**)

Boxplot



Boxplot

Differenza (*range*) interquartile: Q3-Q1: IQR

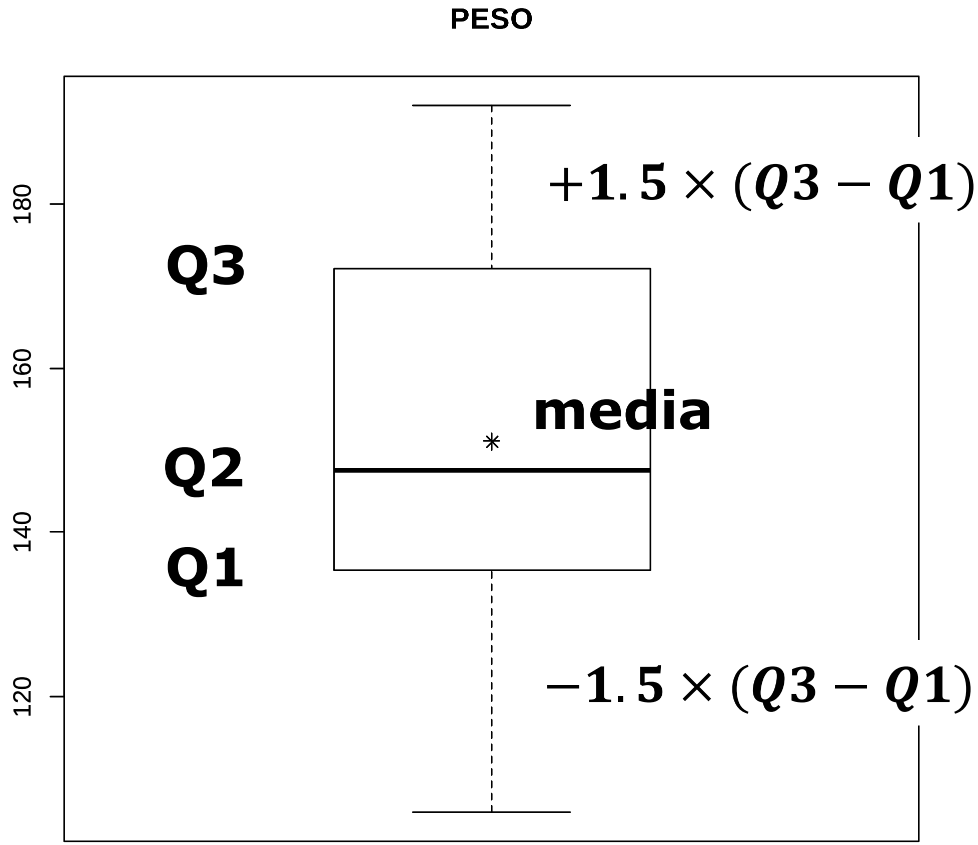


Boxplot

Differenza (*range*) interquartile: Q3-Q1: IQR

$$\mathbf{1.5} \times (\boldsymbol{Q3} - \boldsymbol{Q1})$$

è la
lunghezza **massima**
dei baffi



Boxplot

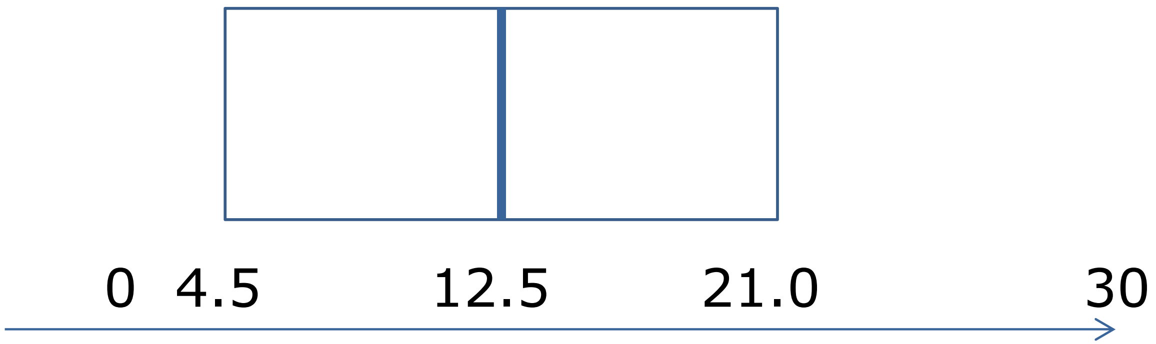
500 dati sul tempo di attesa (in minuti) presso il CUP di un'azienda ospedaliera, dati variano da 0 a 30 minuti (*range*)

Quartili:

$$Q1 = 4.5$$

$$Q2 = 12.5$$

$$Q3 = 21.0$$



Boxplot

500 dati sul tempo di attesa (in minuti) presso il CUP di un'azienda ospedaliera, dati variano da 0 a 30 minuti (*range*)

Quartili:

$Q1 = 4.5$

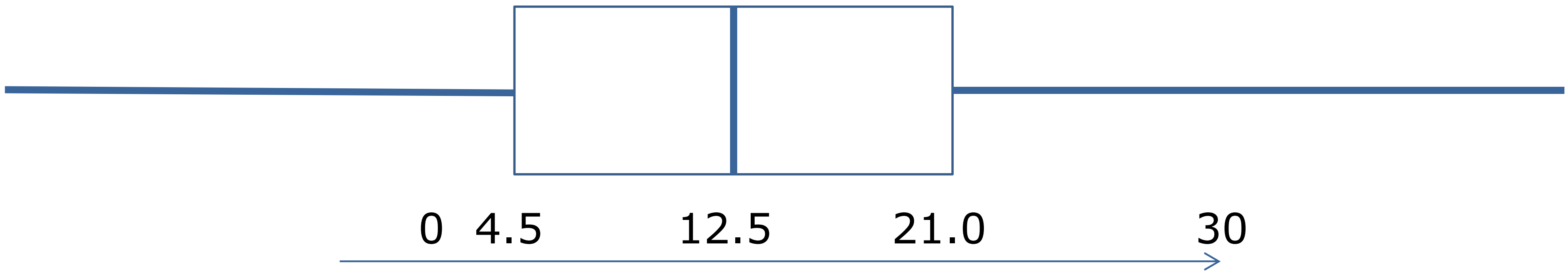
$Q2 = 12.5$

$Q3 = 21.0$

$BAFFI: \quad 1.5 \times (21.0 - 4.5) = 24.75$

$Q1 - 24.75 = -20.25 < 0$

$Q3 + 24.75 = 45.75 > 30$



Boxplot

500 dati sul tempo di attesa (in minuti) presso il CUP di un'azienda ospedaliera, dati variano da 0 a 30 minuti (*range*)

Quartili:

$Q1 = 4.5$

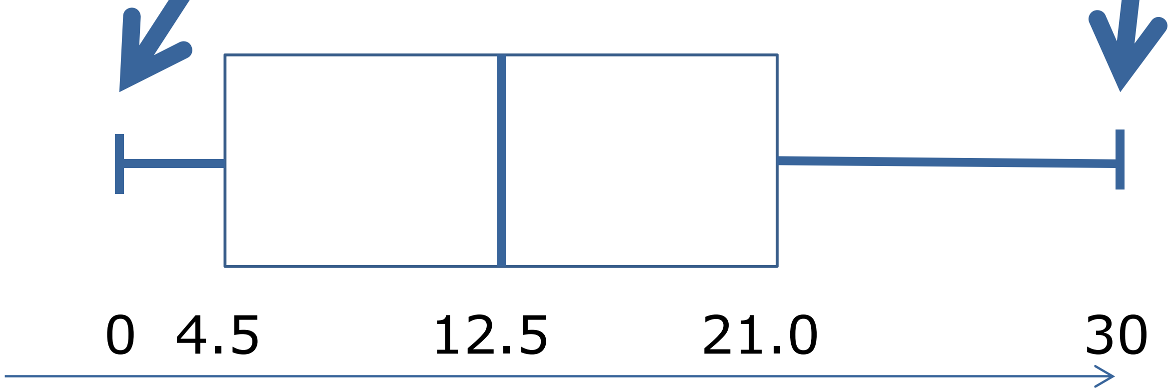
$Q2 = 12.5$

$Q3 = 21.0$

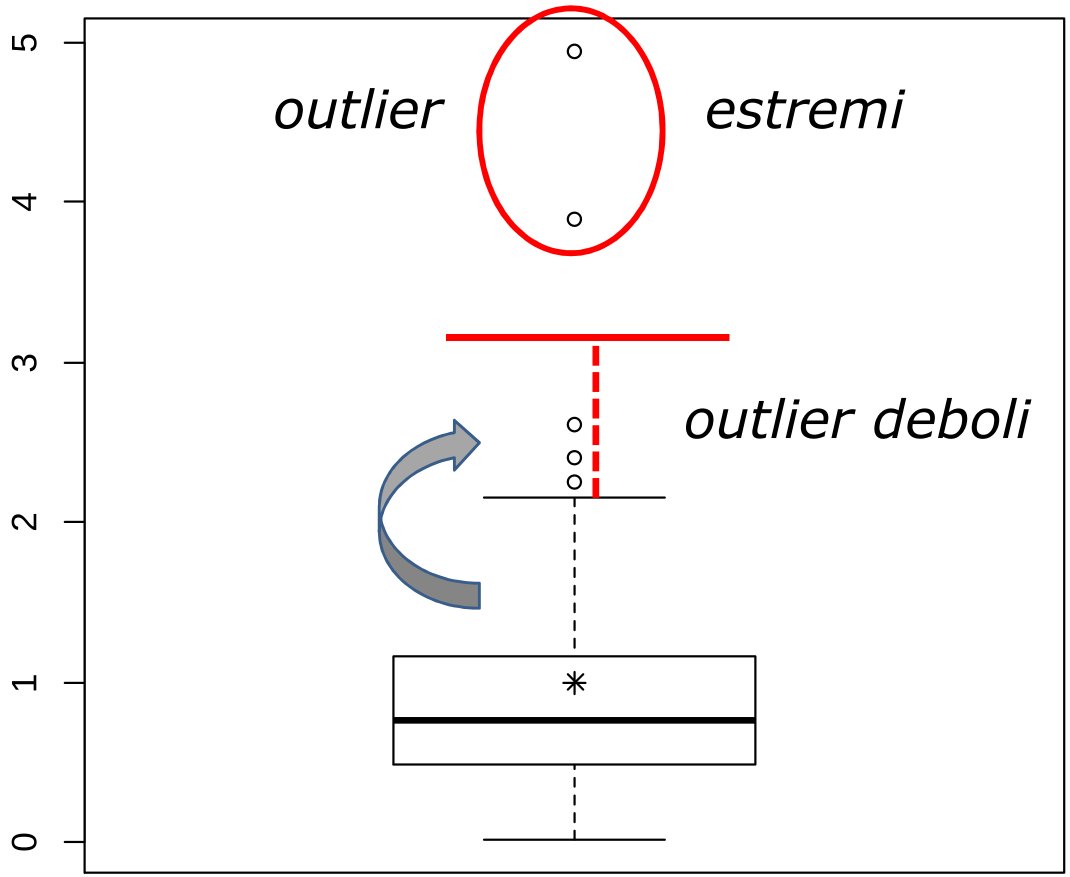
$BAFFI: \quad 1.5 \times (21.0 - 4.5) = 24.75$

$Q1 - 24.75 = -20.25 < 0$

$Q3 + 24.75 = 45.75 > 30$



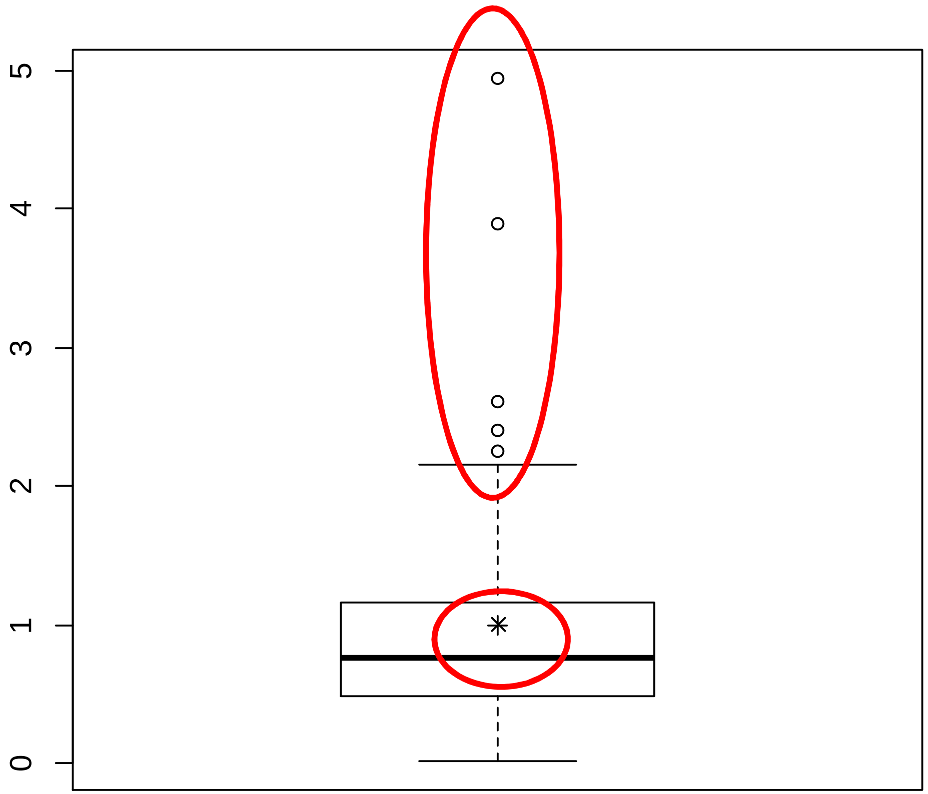
Boxplot



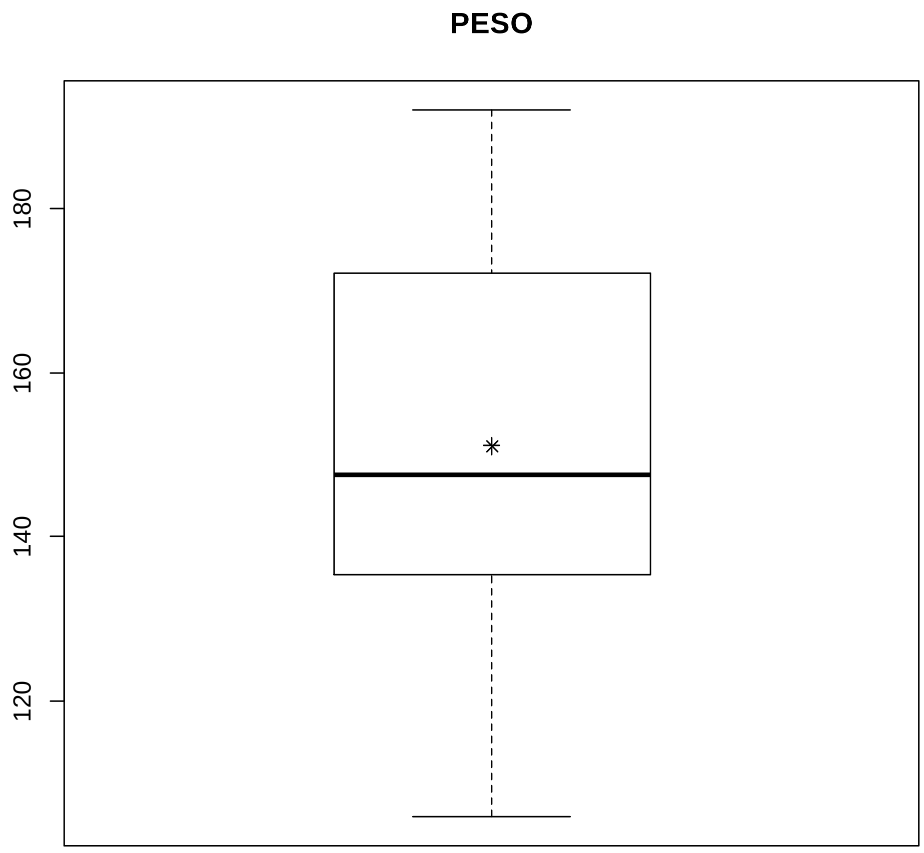
Dati fuori dal baffo (superiore): *outlier*

Baffo più corto della lunghezza massima

Boxplot

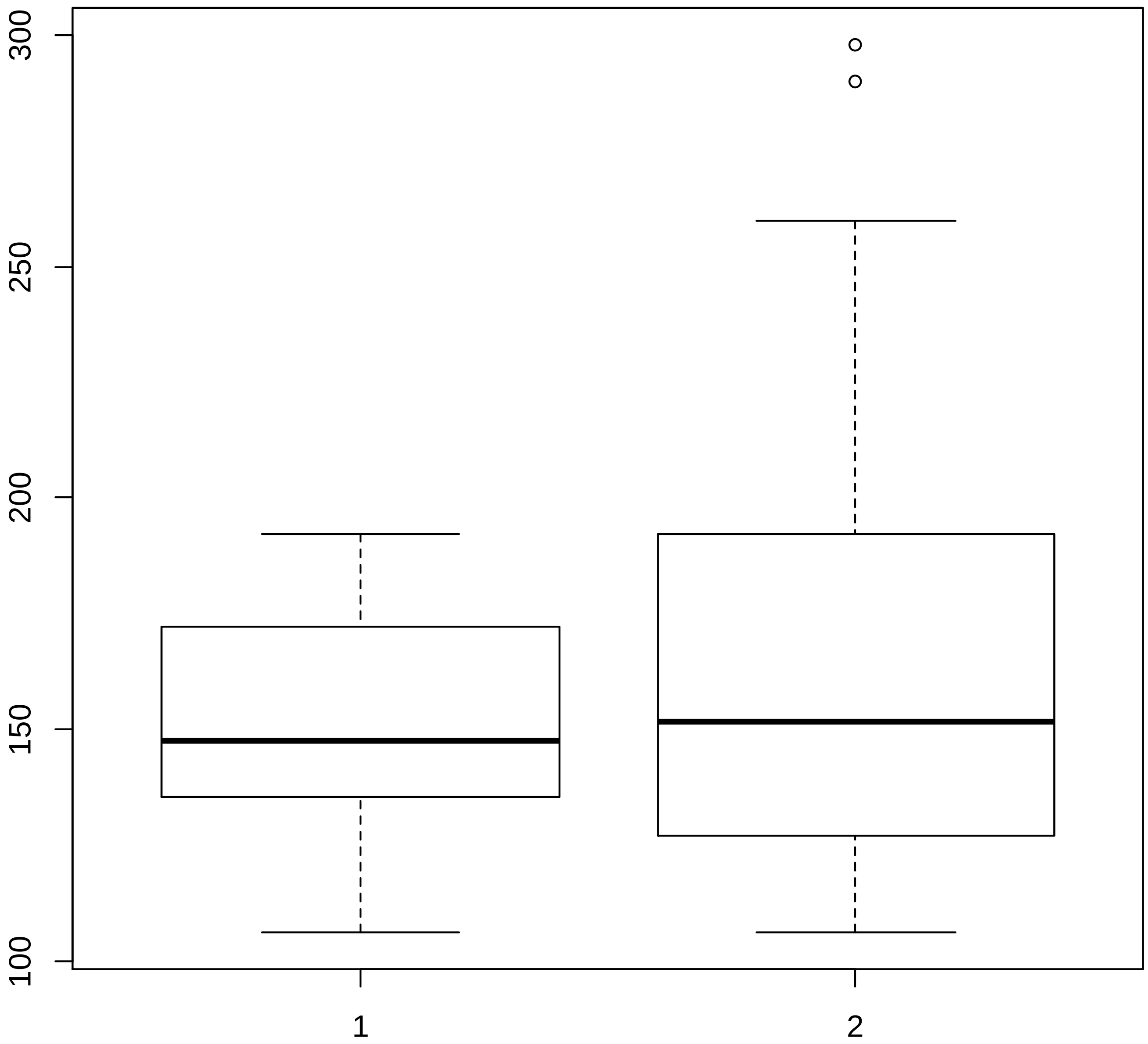


poca simmetria

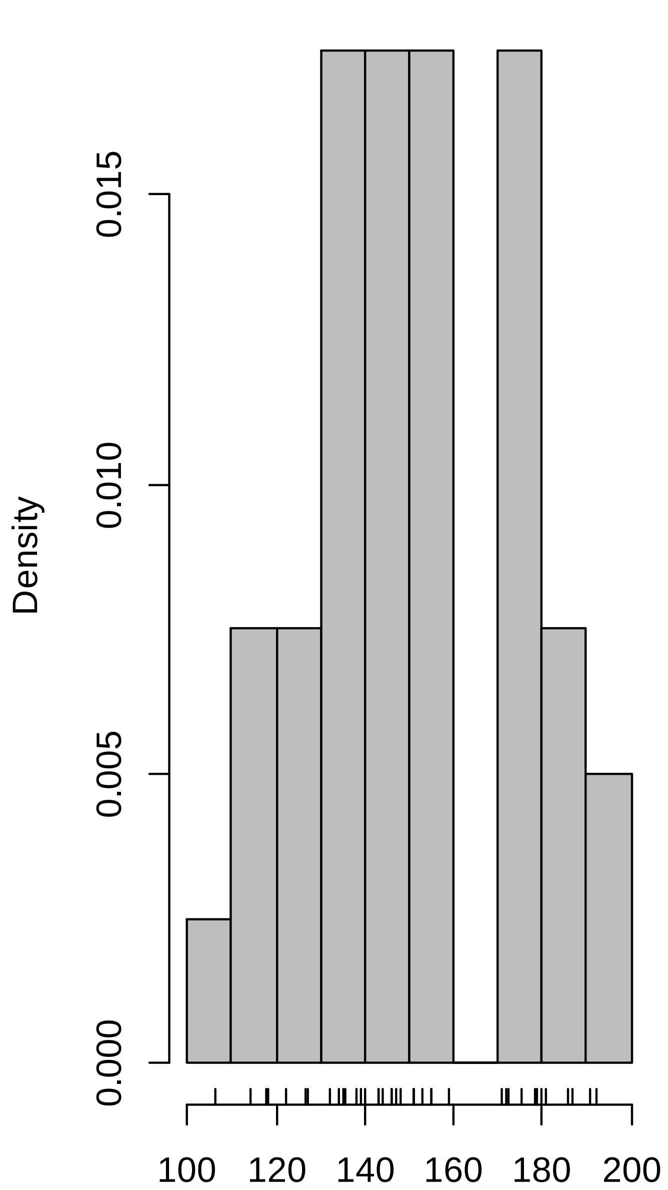


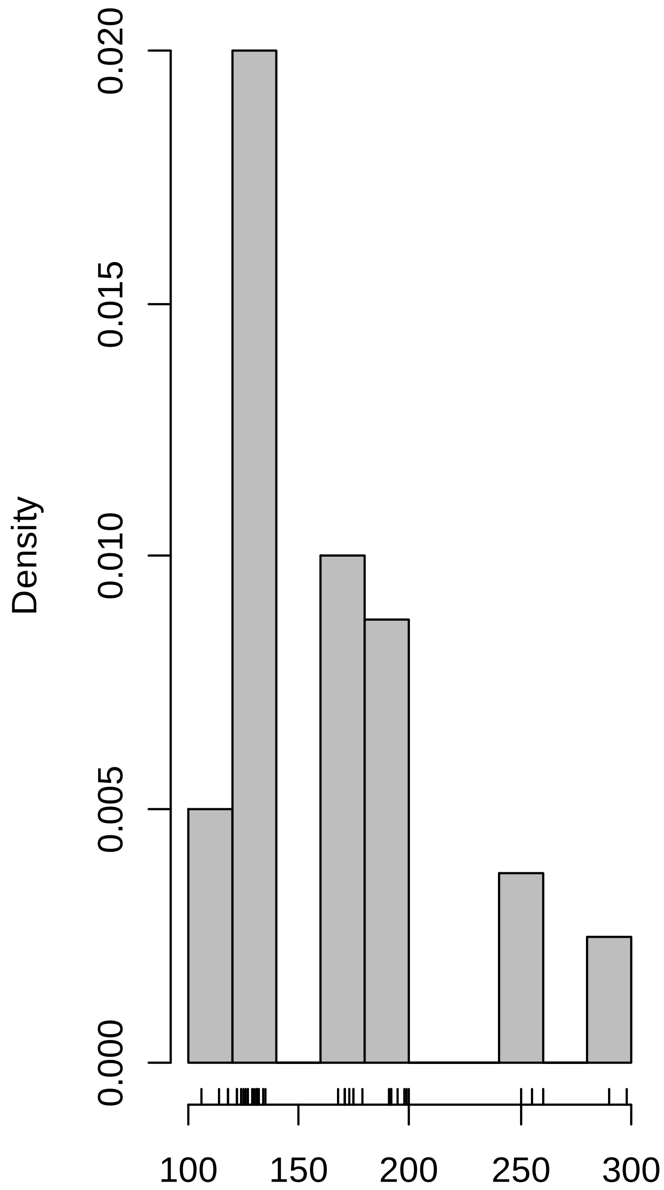
simmetria

Boxplot ...

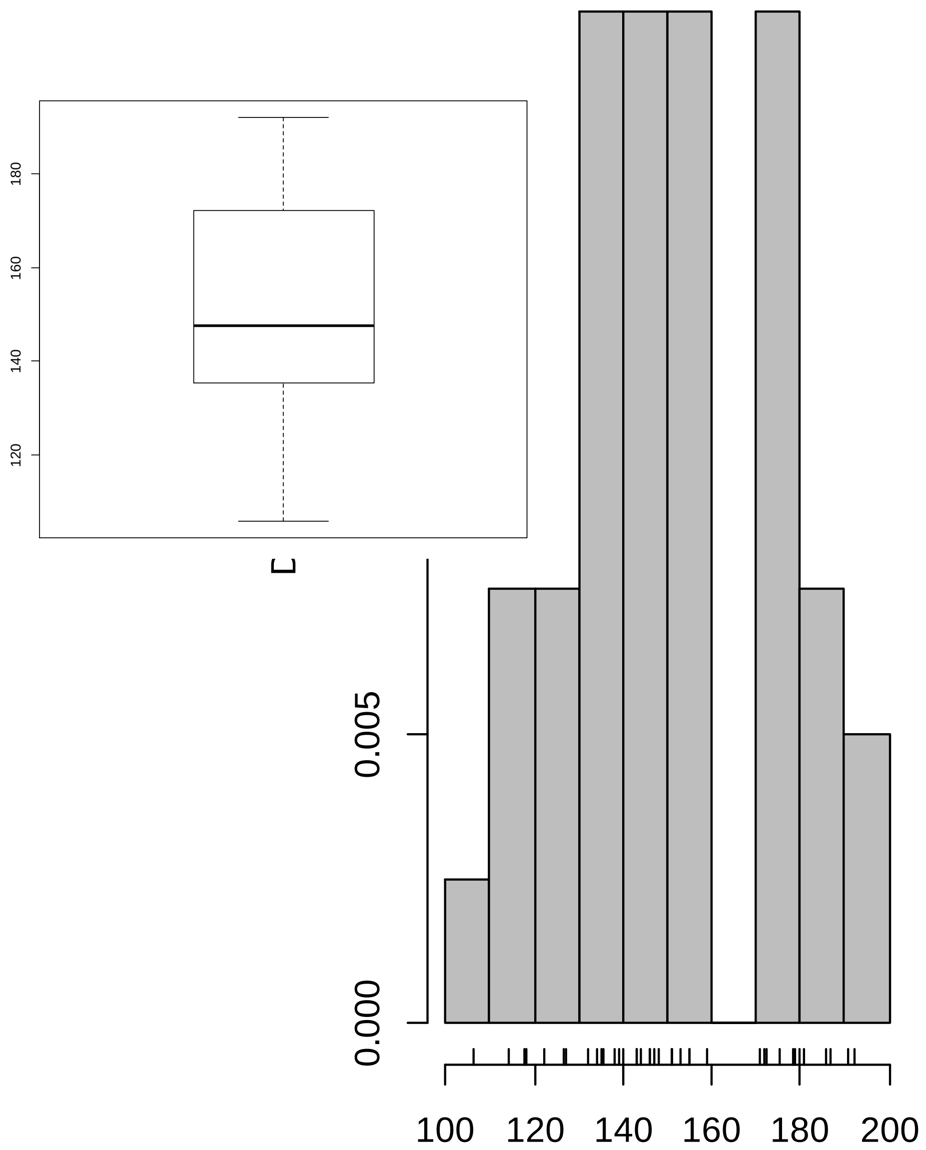


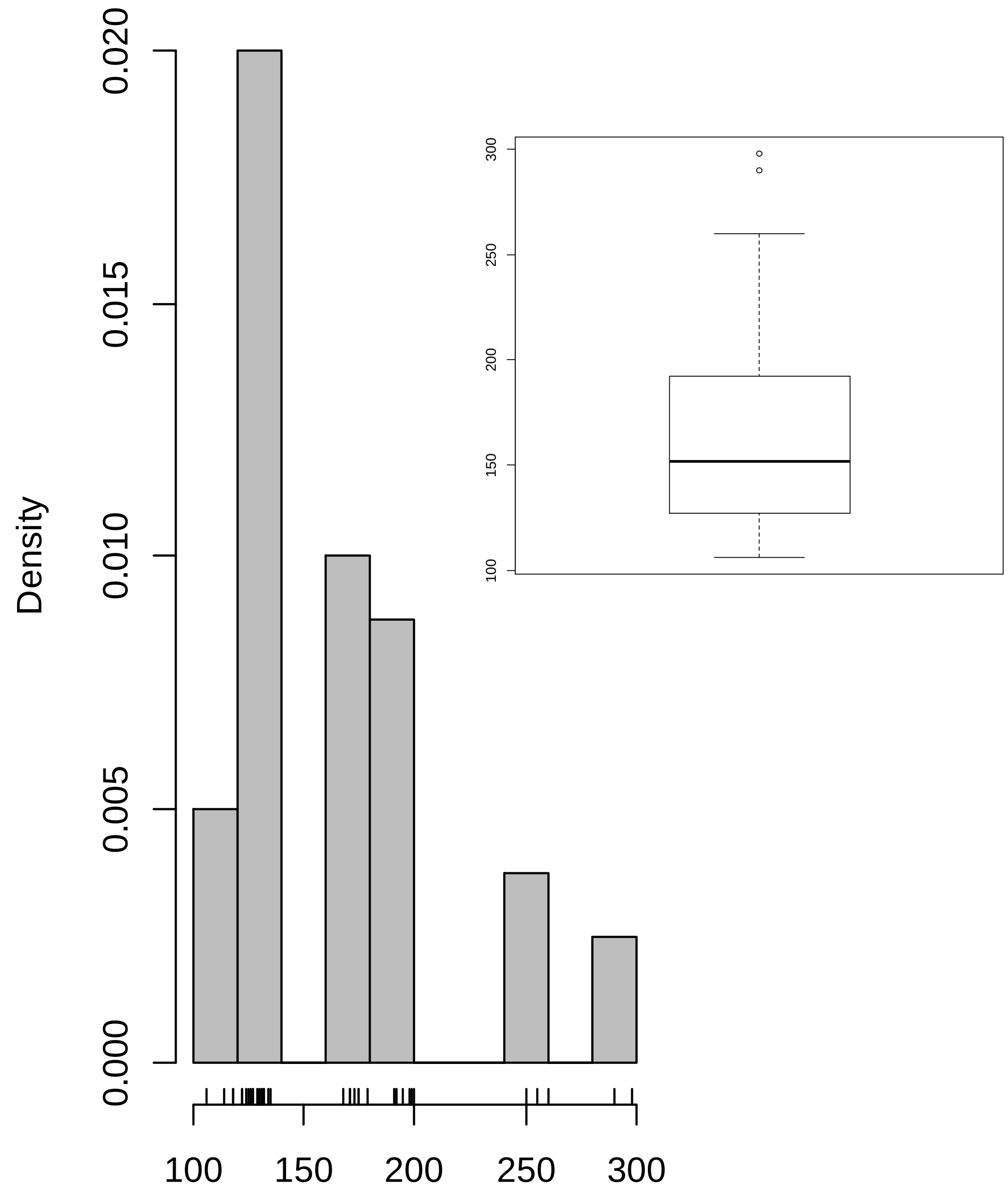
… e istogrammi



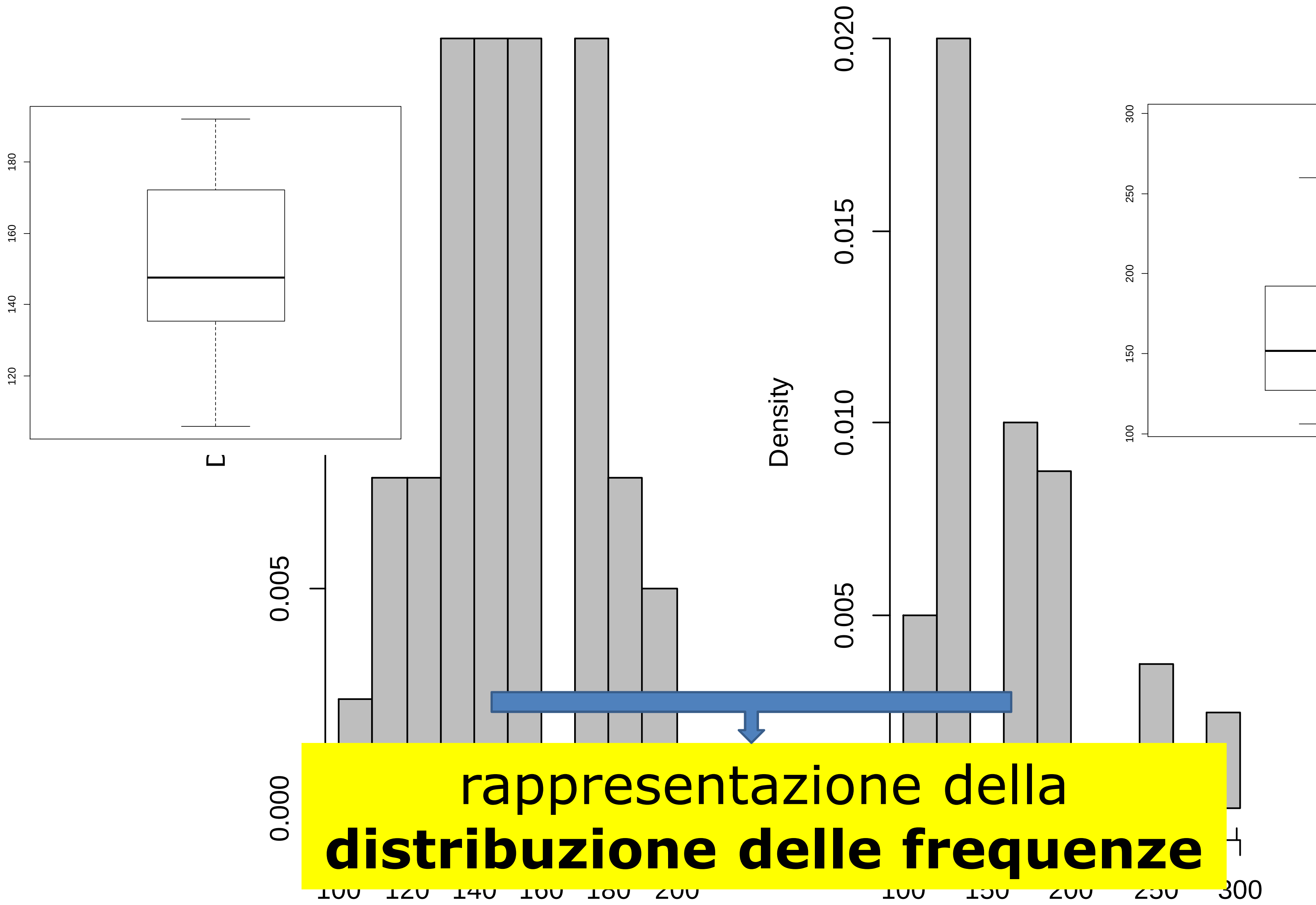


… e istogrammi



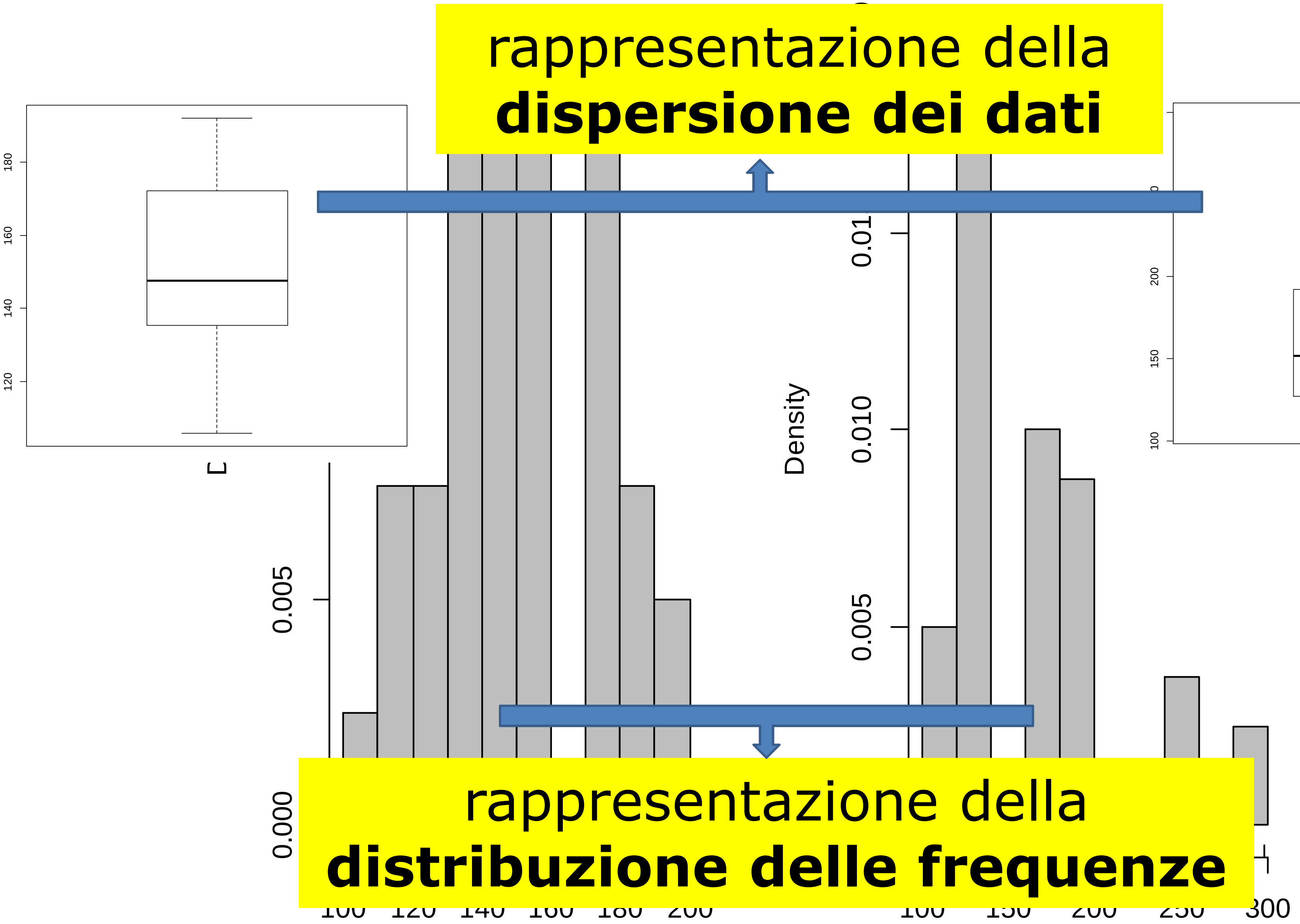


… e istogrammi



rappresentazione della
distribuzione delle frequenze

… e istogrammi



rappresentazione della
dispersione dei dati

rappresentazione della
distribuzione delle frequenze

Percentili

dati : $x_1, x_2, \cdots, x_n$

dati **riordinati**:

$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$

$\mathbf{0 < p < 1}$

Q1: $p = 0.25$
Q2: $p = 0.50$
Q3: $p = 0.75$

Se $(n+1)p$ è intero: $x_{((n+1)p)}$

Se $(n+1)p$ non è intero: $\dfrac{x_{((n+1)p^-)} + x_{((n+1)p^+)}}{2}$

Unità	Peso	**Peso o.**
1	118	106
2	151	114
3	143	118
4	172	118
5	147	122
6	146	127
7	138	127
8	175	132
9	134	134
10	172	135
11	118	136
12	151	138
13	155	139
14	155	140
15	146	143
16	135	143
17	127	144
18	178	146
19	136	146
20	180	147
21	151	148
22	186	151
23	122	151
24	132	151
25	114	153
26	171	155
27	140	155
28	187	159
29	106	171
30	159	172
31	127	172
32	191	175
33	192	178
34	181	179
35	143	180
36	153	181
37	144	186
38	139	187
39	148	191
40	179	192

Media pesata

UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE POLITICHE

SCIENZE POLITICHE
UNIVERSITÀ DEGLI STUDI MILANO

Le Commissioni operano in parallelo utilizzando le aule assegnate dal responsabile dei Servizi Generali di Facoltà, che si impegna a risolvere eventuali sovrapposizioni con la normale attività didattica. Ciascuna Commissione è tenuta a completare il verbale dell'esame finale/esame di laurea utilizzando i moduli precompilati predisposti dall'Ufficio SIFA.

I. – Attribuzione del punteggio e richiesta della lode

Il voto di laurea tiene conto della media dei voti che gli studenti hanno riportato negli esami di profitto, ponderata in base ai crediti degli insegnamenti. La formula per calcolare il voto di media è la seguente:

$$\boxed{\frac{\sum_{i=1}^{n}(\text{voto esame}_i \times \text{numero crediti esame}_i)}{\text{totale numero crediti conseguiti}}} \times \frac{110}{30}$$

Media pesata

Media aritmetica:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \underbrace{x_i}_{\text{dato}} \times \underbrace{\color{red}{\frac{1}{n}}}_{\color{red}{\text{peso}}}$$

Media pesata:

$$\overline{x_w} = \sum_{i=1}^{n} x_i \times \color{red}{w_i}$$

pesi tali che: $w_i > 0, \quad \sum_{i=1}^{n} w_i = 1$

Media pesata

$$\overline{x_w} = \sum_{i=1}^{n} x_i \times w_i \qquad \text{pesi tali che: } \begin{cases} w_i = \dfrac{num\ cred.\ esame\ i-mo}{tot.\ crediti}, \\ \sum_{i=1}^{n} w_i = 1 \end{cases}$$

$$30 \times \frac{3}{12} + 22 \times \frac{3}{12} + 28 \times \frac{6}{12} = \frac{324}{12} = 28 \ \ (/30)$$

Stat. Infor. Matem.

$$\boxed{\frac{\sum_{i=1}^{n}(\text{voto esame}_i \times \text{numero crediti esame}_i)}{\text{totale numero crediti conseguiti}} \times}$$

Statistica

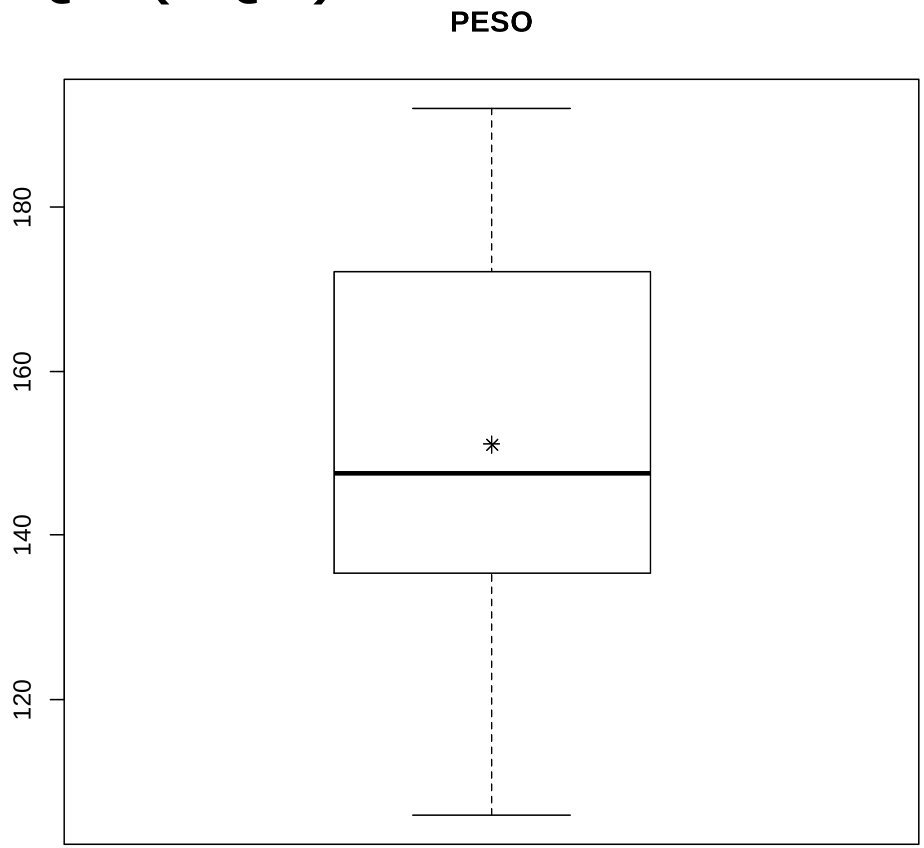
Statistica descrittiva

INDICI

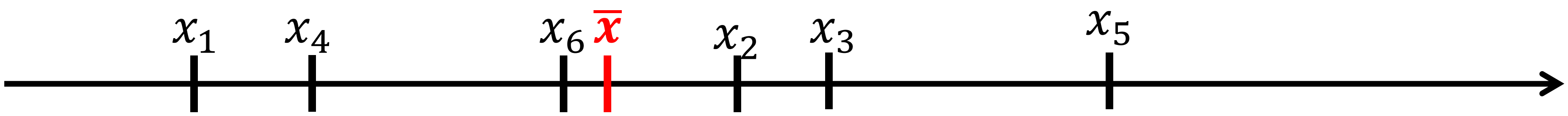
Indici di dispersione

Differenza interquartile: Q3-Q1 (IQR)

Range: Max-Min

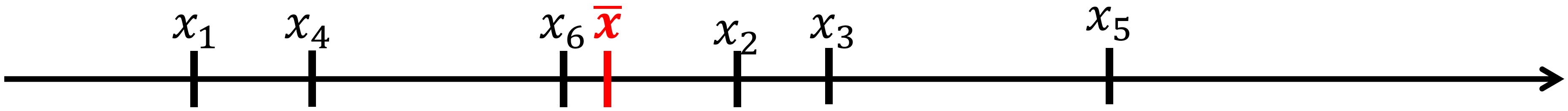


Indici di dispersione



Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

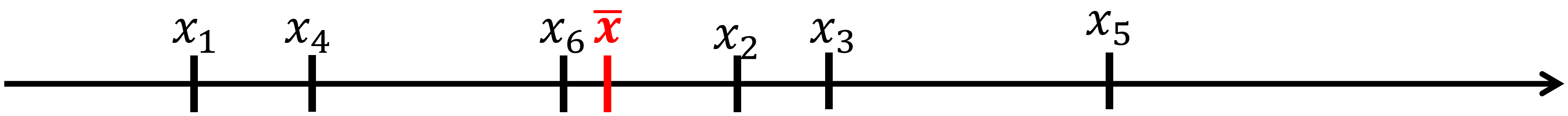
Indici di dispersione



Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

$$\sum_{i=1}^{n}(x_i - \overline{x})$$

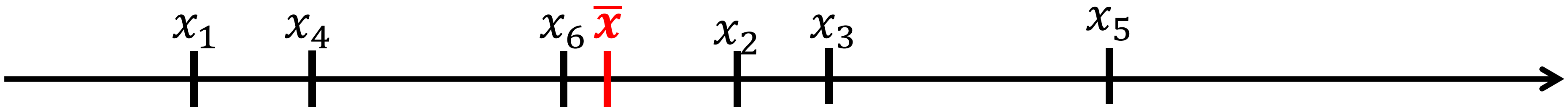
Indici di dispersione



Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \left(\sum_{i=1}^{n} x_i\right) - n\bar{x} \quad = n\bar{x} - n\bar{x} = 0 \text{ !!!!!!!}$$

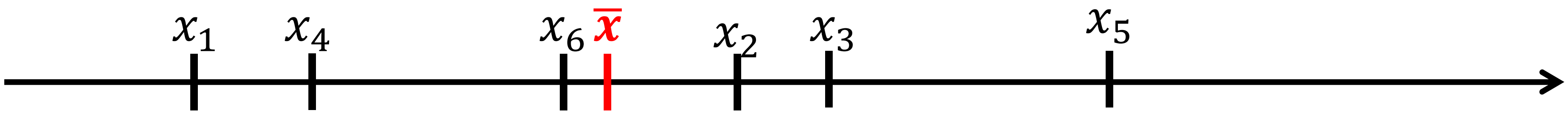
Indici di dispersione



Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

$$\sum_{i=1}^{n} (x_i - \bar{x})^2$$

Indici di dispersione



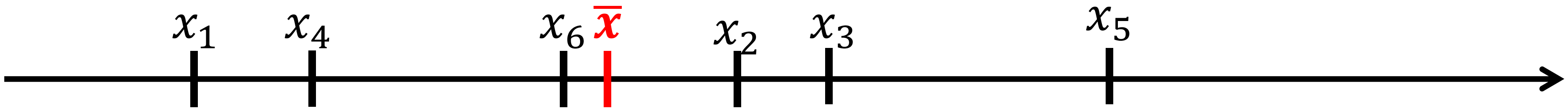
Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

$$0 \leq \sum_{i=1}^{n} (x_i - \overline{x})^2$$

MA $= 0 \Leftrightarrow$ I DATI SONO TUTTI UGUALI!

Ex: 2, 2, 2, 2, 2 $\Longrightarrow \quad \bar{x} = 2$

Indici di dispersione



Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

$$\boxed{\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Varianza, o scarto quadratico medio

Varianza

$$\bar{x} = \frac{19 + 22 + 21 + 23 + 22 + 20}{6} = 21.17$$ anni

	Età (y)	Peso (kg)	Altezza (m)	Sesso	Causa di morte
1	19	50.2	1.65	F	Nat.
2	22	75.6	1.78	M	Inc.
3	21	80.1	1.91	M	Inc.
4	23	56.7	1.72	M	Nat.
5	22	75.0	1.81	M	M.C.
6	20	58.3	1.68	F	Tum.

$$\sigma^2 = \frac{(19-21.17)^2 + (22-21.17)^2 + (21-21.17)^2}{6} +$$

$$+ \frac{(23-21.17)^2 + (22-21.17)^2 + (20-21.17)^2}{6} = 1.81 \quad \text{anni}^2$$

$\bar{x} = 21.17$

	Età (y)	Peso (kg)	Altezza (m)	Sesso	Causa di morte
1	19	50.2	1.65	F	Nat.
2	22	75.6	1.78	M	Inc.
3	21	80.1	1.91	M	Inc.
4	23	56.7	1.72	M	Nat.
5	22	75.0	1.81	M	M.C.
6	20	58.3	1.68	F	Tum.

Varianza

X : n. figli

x_i	n_i
0	4
1	6
2	4
3	1

$\boldsymbol{n = 15}$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$\bar{x} = 1.13$

$\sigma^2 =$

$$\frac{4 \times (0-1.13)^2 + 6 \times (1-1.13)^2 + 4 \times (2-1.13)^2 + 1 \times (3-1.13)^2}{15} = 0.782$$

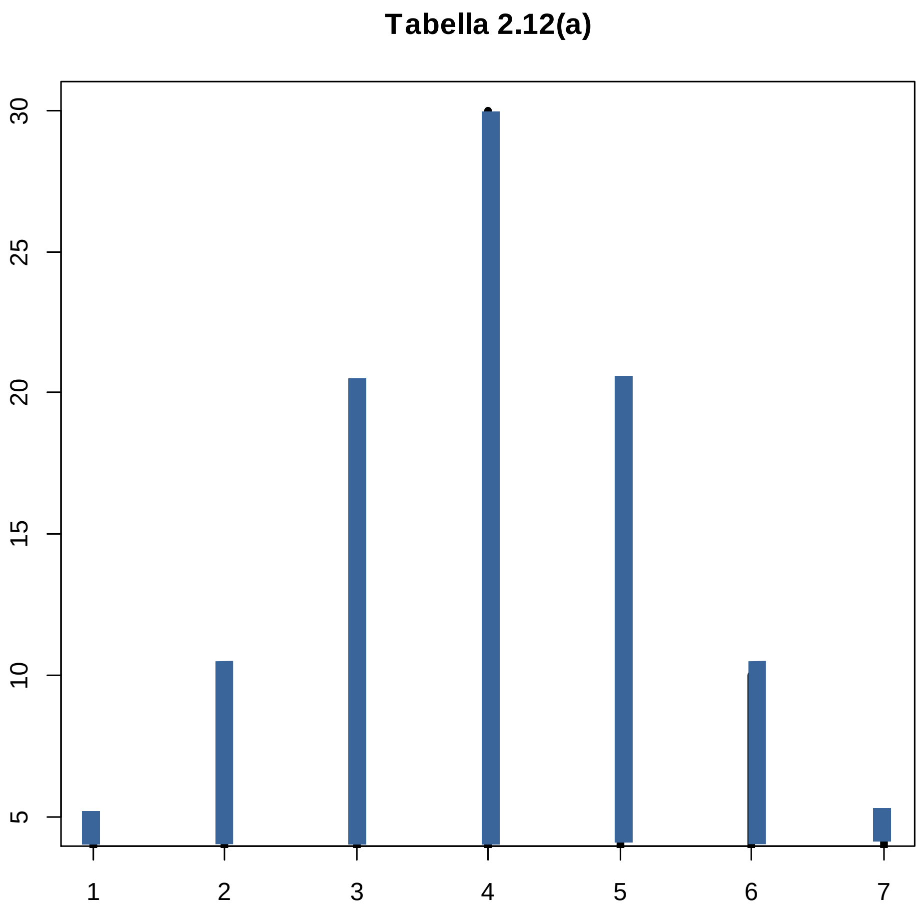
Varianza

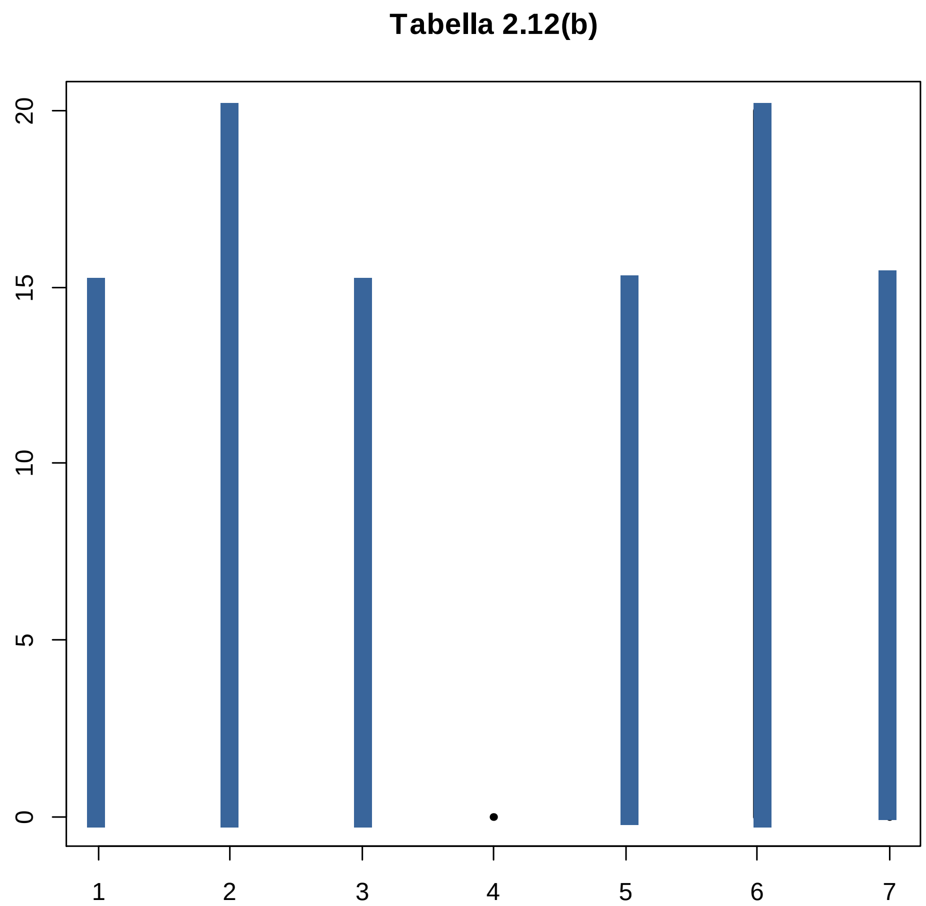
Prendiamo un punto di riferimento, la **media**, e guardiamo a come si disperdono i dati attorno ad esso

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\left(\sum_{i=1}^{n} {x_i}^2\right) - \bar{x}^2$$

(S.M. Iacus, *Statistica*)

Varianza

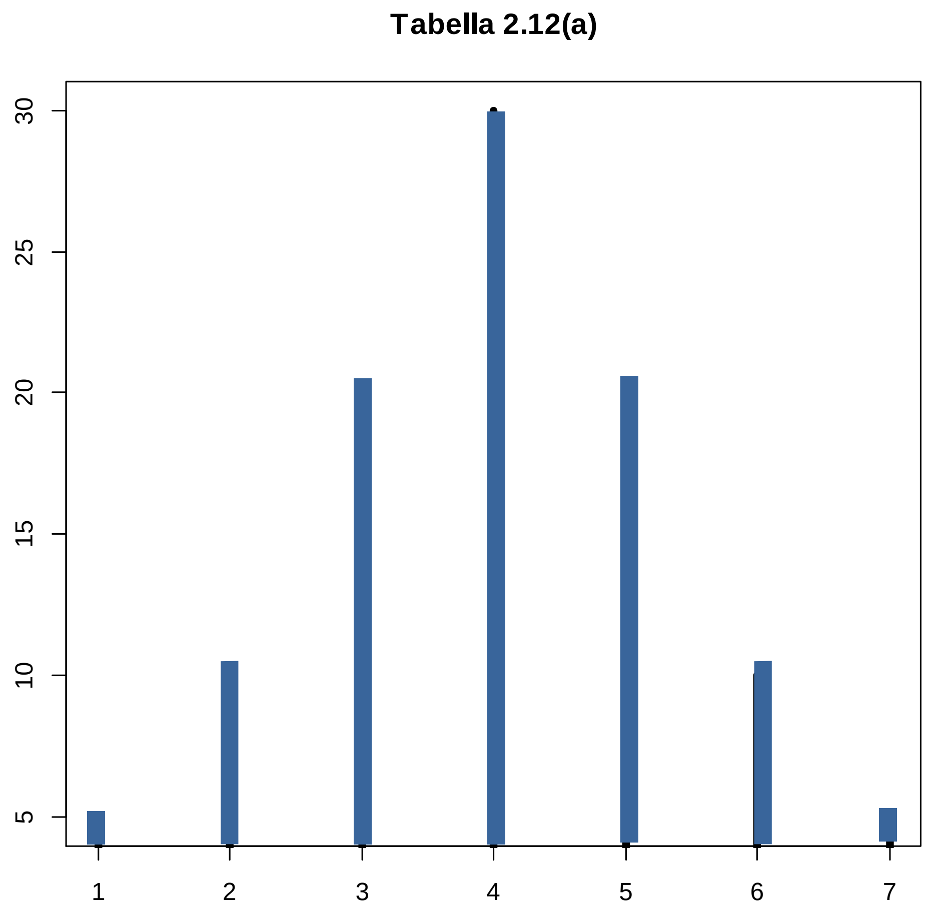


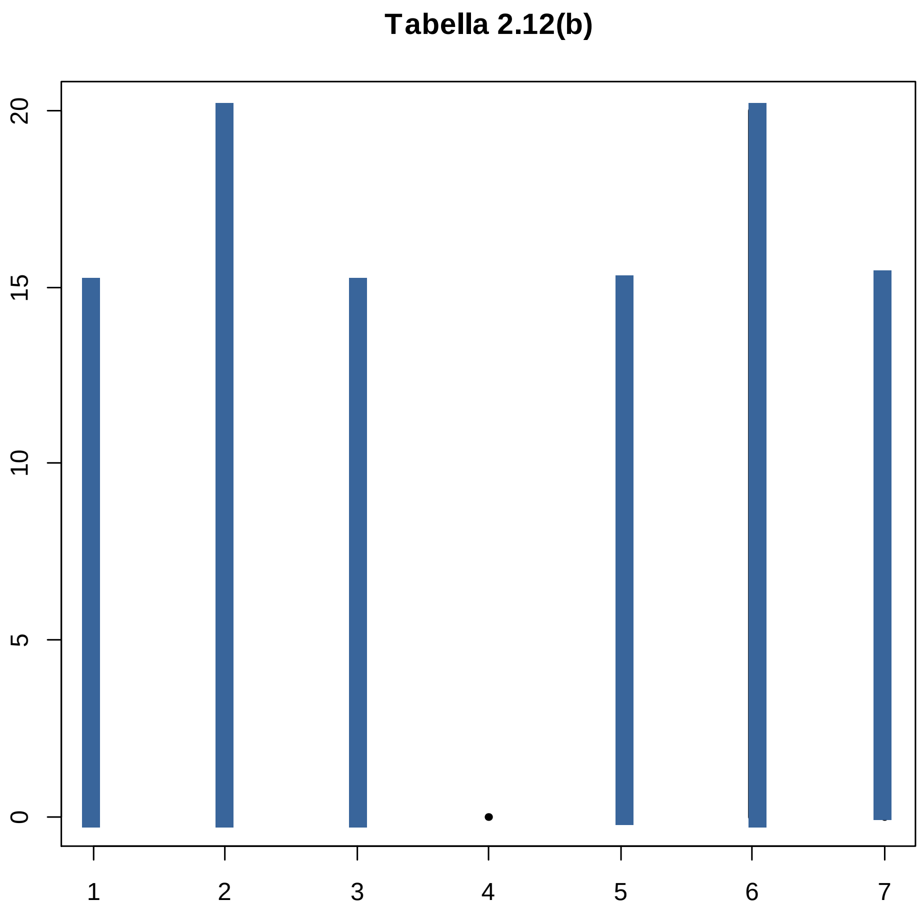


dati più frequenti **vicino a** 4

dati più frequenti **lontano da** 4

Varianza





dati più frequenti **vicino a** 4

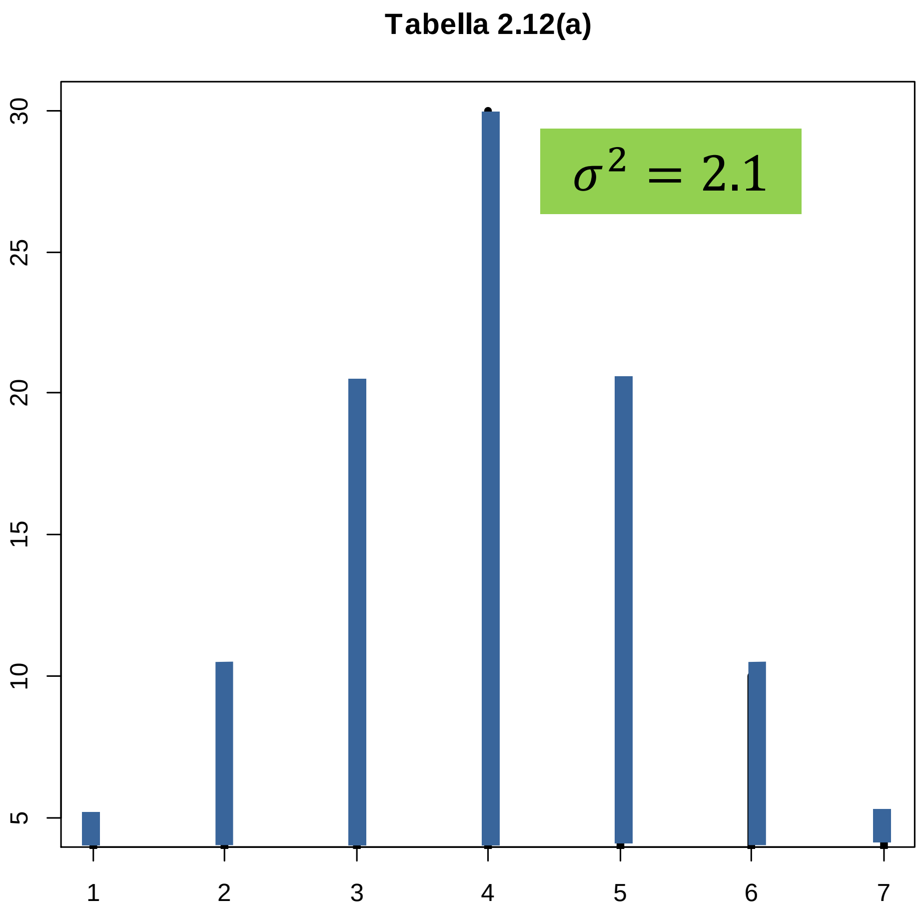
dati più frequenti **lontano da** 4

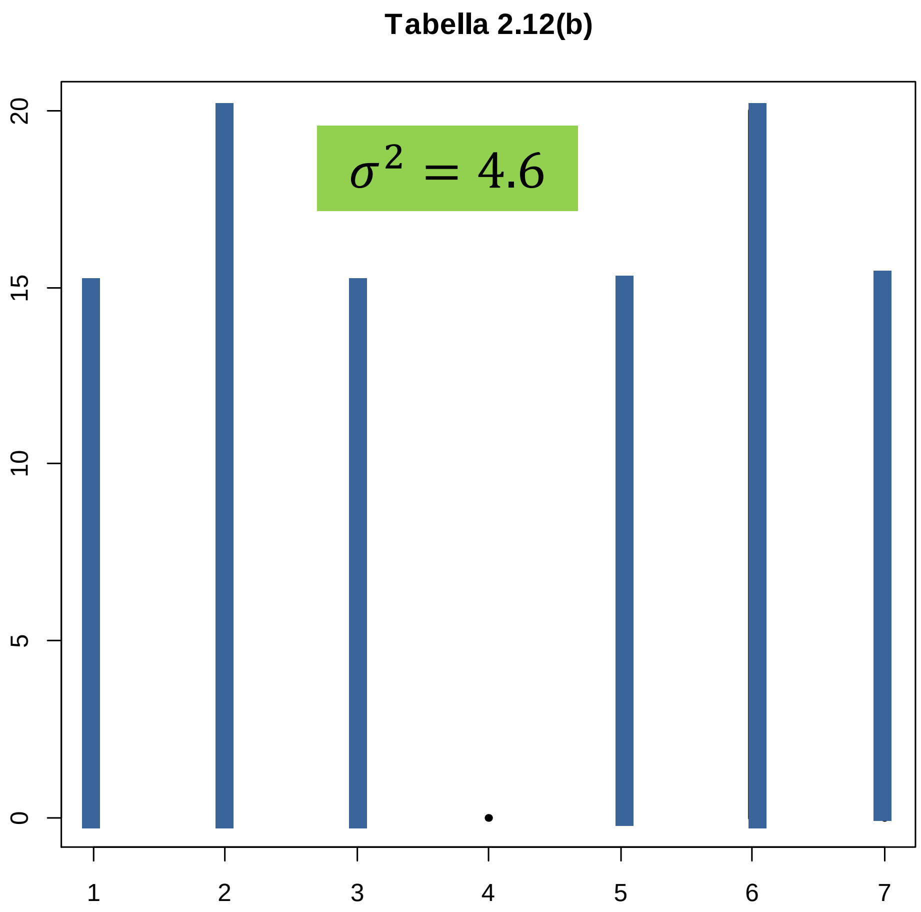
$\bar{x} = 4$

mediane pari a 4

$\bar{y} = 4$

Varianza





dati più frequenti **vicino a** 4

dati più frequenti **lontano da** 4

Proprietà della media

$$y_i = x_i + a \Rightarrow \boxed{\bar{y} = \bar{x} + a}$$

$$y_i = bx_i \Rightarrow \boxed{\bar{y} = b\bar{x}}$$

$$\boxed{y_i = bx_i + a \Rightarrow \bar{y} = b\bar{x} + a}$$

Proprietà della varianza

$$y_i = x_i + a \Rightarrow \boxed{\sigma^2{}_y = \sigma^2{}_x}$$

$$y_i = bx_i \Rightarrow \boxed{\sigma^2{}_y = b^2\sigma^2{}_x}$$

$$\boxed{y_i = bx_i + a \Rightarrow \sigma^2{}_y = b^2\sigma^2{}_x}$$

Varianza campionaria

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

varianza della popolazione

campione ***casuale***

$$\bar{s}_n{}^2 = \frac{1}{\boldsymbol{n-1}}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

varianza ***campionaria***