

Statistica - 1^a lezione (parte I)

23 febbraio 2021

ORARIO E DOCENTI

martedì 8:15 - 11:15
lezione / laboratorio (online)
proff. Alessandro Toigo / Francescopaolo Manicone
<code>alessandro.toigo@polimi.it</code>
<code>francescopaolo.manicone@polimi.it</code>

martedì 16.15 - 18.15 / giovedì 10.15 - 12.15
esercitazione (presenza)
prof. Tommaso Andreoli
<code>tommaso.andreoli@polimi.it</code>

TESTI CONSIGLIATI

- MONTGOMERY D.C., RUNGER G.C., HUBELE N.F., *Statistica per ingegneria - II edizione italiana*, Egea, 2012
- DALGAARD P., *Introductory Statistics with R*, Springer, 2008

MODALITÀ D'ESAME

- **Prova scritta (obbligatoria):**

- 5 appelli a giugno / luglio / settembre / gennaio / febbraio
- voto massimo prova scritta = 29

- **Orale di verifica (a discrezione del docente):**

- dopo la pubblicazione dei risultati dello scritto

- **Progetto di laboratorio (facoltativo):**

- si fa in gruppi di 3 persone
- si presenta a giugno / luglio / settembre
- voto $\in \{-1, 0, +1, +2, +3, +4, +5\}$
- il voto è valido per un anno e si somma al primo scritto sufficiente

ORARIO DI QUESTA SETTIMANA

mar 23 / 02	8:15 - 11:00	lezione	Toigo	online
	16:30 - 18:15			online
gio 25 / 02	10.15 - 12.00			

- **Statistica descrittiva**
(riassumere molti dati attraverso poche caratteristiche essenziali)
- **Probabilità**
(costruire un modello che preveda il risultato di un esperimento)
- **Inferenza statistica**
(tarare i parametri del modello in base ai risultati dell'esperimento)
- **Regressione lineare**
(riconoscere relazioni tra dati di tipo diverso)

- **Statistica descrittiva**
(riassumere molti dati attraverso poche caratteristiche essenziali)
- **Probabilità**
(costruire un modello che preveda il risultato di un esperimento)
- **Inferenza statistica**
(tarare i parametri del modello in base ai risultati dell'esperimento)
- **Regressione lineare**
(riconoscere relazioni tra dati di tipo diverso)

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati	{	CATEGORICI	(sesso, colore degli occhi)
		NUMERICI	{ DISCRETI (numero di figli, età) CONTINUI (altezza, peso)

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati $\left\{ \begin{array}{ll} \text{CATEGORICI} & (\text{sex, colore degli occhi}) \\ \text{NUMERICI} & \left\{ \begin{array}{ll} \text{DISCRETI} & (\text{numero di figli, età}) \\ \text{CONTINUI} & (\text{altezza, peso}) \end{array} \right. \end{array} \right.$

S = insieme dei valori possibili per ciascun dato

C = classe di valori (cioè un qualunque sottoinsieme $C \subseteq S$)

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati $\left\{ \begin{array}{ll} \text{CATEGORICI} & (\text{sezzo, colore degli occhi}) \\ \text{NUMERICI} & \left\{ \begin{array}{ll} \text{DISCRETI} & (\text{numero di figli, età}) \\ \text{CONTINUI} & (\text{altezza, peso}) \end{array} \right. \end{array} \right.$

S = insieme dei valori possibili per ciascun dato

C = classe di valori (cioè un qualunque sottoinsieme $C \subseteq S$)

ESEMPI:

$S = \{\text{maschio, femmina}\}$ $C_1 = \{\text{maschio}\}$, $C_2 = \{\text{femmina}\}$

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati $\left\{ \begin{array}{ll} \text{CATEGORICI} & (\text{sex, colore degli occhi}) \\ \text{NUMERICI} & \left\{ \begin{array}{ll} \text{DISCRETI} & (\text{numero di figli, età}) \\ \text{CONTINUI} & (\text{altezza, peso}) \end{array} \right. \end{array} \right.$

S = insieme dei valori possibili per ciascun dato

C = classe di valori (cioè un qualunque sottoinsieme $C \subseteq S$)

ESEMPI:

$S = \{\text{maschio, femmina}\}$ $C_1 = \{\text{maschio}\}$, $C_2 = \{\text{femmina}\}$

$S = \{0, 1, 2, 3, 4, 5, \dots\}$ $C_1 = \{0\}$, $C_2 = \{1, 2\}$, $C_3 = \{3, 4, 5, \dots\}$

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati $\left\{ \begin{array}{l} \text{CATEGORICI (sesso, colore degli occhi)} \\ \text{NUMERICI} \left\{ \begin{array}{l} \text{DISCRETI (numero di figli, età)} \\ \text{CONTINUI (altezza, peso)} \end{array} \right. \end{array} \right.$

S = insieme dei valori possibili per ciascun dato

C = classe di valori (cioè un qualunque sottoinsieme $C \subseteq S$)

ESEMPI:

$S = \{\text{maschio, femmina}\}$ $C_1 = \{\text{maschio}\}$, $C_2 = \{\text{femmina}\}$

$S = \{0, 1, 2, 3, 4, 5, \dots\}$ $C_1 = \{0\}$, $C_2 = \{1, 2\}$, $C_3 = \{3, 4, 5, \dots\}$

$S = [0, +\infty)$ $C_1 = [1.40, 1.60)$, $C_2 = [1.60, 1.90)$,
 $C_3 = [1.90, +\infty)$

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati $\left\{ \begin{array}{ll} \text{CATEGORICI} & (\text{sex, colore degli occhi}) \\ \text{NUMERICI} & \left\{ \begin{array}{ll} \text{DISCRETI} & (\text{numero di figli, età}) \\ \text{CONTINUI} & (\text{altezza, peso}) \end{array} \right. \end{array} \right.$

S = insieme dei valori possibili per ciascun dato

C = classe di valori (cioè un qualunque sottoinsieme $C \subseteq S$)

FREQUENZA ASSOLUTA della classe C =

$$= \text{FA}(C) := \#\{i \in \{1, 2, \dots, n\} \mid x_i \in C\}$$

I dati

DATO = risultato ottenuto in una singola misura

In genere, si ripetono più misure, ottenendo così n dati x_1, x_2, \dots, x_n

Tipi di dati $\left\{ \begin{array}{ll} \text{CATEGORICI} & (\text{sex, colore degli occhi}) \\ \text{NUMERICI} & \left\{ \begin{array}{ll} \text{DISCRETI} & (\text{numero di figli, età}) \\ \text{CONTINUI} & (\text{altezza, peso}) \end{array} \right. \end{array} \right.$

S = insieme dei valori possibili per ciascun dato

C = classe di valori (cioè un qualunque sottoinsieme $C \subseteq S$)

FREQUENZA ASSOLUTA della classe C =

$$= \text{FA}(C) := \#\{i \in \{1, 2, \dots, n\} \mid x_i \in C\}$$

FREQUENZA RELATIVA della classe C =

$$= \text{FR}(C) := \frac{1}{n} \text{FA}(C)$$

Dati categorici

ESEMPIO: Abbiamo misurato i gruppi sanguigni di $n = 12$ studenti, ottenendo i dati seguenti:

B B AB 0 A 0 A A A B A A

I valori possibili sono $S = \{A, B, AB, 0\}$

Come classi scegliamo insiemi formati da un solo elemento (singletoni)

$$C_1 = \{A\}, \quad C_2 = \{B\}, \quad C_3 = \{AB\}, \quad C_4 = \{0\}$$

Dati categorici

ESEMPIO: Abbiamo misurato i gruppi sanguigni di $n = 12$ studenti, ottenendo i dati seguenti:

B B AB 0 A 0 A A A B A A

I valori possibili sono $S = \{A, B, AB, 0\}$

Come classi scegliamo insiemi formati da un solo elemento (singletoni)

$$C_1 = \{A\}, \quad C_2 = \{B\}, \quad C_3 = \{AB\}, \quad C_4 = \{0\}$$

Ne ricaviamo le frequenze

$$f_A(\{A\}) = \#\{5, 7, 8, 9, 11, 12\} = 6$$

Dati categorici

ESEMPIO: Abbiamo misurato i gruppi sanguigni di $n = 12$ studenti, ottenendo i dati seguenti:

B B AB 0 A 0 A A A B A A

I valori possibili sono $S = \{A, B, AB, 0\}$

Come classi scegliamo insiemi formati da un solo elemento (singletoni)

$$C_1 = \{A\}, \quad C_2 = \{B\}, \quad C_3 = \{AB\}, \quad C_4 = \{0\}$$

Ne ricaviamo le frequenze

$$FA(\{A\}) = \#\{5, 7, 8, 9, 11, 12\} = 6$$

$$FA(\{B\}) = \#\{1, 2, 10\} = 3$$

Dati categorici

ESEMPIO: Abbiamo misurato i gruppi sanguigni di $n = 12$ studenti, ottenendo i dati seguenti:

B B **AB** 0 A 0 A A A B A A

I valori possibili sono $S = \{A, B, AB, 0\}$

Come classi scegliamo insiemi formati da un solo elemento (singletoni)

$$C_1 = \{A\}, \quad C_2 = \{B\}, \quad C_3 = \{AB\}, \quad C_4 = \{0\}$$

Ne ricaviamo le frequenze

$$FA(\{A\}) = \#\{5, 7, 8, 9, 11, 12\} = 6$$

$$FA(\{B\}) = \#\{1, 2, 10\} = 3$$

$$FA(\{\mathbf{AB}\}) = \#\{\mathbf{3}\} = 1$$

Dati categorici

ESEMPIO: Abbiamo misurato i gruppi sanguigni di $n = 12$ studenti, ottenendo i dati seguenti:

B B AB 0 A 0 A A A B A A

I valori possibili sono $S = \{A, B, AB, 0\}$

Come classi scegliamo insiemi formati da un solo elemento (singletoni)

$$C_1 = \{A\}, \quad C_2 = \{B\}, \quad C_3 = \{AB\}, \quad C_4 = \{0\}$$

Ne ricaviamo le frequenze

$$FA(\{A\}) = \#\{5, 7, 8, 9, 11, 12\} = 6$$

$$FA(\{B\}) = \#\{1, 2, 10\} = 3$$

$$FA(\{AB\}) = \#\{3\} = 1$$

$$FA(\{0\}) = \#\{4, 6\} = 2$$

Dati categorici

ESEMPIO: Abbiamo misurato i gruppi sanguigni di $n = 12$ studenti, ottenendo i dati seguenti:

B B AB 0 A 0 A A A B A A

I valori possibili sono $S = \{A, B, AB, 0\}$

Come classi scegliamo insiemi formati da un solo elemento (singletoni)

$$C_1 = \{A\}, \quad C_2 = \{B\}, \quad C_3 = \{AB\}, \quad C_4 = \{0\}$$

Ne ricaviamo le frequenze

$$FA(\{A\}) = \#\{5, 7, 8, 9, 11, 12\} = 6 \Rightarrow FR(\{A\}) = 6/12 = 0.5$$

$$FA(\{B\}) = \#\{1, 2, 10\} = 3 \Rightarrow FR(\{B\}) = 3/12 = 0.25$$

$$FA(\{AB\}) = \#\{3\} = 1 \Rightarrow FR(\{AB\}) = 1/12 = 0.08\bar{3}$$

$$FA(\{0\}) = \#\{4, 6\} = 2 \Rightarrow FR(\{0\}) = 2/12 = 0.1\bar{6}$$

**Tabella
delle frequenze**

Classi	FA	FR
A	6	$0.5 = 50\%$
B	3	$0.25 = 25\%$
AB	1	$0.08\bar{3} = 8.\bar{3}\%$
0	2	$0.1\bar{6} = 16.\bar{6}\%$

**Tabella
delle frequenze**

Classi	FA	FR
A	6	$0.5 = 50\%$
B	3	$0.25 = 25\%$
AB	1	$0.08\bar{3} = 8.\bar{3}\%$
0	2	$0.1\bar{6} = 16.\bar{6}\%$
12 OK!		

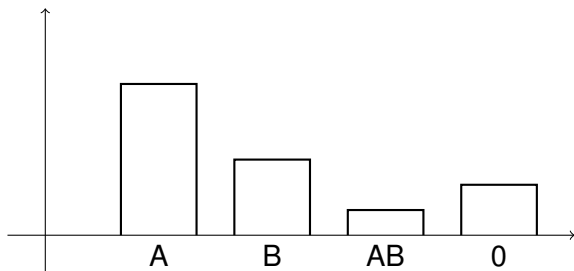
**Tabella
delle frequenze**

Classi	FA	FR
A	6	$0.5 = 50\%$
B	3	$0.25 = 25\%$
AB	1	$0.08\bar{3} = 8.\bar{3}\%$
0	2	$0.1\bar{6} = 16.\bar{6}\%$
	12 OK!	100% OK!

**Tabella
delle frequenze**

Classi	FA	FR
A	6	$0.5 = 50\%$
B	3	$0.25 = 25\%$
AB	1	$0.08\bar{3} = 8.\bar{3}\%$
0	2	$0.1\bar{6} = 16.\bar{6}\%$
	12 OK!	100% OK!

Istogramma

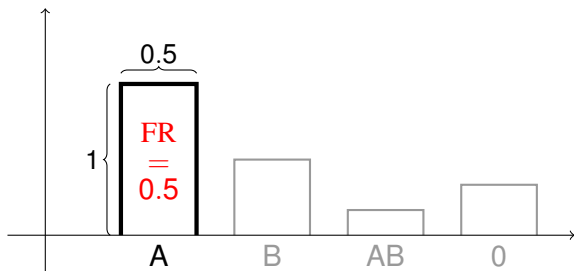


Dati categorici

**Tabella
delle frequenze**

Classi	FA	FR
A	6	$0.5 = 50\%$
B	3	$0.25 = 25\%$
AB	1	$0.08\bar{3} = 8.\bar{3}\%$
0	2	$0.1\bar{6} = 16.\bar{6}\%$
	12 OK!	100% OK!

Istogramma



Dati numerici

I dati numerici si possono

- ordinare:

$$x_1 = 28.4 \quad x_2 = 8.3 \quad x_3 = 15.7$$

dati grezzi

$$x_{(1)} = 8.3 \quad x_{(2)} = 15.7 \quad x_{(3)} = 28.4$$

dati ordinati

Dati numerici

I dati numerici si possono

- ordinare:

$$x_1 = 28.4 \quad x_2 = 8.3 \quad x_3 = 15.7 \quad \text{dati grezzi}$$

$$x_{(1)} = 8.3 \quad x_{(2)} = 15.7 \quad x_{(3)} = 28.4 \quad \text{dati ordinati}$$

FREQUENZA CUMULATA del valore $k =$

$$= FC(k) := \frac{1}{n} \# \{i \in \{1, 2, \dots, n\} \mid x_i \leq k\}$$

Dati numerici

I dati numerici si possono

- ordinare:

$$\begin{array}{llll} x_1 = 28.4 & x_2 = 8.3 & x_3 = 15.7 & \text{dati grezzi} \\ x_{(1)} = 8.3 & x_{(2)} = 15.7 & x_{(3)} = 28.4 & \text{dati ordinati} \end{array}$$

FREQUENZA CUMULATA del valore $k =$

$$= FC(k) := \frac{1}{n} \# \{i \in \{1, 2, \dots, n\} \mid x_i \leq k\}$$

- sommare e moltiplicare per costanti:

$$x_i \longrightarrow y_i = ax_i + b \quad a, b \in \mathbb{R}$$

$$\begin{array}{llll} x_1 = 28.4 & x_2 = 8.3 & x_3 = 15.7 & \text{temperature in } ^\circ\text{C} \\ x_i \longrightarrow y_i = 1.8 \cdot x_i + 32 & & & \text{conversione in } ^\circ\text{F} \\ y_1 = 83.1 & y_2 = 46.9 & y_3 = 60.3 & \text{temperature in } ^\circ\text{F} \end{array}$$

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori possibili sono $S = \{0, 1, 2, 3, \dots, 100\}$

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori **veramente** possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0			
1			
2			
3			
4			
5			
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1			
2			
3			
4			
5			
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1 0 3 1 3 2 2 1 2 5 3 0
1 4 3 7 1 3 1 4 2 3 0 5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2			
3			
4			
5			
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2	4	$4/24 = 0.1\bar{6}$	$0.375 + 0.1\bar{6} = 0.541\bar{6}$
3			
4			
5			
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1 0 3 1 3 2 2 1 2 5 3 0
1 4 3 7 1 3 1 4 2 3 0 5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2	4	$4/24 = 0.1\bar{6}$	$0.375 + 0.1\bar{6} = 0.541\bar{6}$
3	6	$6/24 = 0.25$	$0.541\bar{6} + 0.25 = 0.791\bar{6}$
4			
5			
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2	4	$4/24 = 0.1\bar{6}$	$0.375 + 0.1\bar{6} = 0.541\bar{6}$
3	6	$6/24 = 0.25$	$0.541\bar{6} + 0.25 = 0.791\bar{6}$
4	2	$2/24 = 0.08\bar{3}$	$0.791\bar{6} + 0.08\bar{3} = 0.875$
5			
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2	4	$4/24 = 0.1\bar{6}$	$0.375 + 0.1\bar{6} = 0.541\bar{6}$
3	6	$6/24 = 0.25$	$0.541\bar{6} + 0.25 = 0.791\bar{6}$
4	2	$2/24 = 0.08\bar{3}$	$0.791\bar{6} + 0.08\bar{3} = 0.875$
5	2	$2/24 = 0.08\bar{3}$	$0.875 + 0.08\bar{3} = 0.958\bar{3}$
6			
7			

Dati numerici discreti

ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2	4	$4/24 = 0.1\bar{6}$	$0.375 + 0.1\bar{6} = 0.541\bar{6}$
3	6	$6/24 = 0.25$	$0.541\bar{6} + 0.25 = 0.791\bar{6}$
4	2	$2/24 = 0.08\bar{3}$	$0.791\bar{6} + 0.08\bar{3} = 0.875$
5	2	$2/24 = 0.08\bar{3}$	$0.875 + 0.08\bar{3} = 0.958\bar{3}$
6	0	$0/24 = 0$	$0.958\bar{3} + 0 = 0.958\bar{3}$
7			

Dati numerici discreti

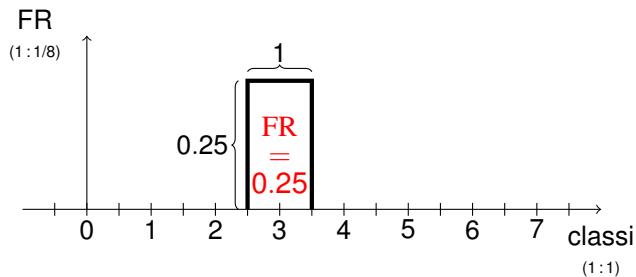
ESEMPIO: Abbiamo contato il numero di CD difettosi in 24 scatole da 100 pezzi ciascuna, ottenendo i dati seguenti:

1	0	3	1	3	2	2	1	2	5	3	0
1	4	3	7	1	3	1	4	2	3	0	5

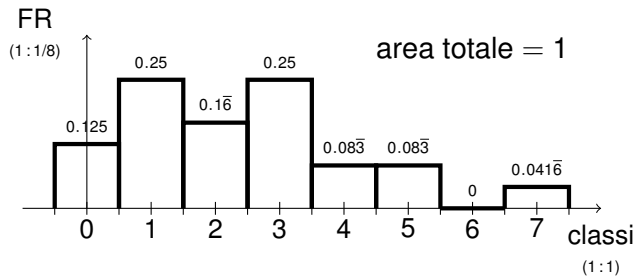
I valori veramente possibili sono $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

Classi	FA	FR	FC
0	3	$3/24 = 0.125$	0.125
1	6	$6/24 = 0.25$	$0.125 + 0.25 = 0.375$
2	4	$4/24 = 0.1\bar{6}$	$0.375 + 0.1\bar{6} = 0.541\bar{6}$
3	6	$6/24 = 0.25$	$0.541\bar{6} + 0.25 = 0.791\bar{6}$
4	2	$2/24 = 0.08\bar{3}$	$0.791\bar{6} + 0.08\bar{3} = 0.875$
5	2	$2/24 = 0.08\bar{3}$	$0.875 + 0.08\bar{3} = 0.958\bar{3}$
6	0	$0/24 = 0$	$0.958\bar{3} + 0 = 0.958\bar{3}$
7	1	$1/24 = 0.041\bar{6}$	$0.958\bar{3} + 0.041\bar{6} = 1$ OK!

Istogramma

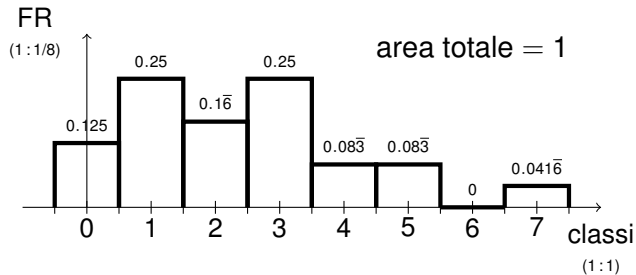


Istogramma

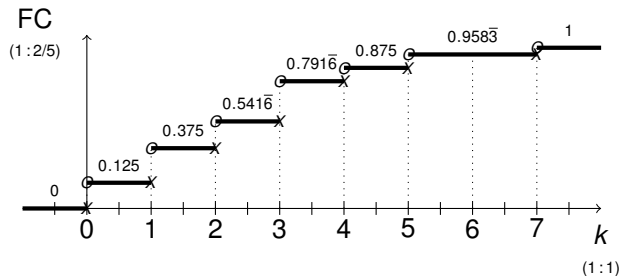


Dati numerici discreti

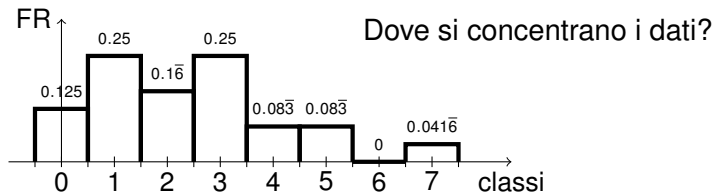
Istogramma



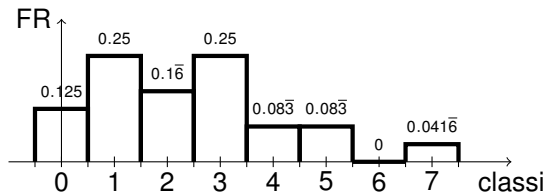
Frequenze
cumulate



Indici di posizione

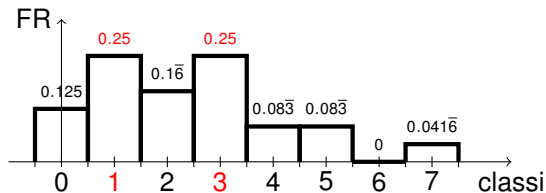


Indici di posizione



MODA = classe (o classi) più frequenti

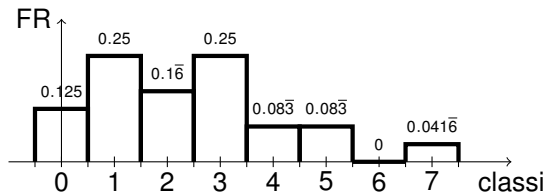
Indici di posizione



MODA = classe (o classi) più frequenti

= {1, 3} \Rightarrow dati **bimodali**

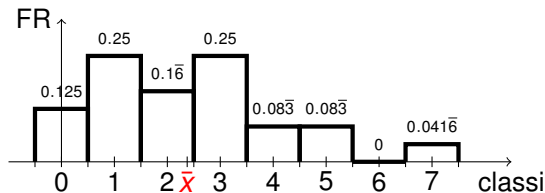
Indici di posizione



MODA = classe (o classi) più frequenti

MEDIA CAMPIONARIA $= \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$

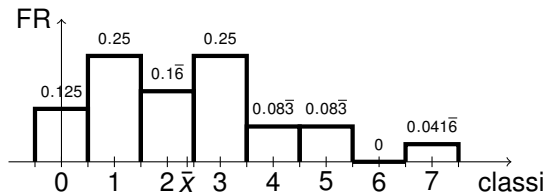
Indici di posizione



MODA = classe (o classi) più frequenti

$$\begin{aligned}\text{MEDIA CAMPIONARIA} &= \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n \\ &= \frac{1 + 0 + 3 + \dots + 5}{24} = 2.375\end{aligned}$$

Indici di posizione



MODA = classe (o classi) più frequenti

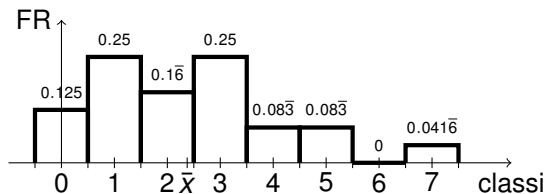
$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

dove $\lfloor \cdot \rfloor$ = intero immediatamente precedente:

$$\lfloor 7.93 \rfloor = 7, \quad \lfloor 2.08 \rfloor = 2, \quad \lfloor 5 \rfloor = 5, \quad \lfloor 0.57 \rfloor = 0$$

Indici di posizione



	x_1, x_2, \dots, x_{24}					
1	0	3	1	3	2	
2	1	2	5	3	0	
1	4	3	7	1	3	
1	4	2	3	0	5	

MODA = classe (o classi) più frequenti

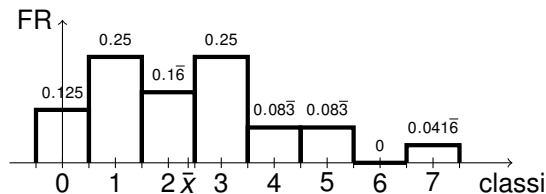
$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

dove $\lfloor \cdot \rfloor$ = intero immediatamente precedente:

$$\lfloor 7.93 \rfloor = 7, \quad \lfloor 2.08 \rfloor = 2, \quad \lfloor 5 \rfloor = 5, \quad \lfloor 0.57 \rfloor = 0$$

Indici di posizione



$X_{(1)}, X_{(2)}, \dots, X_{(24)}$					
0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

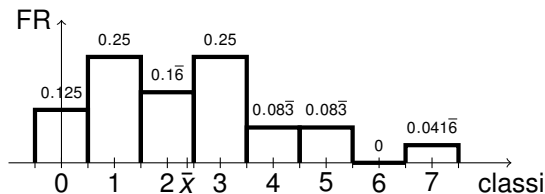
MODA = classe (o classi) più frequenti

$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

$$n/2 = 24/2 = 12 \in \mathbb{N}$$

Indici di posizione



$X_{(1)}, X_{(2)}, \dots, X_{(24)}$					
0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

MODA = classe (o classi) più frequenti

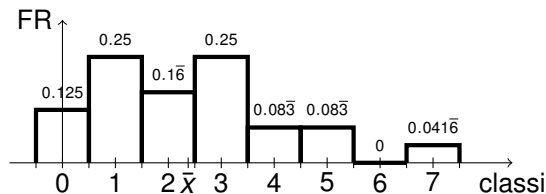
$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

$$n/2 = 24/2 = 12 \in \mathbb{N}$$

$$\Rightarrow m = \frac{1}{2} (x_{(12)} + x_{(13)})$$

Indici di posizione



$X_{(1)}, X_{(2)}, \dots, X_{(24)}$					
0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

MODA = classe (o classi) più frequenti

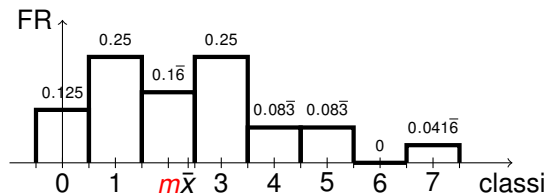
$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

$$n/2 = 24/2 = 12 \in \mathbb{N}$$

$$\Rightarrow m = \frac{1}{2} (x_{(12)} + x_{(13)}) = \frac{1}{2} (2 + 2) = 2$$

Indici di posizione



$X_{(1)}, X_{(2)}, \dots, X_{(24)}$					
0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

MODA = classe (o classi) più frequenti

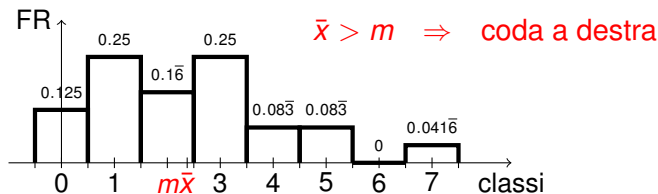
$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

$$n/2 = 24/2 = 12 \in \mathbb{N}$$

$$\Rightarrow m = \frac{1}{2} (x_{(12)} + x_{(13)}) = \frac{1}{2} (2 + 2) = 2$$

Indici di posizione



MODA = classe (o classi) più frequenti

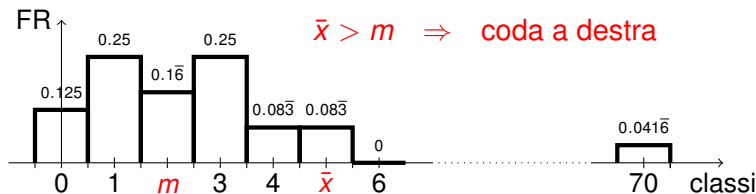
$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

$$n/2 = 24/2 = 12 \in \mathbb{N}$$

$$\Rightarrow m = \frac{1}{2} (x_{(12)} + x_{(13)}) = \frac{1}{2} (2 + 2) = 2$$

Indici di posizione



MODA = classe (o classi) più frequenti

$$\text{MEDIA CAMPIONARIA} = \frac{1}{n} \sum_{i=1}^n x_i =: \bar{x}_n$$

$$\text{MEDIANA} = \begin{cases} x_{(\lfloor n/2 \rfloor + 1)} & \text{se } n/2 \notin \mathbb{N} \\ \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)}) & \text{se } n/2 \in \mathbb{N} \end{cases} =: m$$

$$n/2 = 24/2 = 12 \in \mathbb{N}$$

$$\Rightarrow m = \frac{1}{2} (x_{(12)} + x_{(13)}) = \frac{1}{2} (2 + 2) = 2$$

- Per dati **discreti**:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

Proprietà della media campionaria

- Per dati **discreti**:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

Per esempio:

$$\bar{x} = \frac{1 + 0 + 3 + 1 + 3 + \dots + 0 + 5}{24}$$

Proprietà della media campionaria

- Per dati **discreti**:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

Per esempio:

$$\begin{aligned}\bar{x} &= \frac{1 + 0 + 3 + 1 + 3 + \dots + 0 + 5}{24} \\ &= \frac{\overbrace{0 + 0 + \dots + 0}^{\text{FA}(0) \text{ volte}} + \overbrace{1 + 1 + \dots + 1}^{\text{FA}(1) \text{ volte}} + \dots + \overbrace{7 + 7 + \dots + 7}^{\text{FA}(7) \text{ volte}}}{24}\end{aligned}$$

Proprietà della media campionaria

- Per dati **discreti**:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

Per esempio:

$$\begin{aligned}\bar{x} &= \frac{1 + 0 + 3 + 1 + 3 + \dots + 0 + 5}{24} \\&= \frac{\overbrace{0 + 0 + \dots + 0}^{\text{FA}(0) \text{ volte}} + \overbrace{1 + 1 + \dots + 1}^{\text{FA}(1) \text{ volte}} + \dots + \overbrace{7 + 7 + \dots + 7}^{\text{FA}(7) \text{ volte}}}{24} \\&= 0 \cdot \underbrace{\frac{\text{FA}(0)}{24}}_{\text{FR}(0)} + 1 \cdot \underbrace{\frac{\text{FA}(1)}{24}}_{\text{FR}(1)} + \dots + 7 \cdot \underbrace{\frac{\text{FA}(7)}{24}}_{\text{FR}(7)}\end{aligned}$$

Proprietà della media campionaria

- Per dati **discreti**:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

Per esempio:

$$\begin{aligned}\bar{x} &= \frac{1 + 0 + 3 + 1 + 3 + \dots + 0 + 5}{24} \\&= \frac{\overbrace{0 + 0 + \dots + 0}^{\text{FA}(0) \text{ volte}} + \overbrace{1 + 1 + \dots + 1}^{\text{FA}(1) \text{ volte}} + \dots + \overbrace{7 + 7 + \dots + 7}^{\text{FA}(7) \text{ volte}}}{24} \\&= 0 \cdot \underbrace{\frac{\text{FA}(0)}{24}}_{\text{FR}(0)} + 1 \cdot \underbrace{\frac{\text{FA}(1)}{24}}_{\text{FR}(1)} + \dots + 7 \cdot \underbrace{\frac{\text{FA}(7)}{24}}_{\text{FR}(7)} \\&= \sum_{k=0}^7 k \text{FR}(k)\end{aligned}$$

Proprietà della media campionaria

- Per dati discreti:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$\bar{y}_n = a\bar{x}_n + b$$

Proprietà della media campionaria

- Per dati discreti:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$\bar{y}_n = a\bar{x}_n + b$$

Infatti:

$$\bar{y}_n := \frac{1}{n} \sum_{i=1}^n y_i$$

- Per dati discreti:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$\bar{y}_n = a\bar{x}_n + b$$

Infatti:

$$\bar{y}_n := \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (ax_i + b)$$

Proprietà della media campionaria

- Per dati discreti:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$\bar{y}_n = a\bar{x}_n + b$$

Infatti:

$$\begin{aligned}\bar{y}_n &:= \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (ax_i + b) \\ &= a \frac{1}{n} \sum_{i=1}^n x_i + b \frac{1}{n} \sum_{i=1}^n 1\end{aligned}$$

Proprietà della media campionaria

- Per dati discreti:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$\bar{y}_n = a\bar{x}_n + b$$

Infatti:

$$\begin{aligned}\bar{y}_n &:= \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (ax_i + b) \\ &= a \frac{1}{n} \sum_{i=1}^n x_i + b \frac{1}{n} \sum_{i=1}^n 1 = a\bar{x}_n + b \frac{1}{n} n\end{aligned}$$

Proprietà della media campionaria

- Per dati discreti:

$$\bar{x}_n = \sum_{\text{classi } k} k \text{FR}(k)$$

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$\bar{y}_n = a\bar{x}_n + b$$

Infatti:

$$\begin{aligned}\bar{y}_n &:= \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (ax_i + b) \\ &= a \frac{1}{n} \sum_{i=1}^n x_i + b \frac{1}{n} \sum_{i=1}^n 1 = a\bar{x}_n + b \frac{1}{n} n \\ &= a\bar{x}_n + b\end{aligned}$$

Proprietà della mediana

- Almeno il 50% dei dati è $\leq m$
- Almeno il 50% dei dati è $\geq m$

Proprietà della mediana

- Almeno il 50% dei dati è $\leq m$
- Almeno il 50% dei dati è $\geq m$
- Dalla tabella delle frequenze:

Classi	FA	FR	FC	
0	3	0.125	0.125	solo il 12.5% dei dati ≤ 0
1	6	0.25	0.375	solo il 37.5% dei dati ≤ 1
2	4	0.16	0.5416	$m = 2$
3	6	0.25	0.7916	solo il $(100 - 54.16)\%$ dei dati ≥ 3
4	2	0.083	0.875	solo il $(100 - 79.16)\%$ dei dati ≥ 4
5	2	0.083	0.9583	solo il $(100 - 87.5)\%$ dei dati ≥ 5
6	0	0	0.9583	solo il $(100 - 95.83)\%$ dei dati ≥ 6
7	1	0.0416	1	solo il $(100 - 95.83)\%$ dei dati ≥ 7

Proprietà della mediana

- Almeno il 50% dei dati è $\leq m$
- Almeno il 50% dei dati è $\geq m$
- Dalla tabella delle frequenze:

Classi	FA	FR	FC	
0	3	0.125	0.125	solo il 12.5% dei dati ≤ 0
1	6	0.25	0.375	solo il 37.5% dei dati ≤ 1
2	4	0.16	0.5416	$m = 2$
3	6	0.25	0.7916	solo il $(100 - 54.16)\%$ dei dati ≥ 3
4	2	0.083	0.875	solo il $(100 - 79.16)\%$ dei dati ≥ 4
5	2	0.083	0.9583	solo il $(100 - 87.5)\%$ dei dati ≥ 5
6	0	0	0.9583	solo il $(100 - 95.83)\%$ dei dati ≥ 6
7	1	0.0416	1	solo il $(100 - 95.83)\%$ dei dati ≥ 7

- Per una **trasformazione affine** $y_i = ax_i + b$ con $a \neq 0$:

$$m_y = am_x + b$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

- Almeno il $100 \cdot \gamma$ % dei dati è $\leq q_\gamma$
- Almeno il $100 \cdot (1 - \gamma)$ % dei dati è $\geq q_\gamma$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

- Almeno il $100 \cdot \gamma$ % dei dati è $\leq q_\gamma$
- Almeno il $100 \cdot (1 - \gamma)$ % dei dati è $\geq q_\gamma$
- $q_{0.25} =: Q1$, $q_{0.5} = m =: Q2$, $q_{0.75} =: Q3$: **quartili**

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

- Almeno il $100 \cdot \gamma$ % dei dati è $\leq q_\gamma$
- Almeno il $100 \cdot (1 - \gamma)$ % dei dati è $\geq q_\gamma$
- $q_{0.25} =: Q1$, $q_{0.5} = m =: Q2$, $q_{0.75} =: Q3$: quartili
- q_γ si può ricavare dalla tabella delle frequenze

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

- Almeno il $100 \cdot \gamma$ % dei dati è $\leq q_\gamma$
- Almeno il $100 \cdot (1 - \gamma)$ % dei dati è $\geq q_\gamma$
- $q_{0.25} =: Q1$, $q_{0.5} = m =: Q2$, $q_{0.75} =: Q3$: quartili
- q_γ si può ricavare dalla tabella delle frequenze
- Per una trasformazione affine $y_i = ax_i + b$ con $a \neq 0$:

$$q_\gamma^y = \begin{cases} a q_\gamma^x + b & \text{se } a > 0 \\ a q_{1-\gamma}^x + b & \text{se } a < 0 \end{cases}$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$q_{0.20} = ? ? ? ?$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$q_{0.20} = \text{????}$$

$$n\gamma = 24 \cdot 0.20 = 4.8$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$q_{0.20} = x_{(\lfloor 4.8 \rfloor + 1)}$$

$$n\gamma = 24 \cdot 0.20 = 4.8$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$x_{(1)}, x_{(2)}, \dots, x_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$q_{0.20} = x_{(\lfloor 4.8 \rfloor + 1)} = x_{(5)}$$

$$n\gamma = 24 \cdot 0.20 = 4.8$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$q_{0.20} = x_{(\lfloor 4.8 \rfloor + 1)} = x_{(5)} = 1$$

$$n\gamma = 24 \cdot 0.20 = 4.8$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q1 = q_{0.25} = \text{????}$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q1 = q_{0.25} = \text{????}$$

$$n\gamma = 24 \cdot 0.25 = 6$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$x_{(1)}, x_{(2)}, \dots, x_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q1 = q_{0.25} = \frac{1}{2} (x_{(6)} + x_{(7)})$$

$$n\gamma = 24 \cdot 0.25 = 6$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$x_{(1)}, x_{(2)}, \dots, x_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q1 = q_{0.25} = \frac{1}{2} (x_{(6)} + x_{(7)}) = \frac{1}{2} (1 + 1)$$

$$n\gamma = 24 \cdot 0.25 = 6$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$x_{(1)}, x_{(2)}, \dots, x_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q1 = q_{0.25} = \frac{1}{2} (x_{(6)} + x_{(7)}) = \frac{1}{2} (1 + 1) = 1 \quad n\gamma = 24 \cdot 0.25 = 6$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q3 = q_{0.75} = \text{????}$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q3 = q_{0.75} = \text{????}$$

$$n\gamma = 24 \cdot 0.75 = 18$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$X_{(1)}, X_{(2)}, \dots, X_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q3 = q_{0.75} = \frac{1}{2} (x_{(18)} + x_{(19)})$$

$$n\gamma = 24 \cdot 0.75 = 18$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

ESEMPIO:

$$x_{(1)}, x_{(2)}, \dots, x_{(24)}$$

0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q3 = q_{0.75} = \frac{1}{2} (x_{(18)} + x_{(19)}) = \frac{1}{2} (\mathbf{3} + \mathbf{3})$$

$$n\gamma = 24 \cdot 0.75 = 18$$

QUANTILE DI ORDINE γ (con $\gamma \in (0, 1)$ fissato)

$$= \begin{cases} x_{(\lfloor n\gamma \rfloor + 1)} & \text{se } n\gamma \notin \mathbb{N} \\ \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & \text{se } n\gamma \in \mathbb{N} \end{cases} =: q_\gamma$$

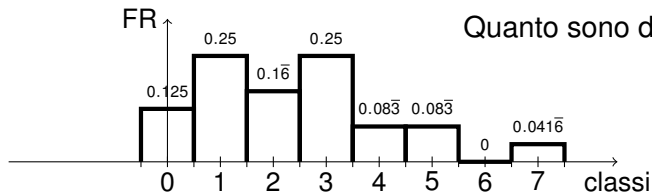
ESEMPIO:

$$x_{(1)}, x_{(2)}, \dots, x_{(24)}$$

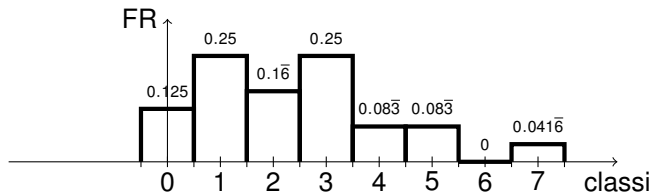
0	0	0	1	1	1
1	1	1	2	2	2
2	3	3	3	3	3
3	4	4	5	5	7

Classi	FA	FR	FC
0	3	0.125	0.125
1	6	0.25	0.375
2	4	0.16	0.5416
3	6	0.25	0.7916
4	2	0.083	0.875
5	2	0.083	0.9583
6	0	0	0.9583
7	1	0.0416	1

$$Q3 = q_{0.75} = \frac{1}{2} (x_{(18)} + x_{(19)}) = \frac{1}{2} (3 + 3) = 3 \quad n\gamma = 24 \cdot 0.75 = 18$$

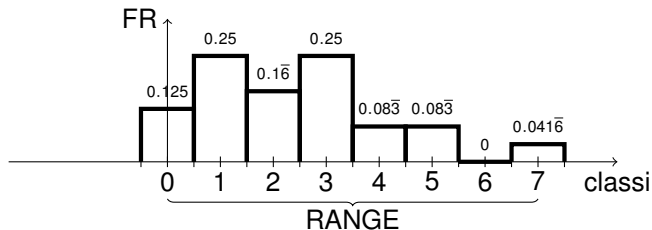


Indici di dispersione



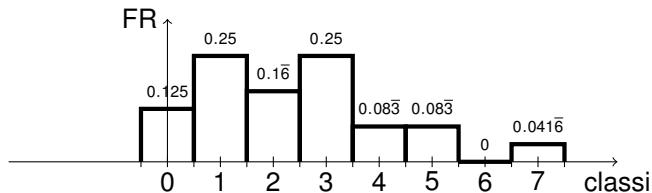
$$\text{RANGE} = x_{(n)} - x_{(1)}$$

Indici di dispersione



$$\begin{aligned}\text{RANGE} &= x_{(n)} - x_{(1)} \\ &= 7 - 0 = 7\end{aligned}$$

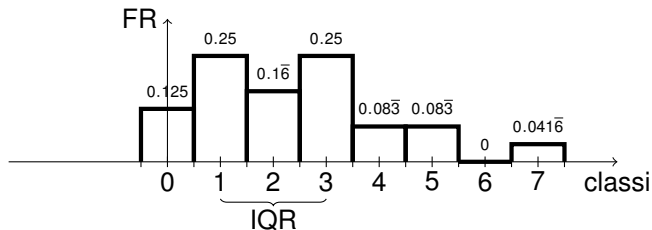
Indici di dispersione



$$\text{RANGE} = x_{(n)} - x_{(1)}$$

$$\text{RANGE INTERQUARTILE} = Q3 - Q1 =: \text{IQR}$$

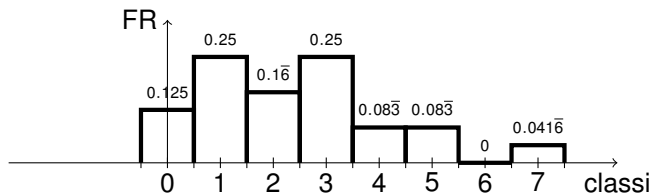
Indici di dispersione



$$\text{RANGE} = x_{(n)} - x_{(1)}$$

$$\begin{aligned}\text{RANGE INTERQUARTILE} &= Q3 - Q1 =: \text{IQR} \\ &= 3 - 1 = 2\end{aligned}$$

Indici di dispersione

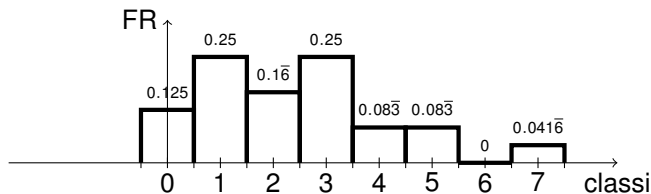


$$\text{RANGE} = x_{(n)} - x_{(1)}$$

$$\text{RANGE INTERQUARTILE} = Q3 - Q1 =: \text{IQR}$$

$$\text{VARIANZA CAMPIONARIA} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 =: s_n^2$$

Indici di dispersione

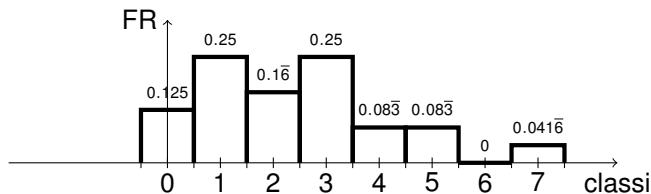


$$\text{RANGE} = x_{(n)} - x_{(1)}$$

$$\text{RANGE INTERQUARTILE} = Q3 - Q1 =: \text{IQR}$$

$$\begin{aligned}\text{VARIANZA CAMPIONARIA} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 =: s_n^2 \\ &= \frac{(1 - 2.375)^2 + (0 - 2.375)^2 + \dots + (5 - 2.375)^2}{24 - 1} \simeq 3.1141\end{aligned}$$

Indici di dispersione



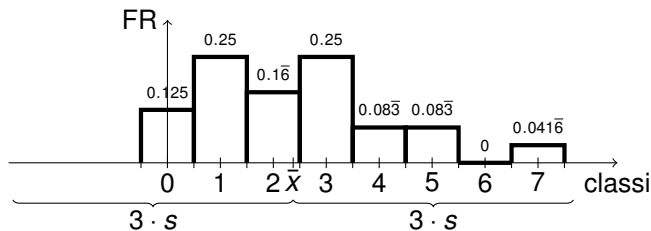
$$\text{RANGE} = x_{(n)} - x_{(1)}$$

$$\text{RANGE INTERQUARTILE} = Q3 - Q1 =: \text{IQR}$$

$$\text{VARIANZA CAMPIONARIA} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 =: s_n^2$$

$$\text{DEVIAZIONE STANDARD CAMPIONARIA} = \sqrt{s_n^2} =: s_n$$

Indici di dispersione



$$\text{RANGE} = x_{(n)} - x_{(1)}$$

$$\text{RANGE INTERQUARTILE} = Q3 - Q1 =: \text{IQR}$$

$$\text{VARIANZA CAMPIONARIA} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 =: s_n^2$$

$$\begin{aligned} \text{DEVIAZIONE STANDARD CAMPIONARIA} &= \sqrt{s_n^2} =: s_n \\ &\simeq 1.7647 \end{aligned}$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la **disuguaglianza di Chebyshev**

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la **disuguaglianza di Chebyshev**

$$\begin{aligned}\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} &\leq \frac{n-1}{k^2} \quad \forall k > 0 \\ &= \frac{24-1}{3^2} \leq 2 \quad \text{con } k = 3\end{aligned}$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la disuguaglianza di Chebyshev

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

- Vale la **formula alternativa**

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

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la disuguaglianza di Chebyshev

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

- Vale la **formula alternativa**

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

Infatti:

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la disuguaglianza di Chebyshev

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

- Vale la **formula alternativa**

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

Infatti:

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{n-1} \sum (x_i^2 - 2x_i \bar{x} + \bar{x}^2)$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la disuguaglianza di Chebyshev

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

- Vale la **formula alternativa**

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

Infatti:

$$\begin{aligned} s^2 &= \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{n-1} \sum (x_i^2 - 2x_i \bar{x} + \bar{x}^2) \\ &= \frac{1}{n-1} \left(\sum x_i^2 - 2\bar{x} \sum x_i + \bar{x}^2 \sum 1 \right) \end{aligned}$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la disuguaglianza di Chebyshev

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

- Vale la **formula alternativa**

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

Infatti:

$$\begin{aligned} s^2 &= \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{n-1} \sum (x_i^2 - 2x_i \bar{x} + \bar{x}^2) \\ &= \frac{1}{n-1} \left(\sum x_i^2 - 2\bar{x} \sum x_i + \bar{x}^2 \sum 1 \right) \\ &= \frac{1}{n-1} \left(\sum x_i^2 - 2n\bar{x}\bar{x} + \bar{x}^2 n \right) \end{aligned}$$

Proprietà di s^2 e di s

- $[s^2] = [x_i]^2$, mentre $[s] = [\bar{x}] = [x_i]$.
- Vale la disuguaglianza di Chebyshev

$$\#\{i \in \{1, 2, \dots, n\} \mid |x_i - \bar{x}| \geq k s\} \leq \frac{n-1}{k^2} \quad \forall k > 0$$

- Vale la formula alternativa

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

- Per dati **discreti**:

$$s_n^2 = \frac{n}{n-1} \sum_{\text{classi } k} (k - \bar{x})^2 \text{FR}(k)$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

Infatti:

$$s_y^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

Infatti:

$$\begin{aligned} s_y^2 &= \frac{1}{n-1} \sum (y_i - \bar{y})^2 \\ &= \frac{1}{n-1} \sum [(ax_i + b) - (a\bar{x} + b)]^2 \end{aligned}$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

Infatti:

$$\begin{aligned} s_y^2 &= \frac{1}{n-1} \sum (y_i - \bar{y})^2 \\ &= \frac{1}{n-1} \sum [(ax_i + b) - (a\bar{x} + b)]^2 \\ &= \frac{1}{n-1} \sum [a(x_i - \bar{x})]^2 \end{aligned}$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

Infatti:

$$\begin{aligned} s_y^2 &= \frac{1}{n-1} \sum (y_i - \bar{y})^2 \\ &= \frac{1}{n-1} \sum [(ax_i + b) - (a\bar{x} + b)]^2 \\ &= \frac{1}{n-1} \sum [a(x_i - \bar{x})]^2 = \frac{a^2}{n-1} \sum (x_i - \bar{x})^2 \end{aligned}$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

Infatti:

$$\begin{aligned} s_y^2 &= \frac{1}{n-1} \sum (y_i - \bar{y})^2 \\ &= \frac{1}{n-1} \sum [(ax_i + b) - (a\bar{x} + b)]^2 \\ &= \frac{1}{n-1} \sum [a(x_i - \bar{x})]^2 = \frac{a^2}{n-1} \sum (x_i - \bar{x})^2 \\ &= a^2 s_x^2 \end{aligned}$$

Proprietà di s^2 e di s

- Per una **trasformazione affine** $y_i = ax_i + b$:

$$s_y^2 = a^2 s_x^2$$

$$s_y = |a| s_x$$

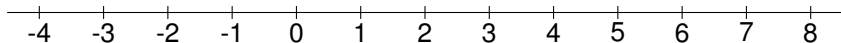
Infatti:

$$\begin{aligned} s_y^2 &= \frac{1}{n-1} \sum (y_i - \bar{y})^2 \\ &= \frac{1}{n-1} \sum [(ax_i + b) - (a\bar{x} + b)]^2 \\ &= \frac{1}{n-1} \sum [a(x_i - \bar{x})]^2 = \frac{a^2}{n-1} \sum (x_i - \bar{x})^2 \\ &= a^2 s_x^2 \end{aligned}$$

$$s_y = \sqrt{s_y^2} = \sqrt{a^2 s_x^2} = |a| \sqrt{s_x^2} = |a| s_x$$

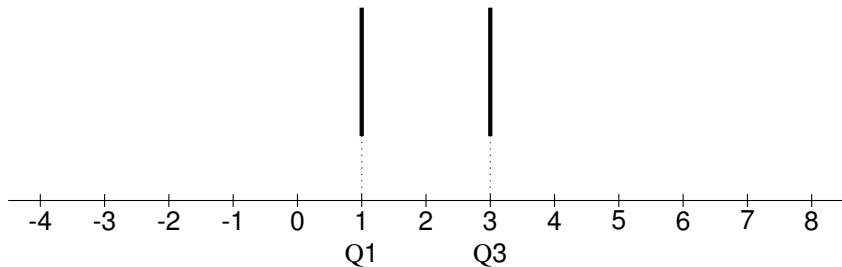
Boxplot

Si disegna una riga (orizzontale o verticale):



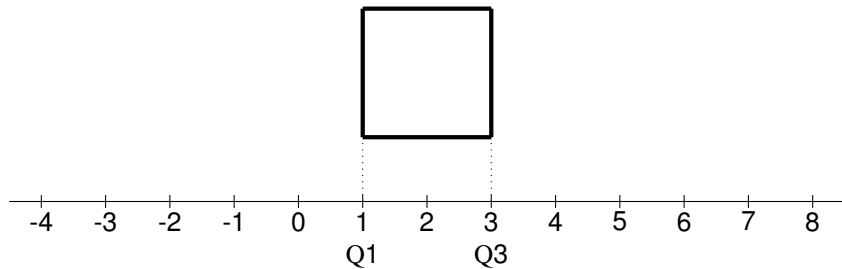
Boxplot

Si tirano due righe in corrispondenza di Q1 e di Q3:



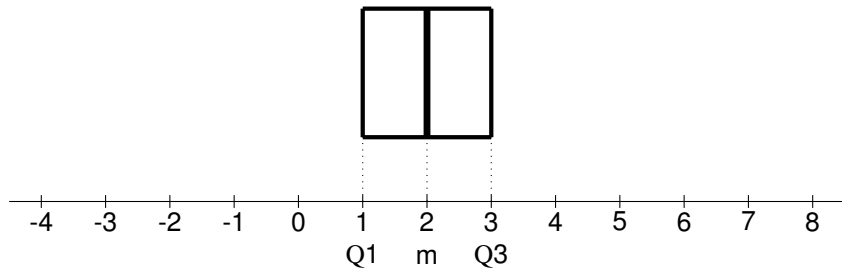
Boxplot

Si chiude la scatola:



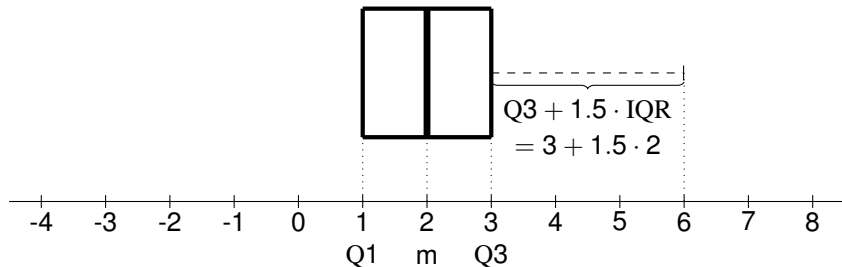
Boxplot

Si tira una riga spessa in corrispondenza di m :



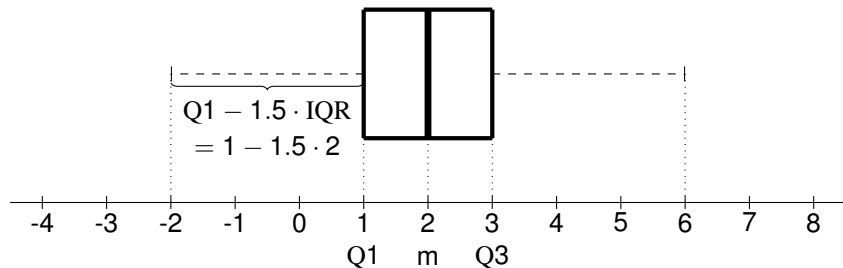
Boxplot

Si disegna un baffo da Q3 a $Q3 + 1.5 \cdot IQR \dots$



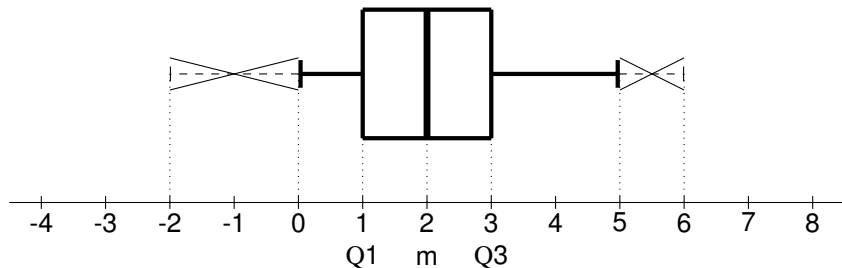
Boxplot

... e un altro baffo da $Q1$ a $Q1 - 1.5 \cdot IQR$:



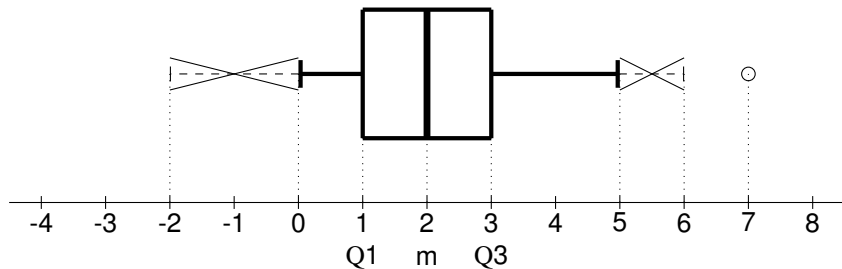
Boxplot

Dopodiché si tagliano i baffi in corrispondenza dei dati più esterni compresi entro di essi:



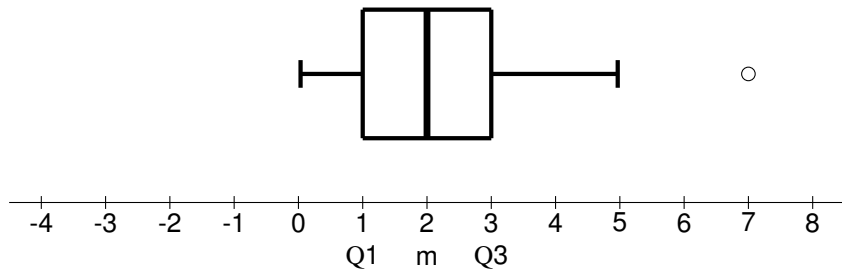
Boxplot

Eventuali dati esterni ai baffi si segnalano come *outlier*.

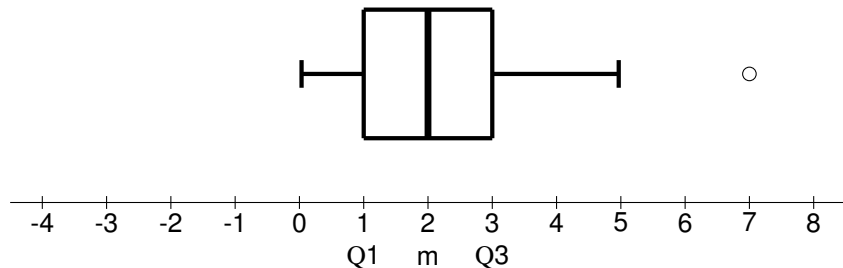


Boxplot

Ecco il risultato finale:



Boxplot



- Almeno il 50% dei dati è nella scatola
- Si vede un accenno di coda sulla destra

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati x_1, x_2, \dots, x_{30} :

14.7	12.3	18.3	10.2	11.5	15.1	14.2	14.7	13.8	17.3
25.3	26.4	31.0	19.4	17.5	17.6	16.8	18.0	13.8	10.7
12.6	14.5	17.8	19.6	17.2	13.1	13.9	14.2	13.7	18.1

I valori possibili sono $S = \mathbb{R}$

Tutti gli indici di posizione e di dispersione si calcolano come prima:

$$\bar{x} = \frac{14.7 + 12.3 + \dots + 18.1}{30} = 16.44\bar{3}$$

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati **ordinati** $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

I valori possibili sono $S = \mathbb{R}$

Tutti gli indici di posizione e di dispersione si calcolano come prima:

$$\bar{x} = \frac{14.7 + 12.3 + \dots + 18.1}{30} = 16.44\bar{3}$$

$$\begin{aligned} q_{0.5} &= \frac{1}{2} (x_{(n\gamma)} + x_{(n\gamma+1)}) & n\gamma &= 30 \cdot 0.5 = 15 \in \mathbb{N} \\ &= \frac{1}{2} (x_{(15)} + x_{(16)}) = \frac{1}{2} (14.7 + 15.1) \\ &= 14.9 \end{aligned}$$

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

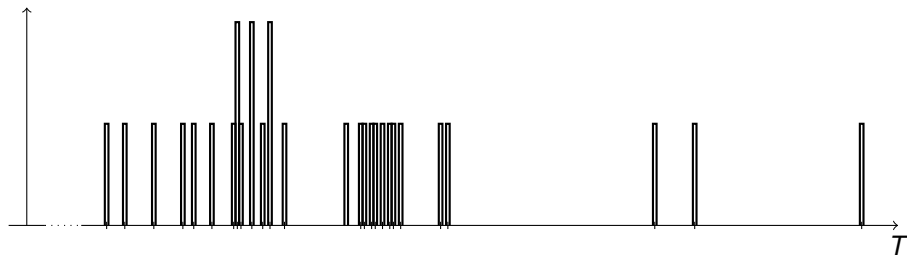
Ma come si disegna l'istogramma?

Dati continui

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

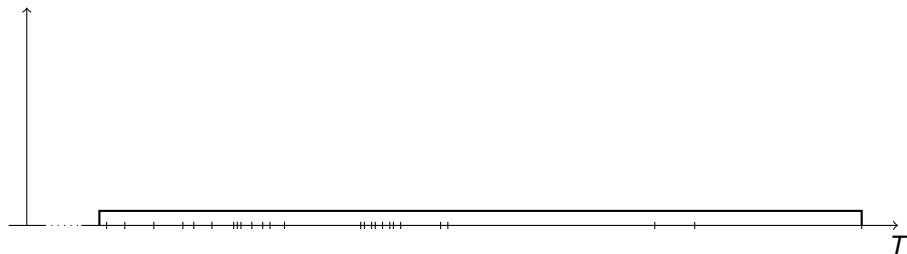
Se uso classi **troppo numerose**:



ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Se uso classi **troppo poco numerose**:



ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Il numero giusto di classi è

$$\# \text{classi} = \sqrt{n}$$

oppure

$$\# \text{classi} = 1 + \log_2 n = 1 + \frac{\log_z n}{\log_z 2}$$

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Il numero giusto di classi è

$$\# \text{classi} = \sqrt{n} = \sqrt{30} \simeq 5.477$$

oppure

$$\# \text{classi} = 1 + \log_2 n = 1 + \frac{\log_z n}{\log_z 2} = 1 + \frac{\ln 30}{\ln 2} \simeq 5.907$$

$$\Rightarrow \text{ scelgo 6 classi di ampiezza } \frac{31.0 - 10.2}{6} = 3.5\bar{6} \simeq 4$$

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi			
(10, 14]			
(14, 18]			
(18, 22]			
(22, 26]			
(26, 30]			
(30, 34]			

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA		
(10, 14]	10		
(14, 18]			
(18, 22]			
(22, 26]			
(26, 30]			
(30, 34]			

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA		
(10, 14]	10		
(14, 18]	13		
(18, 22]			
(22, 26]			
(26, 30]			
(30, 34]			

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA		
(10, 14]	10		
(14, 18]	13		
(18, 22]	4		
(22, 26]			
(26, 30]			
(30, 34]			

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA		
(10, 14]	10		
(14, 18]	13		
(18, 22]	4		
(22, 26]	1		
(26, 30]			
(30, 34]			

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA		
(10, 14]	10		
(14, 18]	13		
(18, 22]	4		
(22, 26]	1		
(26, 30]	1		
(30, 34]			

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA		
(10, 14]	10		
(14, 18]	13		
(18, 22]	4		
(22, 26]	1		
(26, 30]	1		
(30, 34]	1		

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA	FR	
(10, 14]	10	$10/30 \simeq 0.33$	
(14, 18]	13	$13/30 \simeq 0.43$	
(18, 22]	4	$4/30 \simeq 0.13$	
(22, 26]	1	$1/30 \simeq 0.03$	
(26, 30]	1	$1/30 \simeq 0.03$	
(30, 34]	1	$1/30 \simeq 0.03$	

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA	FR	FC
(10, 14]	10	$10/30 \simeq 0.33$	0.33
(14, 18]	13	$13/30 \simeq 0.43$	$0.33 + 0.43 = 0.76$
(18, 22]	4	$4/30 \simeq 0.13$	$0.76 + 0.13 = 0.89$
(22, 26]	1	$1/30 \simeq 0.03$	$0.89 + 0.03 = 0.92$
(26, 30]	1	$1/30 \simeq 0.03$	$0.92 + 0.03 = 0.95$
(30, 34]	1	$1/30 \simeq 0.03$	$0.95 + 0.03 = 0.98$

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

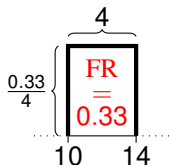
Classi	FA	FR	FC
(10, 14]	10	$10/30 \simeq 0.33$	0.33
(14, 18]	13	$13/30 \simeq 0.43$	$0.33 + 0.43 = 0.76$
(18, 22]	4	$4/30 \simeq 0.14$	$0.76 + 0.14 = 0.90$
(22, 26]	1	$1/30 \simeq 0.04$	$0.90 + 0.04 = 0.94$
(26, 30]	1	$1/30 \simeq 0.03$	$0.94 + 0.03 = 0.97$
(30, 34]	1	$1/30 \simeq 0.03$	$0.97 + 0.03 = 1.00$

Dati continui

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

Classi	FA	FR	FC
$(10, 14]$	10	$10/30 \simeq 0.33$	0.33
$(14, 18]$	13	$13/30 \simeq 0.43$	$0.33 + 0.43 = 0.76$
$(18, 22]$	4	$4/30 \simeq 0.14$	$0.76 + 0.14 = 0.90$
$(22, 26]$	1	$1/30 \simeq 0.04$	$0.90 + 0.04 = 0.94$
$(26, 30]$	1	$1/30 \simeq 0.03$	$0.94 + 0.03 = 0.97$
$(30, 34]$	1	$1/30 \simeq 0.03$	$0.97 + 0.03 = 1.00$



$$\text{DENSITÀ}(C) = \frac{\text{FR}(C)}{\text{ampiezza}(C)}$$

ESEMPIO: Abbiamo misurato le temperature massime giornaliere (in °C) dell'ultimo mese, ottenendo questi dati ordinati $x_{(1)}, x_{(2)}, \dots, x_{(30)}$:

10.2	10.7	11.5	12.3	12.6	13.1	13.7	13.8	13.8	13.9
14.2	14.2	14.5	14.7	14.7	15.1	16.8	17.2	17.3	17.5
17.6	17.8	18.0	18.1	18.3	19.4	19.6	25.3	26.4	31.0

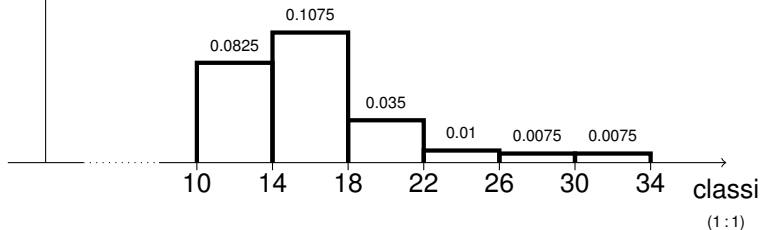
Classi	FA	FR	FC	Densità
(10, 14]	10	$10/30 \simeq 0.33$	0.33	$0.33/4 \simeq 0.0825$
(14, 18]	13	$13/30 \simeq 0.43$	$0.33 + 0.43 = 0.76$	$0.43/4 \simeq 0.1075$
(18, 22]	4	$4/30 \simeq 0.14$	$0.76 + 0.14 = 0.90$	$0.14/4 \simeq 0.035$
(22, 26]	1	$1/30 \simeq 0.04$	$0.90 + 0.04 = 0.94$	$0.04/4 \simeq 0.01$
(26, 30]	1	$1/30 \simeq 0.03$	$0.94 + 0.03 = 0.97$	$0.03/4 \simeq 0.0075$
(30, 34]	1	$1/30 \simeq 0.03$	$0.97 + 0.03 = 1.00$	$0.03/4 \simeq 0.0075$

$$\text{DENSITÀ}(C) = \frac{\text{FR}(C)}{\text{ampiezza}(C)}$$

Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075

densità

(1 : 1/16)



Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075

$$\bar{x} \simeq \sum_{\text{classi } C} \text{punto medio}(C) \cdot \text{FR}(C)$$

Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075

$$\begin{aligned}\bar{x} &\simeq \sum_{\text{classi } C} \text{punto medio}(C) \cdot \text{FR}(C) \\ &= \frac{10 + 14}{2} \cdot 0.33 + \frac{14 + 18}{2} \cdot 0.43 + \dots + \frac{30 + 34}{2} \cdot 0.03 \\ &= 16.4\end{aligned}$$

da confrontarsi con

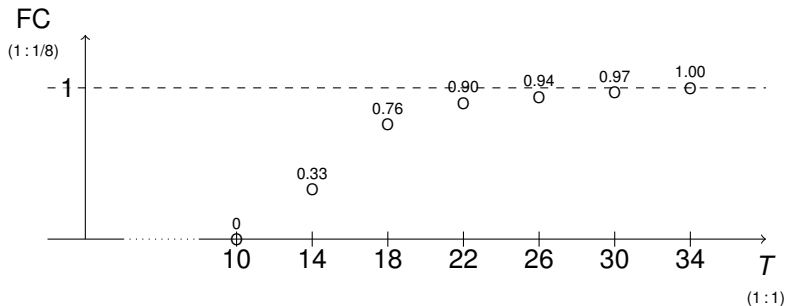
$$\bar{x} = 16.44\bar{3}$$

Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075

$$\left. \begin{array}{l} FC(14) = 0.33 < 0.50 \Rightarrow m \text{ non può essere } \leq 14 \\ 1 - FC(18) = 0.24 < 0.50 \Rightarrow m \text{ non può essere } > 18 \end{array} \right\} \Rightarrow$$
$$\Rightarrow m \in (14, 18]$$

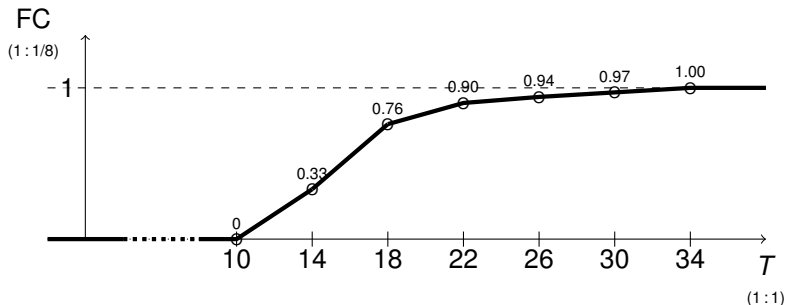
Ma se volessi dare un valore più preciso?

Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075



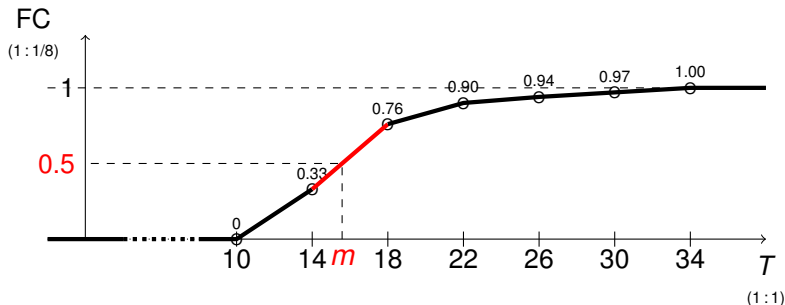
Dati continui

Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075

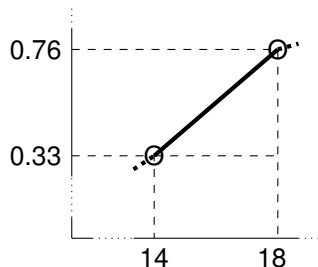


Dati continui

Classi	FA	FR	FC	Densità
(10, 14]	10	0.33	0.33	0.0825
(14, 18]	13	0.43	0.76	0.1075
(18, 22]	4	0.14	0.90	0.035
(22, 26]	1	0.04	0.94	0.01
(26, 30]	1	0.03	0.97	0.0075
(30, 34]	1	0.03	1.00	0.0075

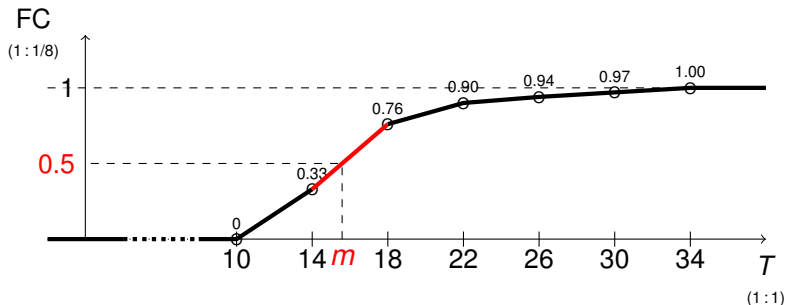


Dati continui

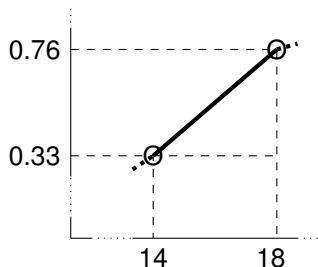


$$FC(x) \simeq ax + b \quad \text{per } x \in (14, 18]$$

$$\begin{cases} FC(14) = 0.33 \\ FC(18) = 0.76 \end{cases} \Rightarrow \begin{cases} 0.33 = 14a + b \\ 0.76 = 18a + b \end{cases}$$



Dati continui



$$FC(x) \simeq ax + b$$

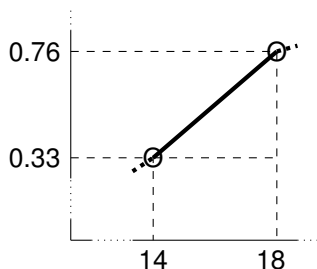
per $x \in (14, 18]$

$$\begin{cases} FC(14) = 0.33 \\ FC(18) = 0.76 \end{cases}$$

$$\Rightarrow \begin{cases} 0.33 = 14a + b \\ 0.76 = 18a + b \end{cases}$$

$$\Rightarrow \begin{cases} a = \frac{0.76 - 0.33}{18 - 14} \\ b = 0.33 - 14 \cdot \frac{0.76 - 0.33}{18 - 14} \end{cases}$$

Dati continui



$$FC(x) \simeq ax + b$$

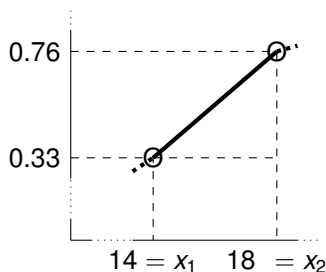
per $x \in (14, 18]$

$$\begin{cases} FC(14) = 0.33 \\ FC(18) = 0.76 \end{cases} \Rightarrow \begin{cases} 0.33 = 14a + b \\ 0.76 = 18a + b \end{cases}$$

$$\Rightarrow \begin{cases} a = \frac{0.76 - 0.33}{18 - 14} \\ b = 0.33 - 14 \cdot \frac{0.76 - 0.33}{18 - 14} \end{cases}$$

$$\Rightarrow FC(x) = \frac{0.76 - 0.33}{18 - 14} (x - 14) + 0.33$$

Dati continui

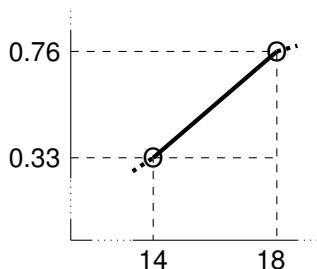


$$FC(x) \simeq ax + b \quad \text{per } x \in (14, 18]$$

$$\begin{cases} FC(14) = 0.33 \\ FC(18) = 0.76 \end{cases} \Rightarrow \begin{cases} 0.33 = 14a + b \\ 0.76 = 18a + b \end{cases}$$

$$\Rightarrow \begin{cases} a = \frac{0.76 - 0.33}{18 - 14} \\ b = 0.33 - 14 \cdot \frac{0.76 - 0.33}{18 - 14} \end{cases}$$

$$\begin{aligned} \Rightarrow FC(x) &= \frac{0.76 - 0.33}{18 - 14} (x - 14) + 0.33 \\ &= \frac{FC(x_2) - FC(x_1)}{x_2 - x_1} (x - x_1) + FC(x_1) \end{aligned}$$

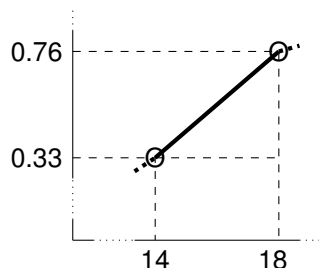


$$FC(x) = \frac{0.76 - 0.33}{18 - 14} (x - 14) + 0.33$$

Per calcolare la mediana:

$$FC(m) \equiv 0.50 \quad \Leftrightarrow \quad \frac{0.76 - 0.33}{18 - 14} (m - 14) + 0.33 = 0.50$$

$$\begin{aligned} \Leftrightarrow \quad m &= (0.50 - 0.33) \cdot \frac{18 - 14}{0.76 - 0.33} + 14 \\ &\simeq 15.5814 \end{aligned}$$



$$FC(x) = \frac{0.76 - 0.33}{18 - 14} (x - 14) + 0.33$$

Per calcolare la mediana:

$$FC(m) \equiv 0.50 \quad \Leftrightarrow \quad \frac{0.76 - 0.33}{18 - 14} (m - 14) + 0.33 = 0.50$$

$$\Leftrightarrow \quad m = (0.50 - 0.33) \cdot \frac{18 - 14}{0.76 - 0.33} + 14$$
$$\simeq 15.5814$$

Il valore esatto era $m = 14.9$