

# Corso di Statistica Psicometrica

Canale 1 - Matricole Dispari

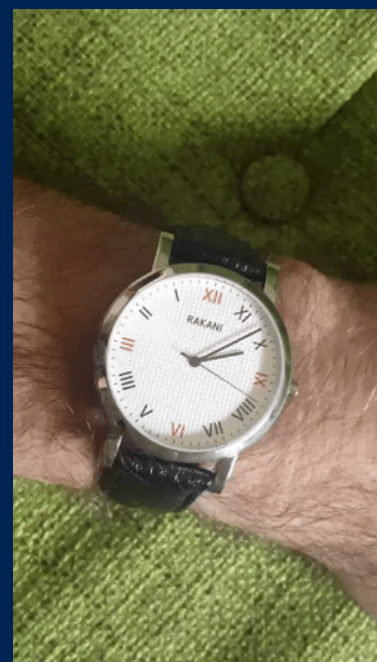
CdS triennale in  
Scienze e Tecniche  
Psicologiche

**Domenico Vistocco**

Dipartimento di Scienze Politiche  
Università degli Studi di Napoli Federico II



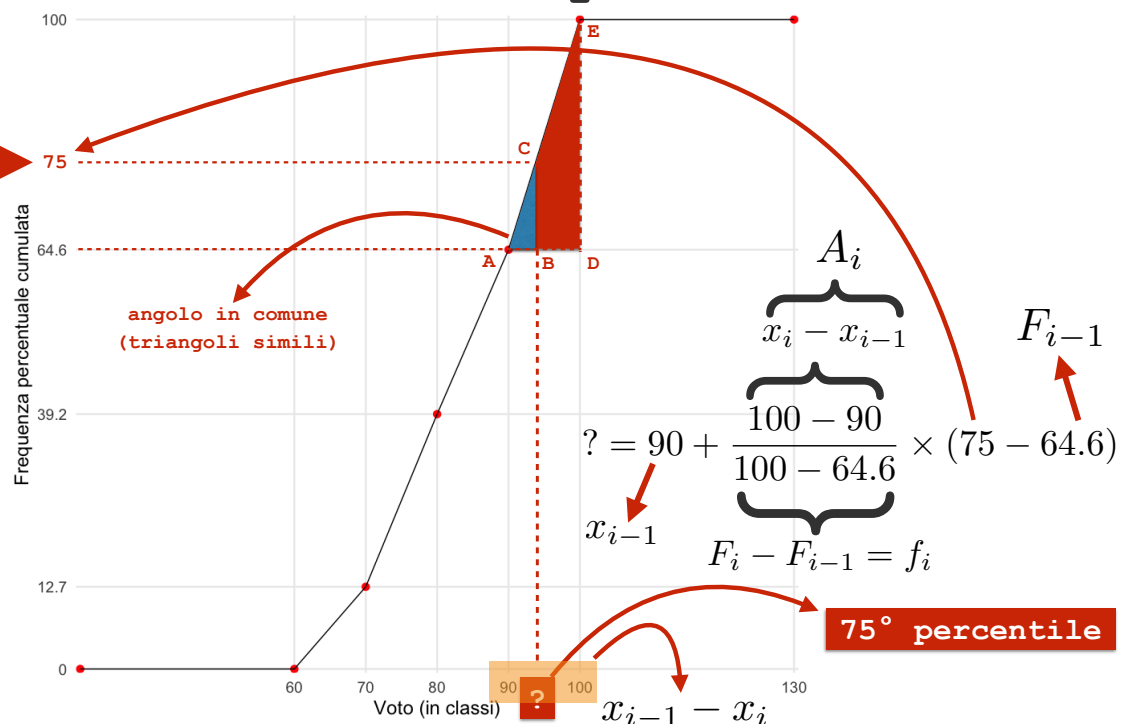
Nella  
puntata  
precedente...



- Tipologia di variabili
- Tipologia di tabelle
- Frequenza assolute, relative e percentuali
- Frequenze cumulate
- Densità di frequenza
- Istogramma e funzione di ripartizione empirica
- Indici assoluti ed indici relativi
- Classificazione degli indici univariati in base all'obiettivo di studio
- Quantile e rango quantile (più comunemente: percentile e rango percentile)

## Giochiamo un poco?

Qual è il voto superato solo dal 25% degli studenti?



## In sintesi...

$$q_{0.75} = x_{i-1} + \frac{x_i - x_{i-1}}{p_i} (75 - P_{i-1}) = q_3$$


$100 \times \frac{75}{100}$

$$q_{0.75} = x_{i-1} + \frac{x_i - x_{i-1}}{n_i} \left( n \frac{75}{100} - N_{i-1} \right) = q_3$$

$$q_{0.75} = x_{i-1} + \frac{x_i - x_{i-1}}{f_i} (0.75 - F_{i-1}) = q_3$$


$1 \times \frac{75}{100} = \frac{3}{4}$

## In generale...

$$\tau \in [0, 1]$$

$$q_\tau = x_{i-1} + \frac{x_i - x_{i-1}}{f_i} (\tau - F_{i-1})$$

$$q_\tau = x_{i-1} + \frac{x_i - x_{i-1}}{p_i} (\tau \times 100 - P_{i-1})$$

$$q_\tau = x_{i-1} + \frac{x_i - x_{i-1}}{n_i} (\tau \times n - N_{i-1})$$

## Alternativa?



$i \in [0, 1, 2, \dots, k-1, k]$

minimo  massimo

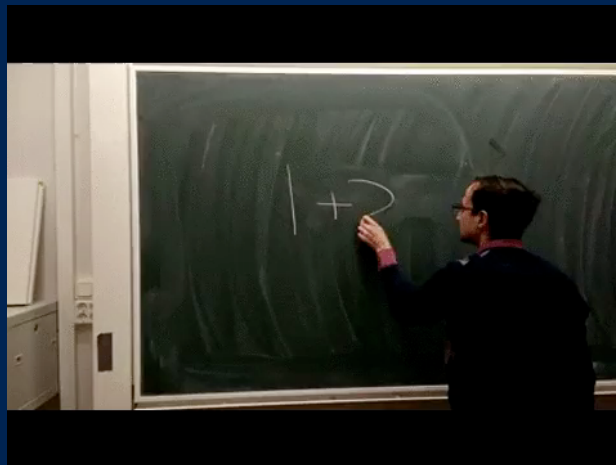
$k$  è il numero di parti in cui vogliamo dividere i dati

$i$  è la parte che ci interessa

$k = 4, i = 3$  **Terzo quartile**

$k = 10, i = 9$  **Nono decile**

$k = 100, i = 20$  **Ventesimo percentile**



# Si riparte?

## Tecnicismi...

Punteggio su una scala di ansia rilevata in 10 studenti dopo questa lezione

Serie grezza

41, 24, 72, 12, 18, 50, 90, 67, 34, 93

Posizione Me

$$\frac{n}{2} = \frac{10}{2} = 5$$

Me

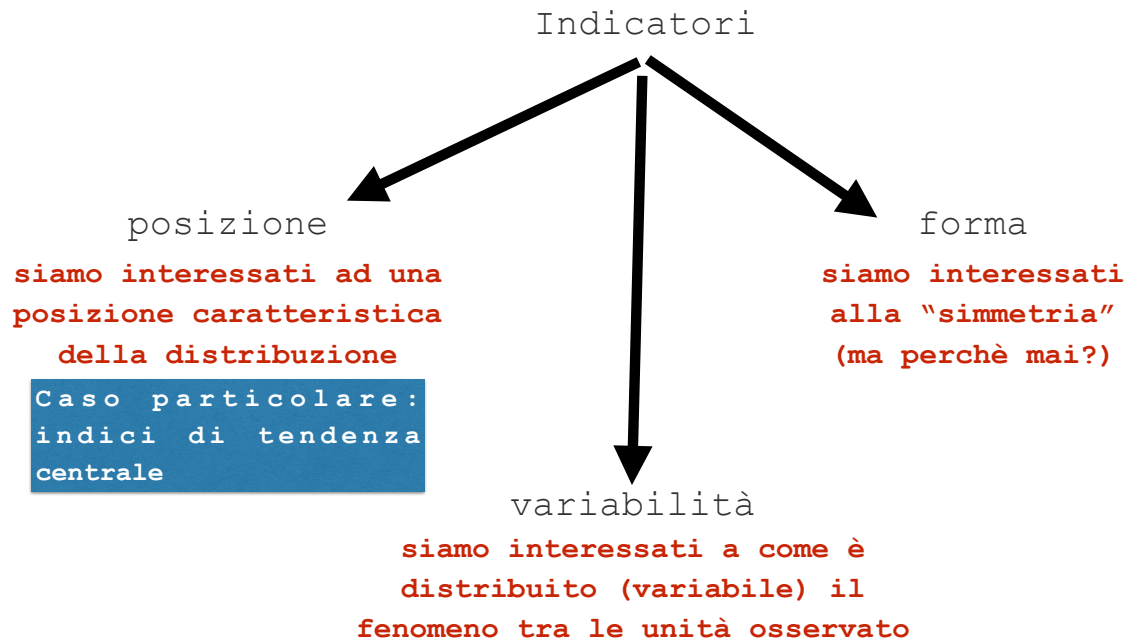
$$x\left(\frac{n}{2}\right) = x_{(5)}$$

Serie ordinata

12, 18, 24, 34, 41, 50, 67, 72, 90, 93

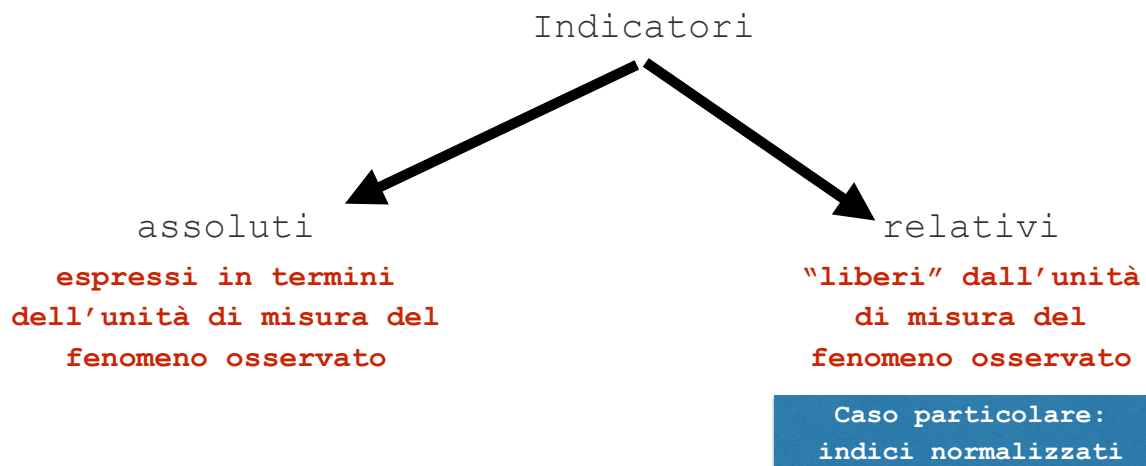
E tu che cosa  
guardi?

# Tipologia di indici univariati

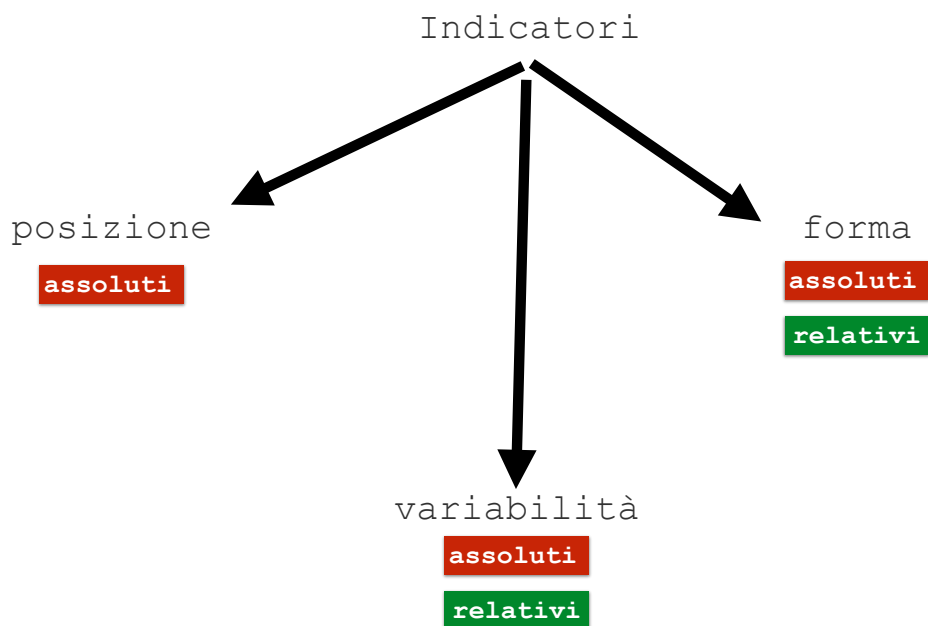


E con che occhi  
lo guardi?

# Tipologia di indici univariati



# Tipologia di indici univariati



## Tecnicismi...



12, 18, 24, 34, 41, 50, 67, 72, 90, 93

← 4 a sinistra      6 a destra →

12, 18, 24, 34, 41, 50, 67, 72, 90, 93

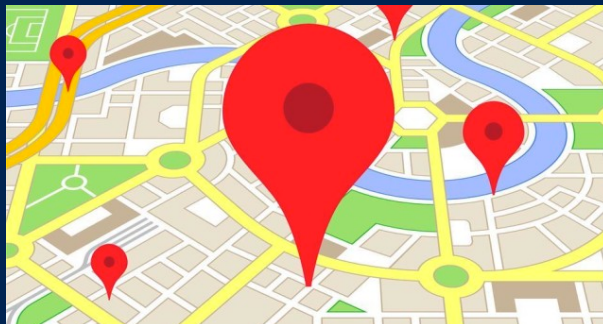
← 6 a sinistra      4 a destra →

12, 18, 24, 34, 41, 50, 67, 72, 90, 93

← 5 a sinistra      5 a destra →

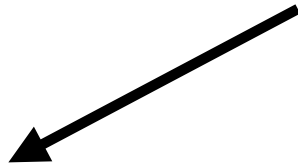
$x_{\left(\frac{n+1}{2}\right)} = x_{(5.5)} = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}$

# Qual è la tua posizione?



## Moda

Indicatori



in generale

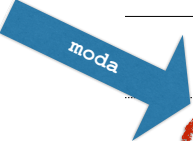
$$Mo = \operatorname{argmax} \{n_i\}$$



$$Mo = \operatorname{argmax} \{n_i\} = \{x_i \mid n_i = \max(n_i)\}$$

## Tecnicismi...

| VOTO | Frequenza assoluta |
|------|--------------------|
| 18   | 5                  |
| 23   | 21                 |
| 25   | 10                 |
| 27   | 14                 |
| 28   | 18                 |
| 30   | 11                 |
|      | 79                 |




 $\max(n_i)$

$\operatorname{argmax}(n_i)$

Ufficio degli affari complicati?

## Moda

Indicatori

in generale

caso particolare  
tabelle per classi di  
diversa ampiezza

$$Mo = \operatorname{argmax} \{d_i\}$$



$$Mo = \operatorname{argmax} \{d_i\} = \{x_i \mid d_i = \max(d_i)\}$$

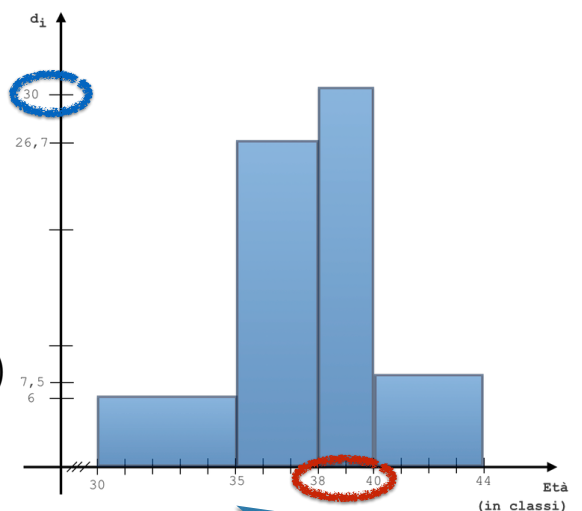
## Mediana

| ETA'     | Frequenza assoluta | Ampiezza | Densità |
|----------|--------------------|----------|---------|
| 30 I- 35 | 30                 | 5        | 6       |
| 35 I- 38 | 80                 | 3        | 26,7    |
| 38 I- 40 | 60                 | 2        | 30      |
| 40 I- 44 | 30                 | 4        | 7,5     |
|          | 200                |          |         |

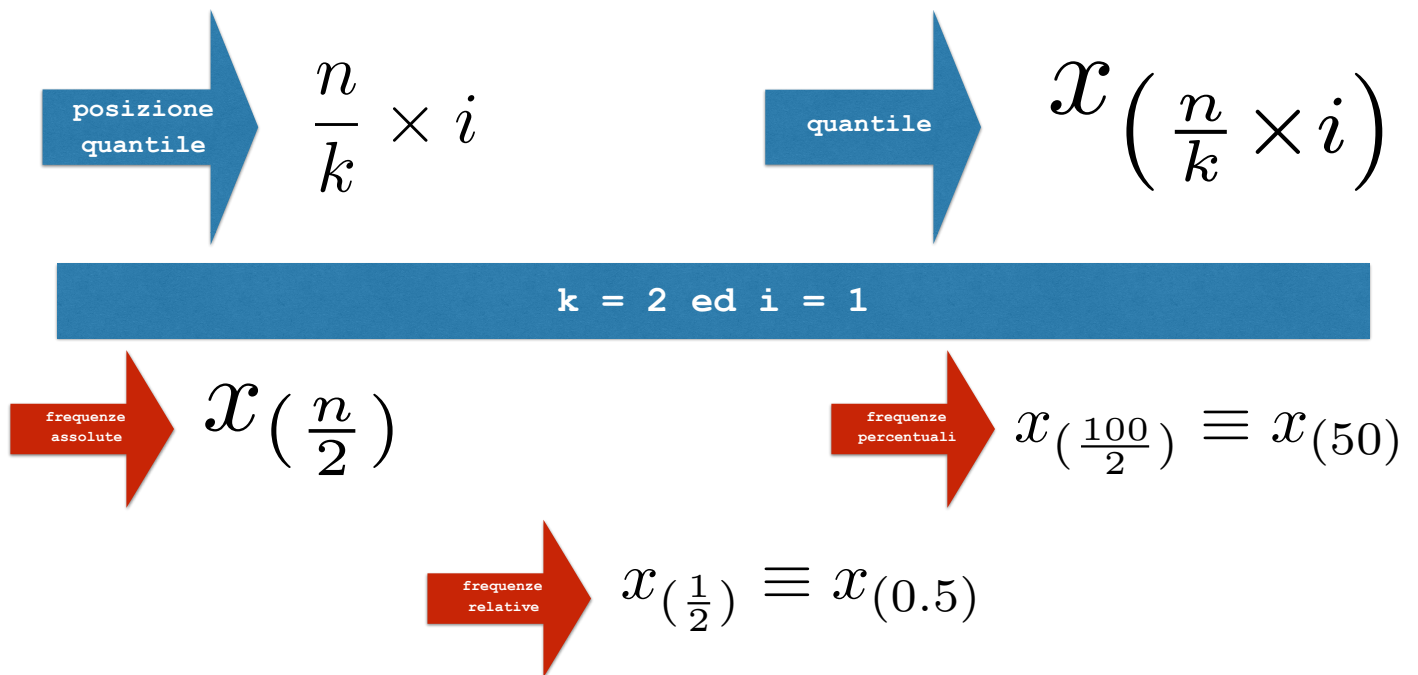
classe modale

$\operatorname{argmax}(d_i)$

$\max(d_i)$



## Mediana



## Tecnicismi...

Distribuzioni bimodali

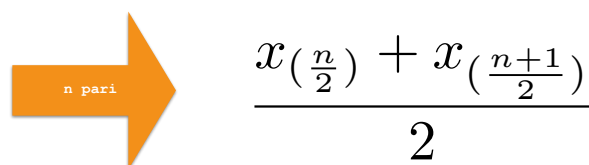
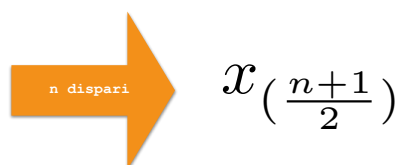
Distribuzioni plurimodali

Assenza di moda

## Mediana (complicazioni?)



$k = 2$  ed  $i = 1$



## Tecnicismi...

$$Me = x\left(\frac{n+1}{2}\right) = x_{(40)}$$

| VOTO | Frequenza assoluta | Frequenza cumulata           |
|------|--------------------|------------------------------|
| 18   | 5                  | 5 (dal 1° al 5° studente)    |
| 23   | 21                 | 26 (dal 6° al 26° studente)  |
| 25   | 10                 | 36 (dal 27° al 36° studente) |
| 27   | 14                 | 50 (dal 37° al 50° studente) |
| 28   | 18                 | 68 (dal 51° al 68° studente) |
| 30   | 11                 | 79 (dal 69° al 79° studente) |
|      | 79                 |                              |

## Tecnicismi...

$$Me = x_{\left(\frac{n+1}{2}\right)} = x_{(40)}$$



| VOTO      | Frequenza assoluta | Frequenza cumulata |
|-----------|--------------------|--------------------|
| 60 I- 70  | 10                 | 10                 |
| 70 I- 80  | 21                 | 31                 |
| 80 I- 90  | 20                 | 51                 |
| 90 I- 100 | 28                 | 79                 |
|           | 79                 |                    |

NOTA:

va quindi applicato il procedimento di interpolazione per trovare un'approssimazione della mediana

## La media aritmetica (questa sì che è conosciuta)

Serie grezza:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

Tabella di frequenza  
per singole modalità:

$$\bar{x} = \frac{\sum_{i=1}^n x_i n_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i n_i$$

Tabella di frequenza  
per classi di modalità:

$$\bar{x} = \frac{\sum_{i=1}^n c_i n_i}{n} = \frac{1}{n} \sum_{i=1}^n c_i n_i$$

$$c_i = \hat{x}_i = \frac{x_{i-1} + x_i}{2}$$

# La media aritmetica (esercizio)

## Serie grezza:

Calcolare la media aritmetica sulla serie grezza dei 10 punteggi di ansia

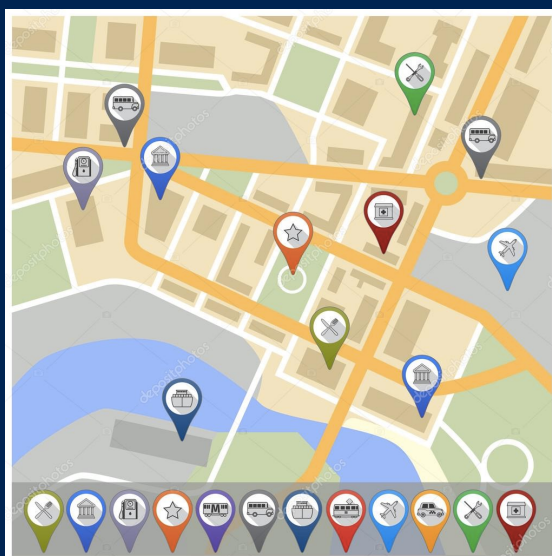
## Tabella di frequenza per singole modalità:

Calcolare la media aritmetica sulla tabella di frequenza per singole modalità relativi ai voti all'esame

## Tabella di frequenza per classi di modalità:

Calcolare la media aritmetica sulla tabella di frequenza per classi di modalità relativi ai voti conseguiti al diploma

# Altre posizioni



# Altri centri

Intervallo medio

$$\frac{x_{(1)} + x_{(n)}}{2} \equiv \frac{\min(X) + \max(X)}{2}$$

Intervallo medio interquartile

$$\frac{q_{(0.25)} + q_{(0.75)}}{2} \equiv \frac{q_1 + q_3}{2}$$

Intervallo medio interquartile

$$\frac{q_{(\tau)} + q_{(1-\tau)}}{2}$$

Medie troncate

## Altre posizioni (non "centrali")

Minimo

$$x_{(1)} \equiv \min(X) \equiv \min$$

Massimo

$$x_{(n)} \equiv \max(X) \equiv \max$$

Quantili

$$q_{\tau}$$

$$\tau \in [0, 1]$$

# Proprietà

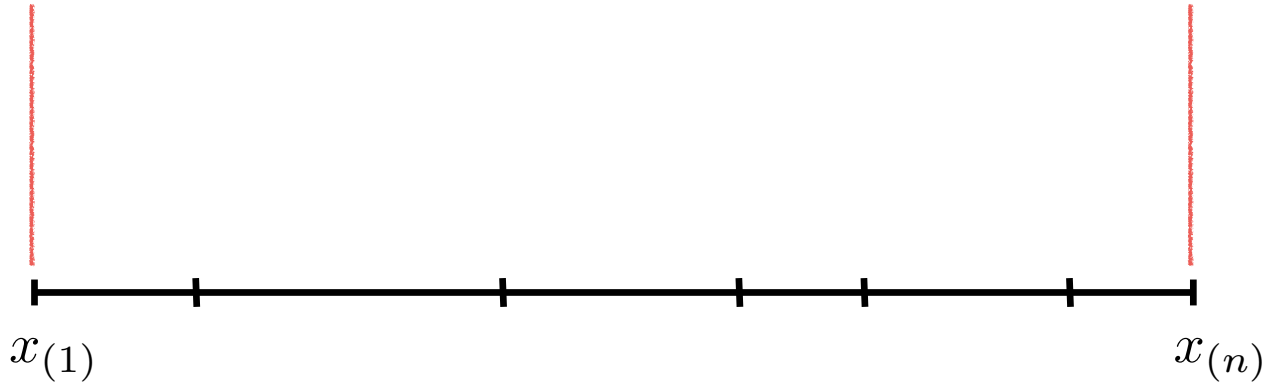


# Tutti gli indici di posizione

Internalità



$$x_{(1)} \leq \text{indicatore} \leq x_{(n)}$$



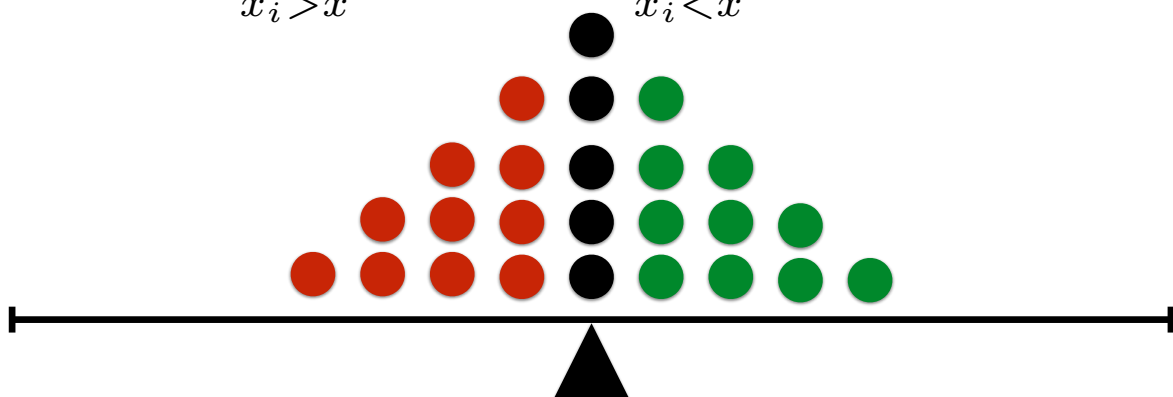
# Media aritmetica

Baricentricità



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

$$\sum_{x_i > \bar{x}} (x_i - \bar{x}) = \sum_{x_i < \bar{x}} (x_i - \bar{x})$$



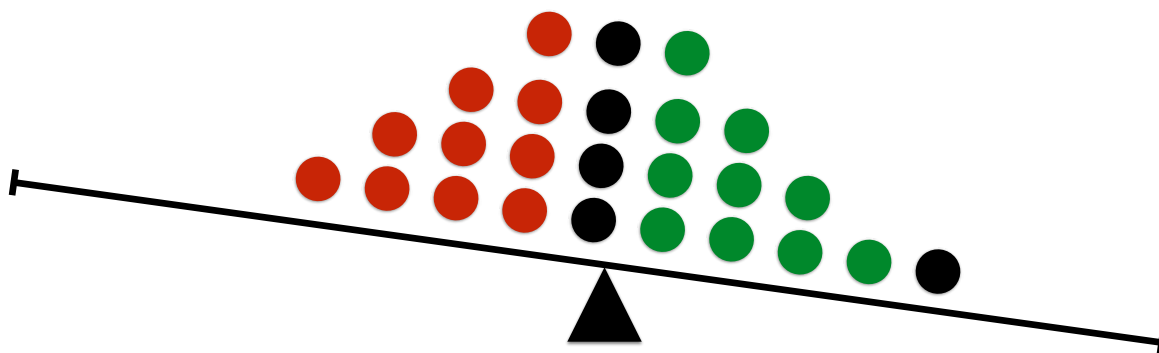
## Media aritmetica

Baricentricità



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

$$\sum_{x_i > \bar{x}} (x_i - \bar{x}) = \sum_{x_i < \bar{x}} (x_i - \bar{x})$$



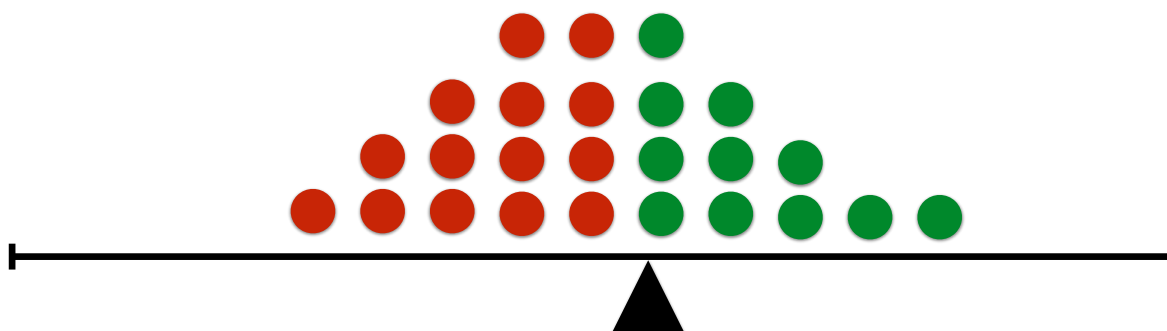
## Media aritmetica

Baricentricità



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

$$\sum_{x_i > \bar{x}} (x_i - \bar{x}) = \sum_{x_i < \bar{x}} (x_i - \bar{x})$$



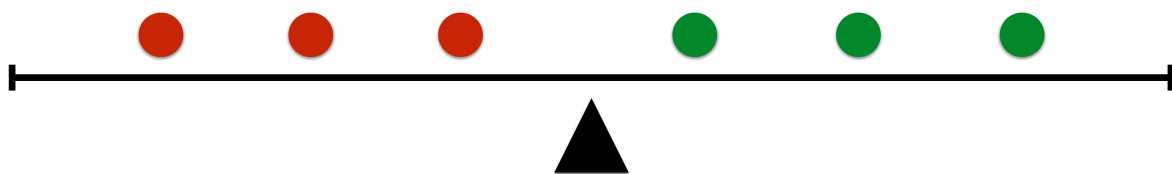
## Media aritmetica

Baricentricità



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

$$\sum_{x_i > \bar{x}} (x_i - \bar{x}) = \sum_{x_i < \bar{x}} (x_i - \bar{x})$$



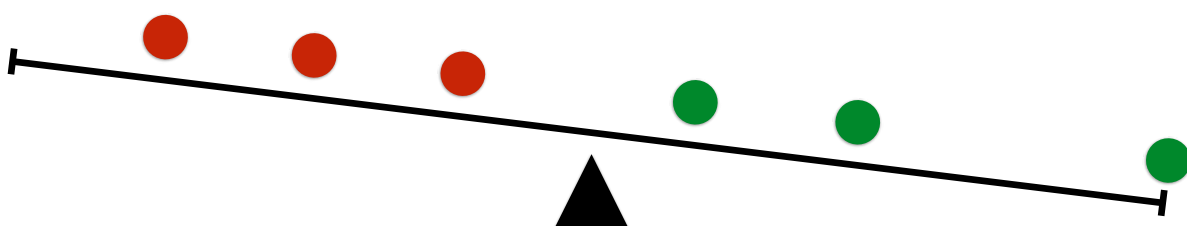
## Media aritmetica

Baricentricità



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

$$\sum_{x_i > \bar{x}} (x_i - \bar{x}) = \sum_{x_i < \bar{x}} (x_i - \bar{x})$$



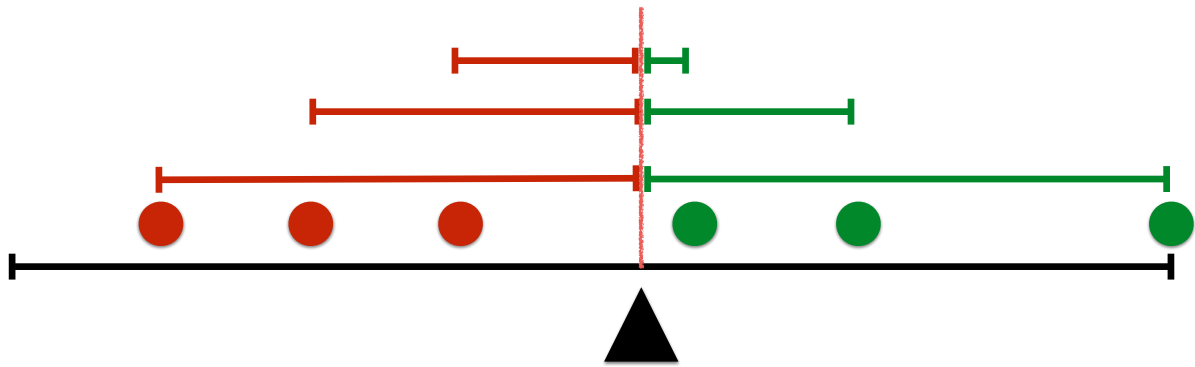
## Media aritmetica

Baricentricità



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

$$\sum_{x_i > \bar{x}} (x_i - \bar{x}) = \sum_{x_i < \bar{x}} (x_i - \bar{x})$$



## Media aritmetica

Baricentricità



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

**Dimostrazione**

# Media aritmetica

Criterio di ottimo   $\bar{x} = \arg \min_c \sum_{i=1}^n (x_i - c)^2$

**Dimostrazione**

## Dati divisi in gruppi: struttura

| Gruppo   | Unità                                                | Cardinalità |
|----------|------------------------------------------------------|-------------|
| 1        | $x_{1,1}, x_{1,2}, \dots, x_{1,j}, \dots, x_{1,n_1}$ | $n_1$       |
| 2        | $x_{2,1}, x_{2,2}, \dots, x_{2,j}, \dots, x_{2,n_2}$ | $n_2$       |
| $\vdots$ | $\vdots$                                             | $\vdots$    |
| i        | $x_{i,1}, x_{i,2}, \dots, x_{i,j}, \dots, x_{i,n_i}$ | $n_i$       |
| $\vdots$ | $\vdots$                                             | $\vdots$    |
| G        | $x_{G,1}, x_{G,2}, \dots, x_{G,j}, \dots, x_{G,n_G}$ | $n_G$       |

$$n = \sum_{i=1}^G n_i$$

# La media aritmetica (questa sì che è conosciuta)

**MEMO**  
(lezione 3)

Voti all'ultima  
sessione di esame:

28

25

26

30

27

Maria

Giovanni

Simona

Laura

Giuseppe

$$\bar{x} = 27.2$$

$$\bar{x}_F = 28$$

$$\bar{x}_M = 26$$

Qui sfruttiamo:

$$\sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i$$

## Dati divisi in gruppi: sintesi

| Gruppo                                                      | Unità                                       | Cardinalità            | Medie                                              |
|-------------------------------------------------------------|---------------------------------------------|------------------------|----------------------------------------------------|
| 1                                                           | $x_{1,1}, \dots, x_{1,j}, \dots, x_{1,n_1}$ | $n_1$                  | $\bar{x}_1$                                        |
| 2                                                           | $x_{2,1}, \dots, x_{2,j}, \dots, x_{2,n_2}$ | $n_2$                  | $\bar{x}_2$                                        |
| $\vdots$                                                    | $\vdots$                                    | $\vdots$               |                                                    |
| i                                                           | $x_{i,1}, \dots, x_{i,j}, \dots, x_{i,n_i}$ | $n_i$                  | $\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{i,j}}{n_i}$ |
| $\vdots$                                                    | $\vdots$                                    | $\vdots$               |                                                    |
| G                                                           | $x_{G,1}, \dots, x_{G,j}, \dots, x_{G,n_G}$ | $n_G$                  | $\bar{x}_G$                                        |
| $\bar{x} = \frac{\sum_{i=1}^G \sum_{j=1}^{n_i} x_{i,j}}{n}$ |                                             | $n = \sum_{i=1}^G n_i$ |                                                    |

# Proprietà di associatività della media

La media complessiva:

$$\bar{x} = \frac{\sum_{i=1}^G \sum_{j=1}^{n_i} x_{i,j}}{n}$$

Può essere ottenuta come media ponderata delle medie di gruppo con pesi dati dalle numerosità di gruppo:

$$\bar{x} = \frac{\sum_{i=1}^G \bar{x}_i n_i}{n}$$

## Media aritmetica

Linearità (comportamento rispetto a trasformazioni lineari)

$$\begin{matrix} X \\ (x_1, \dots, x_i, \dots, x_n) \end{matrix} \quad \longrightarrow \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

# Media aritmetica

Linearità (comportamento rispetto a trasformazioni lineari)

$$X \longrightarrow \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$(x_1, \dots, x_i, \dots, x_n)$



$$Y = a + bX$$

$(y_1, \dots, y_i, \dots, y_n)$

$a + bx_1$        $a + bx_i$        $a + bx_n$

## La media aritmetica (questa sì che è conosciuta)

MEMO

(lezione 3)

Voti al diploma:  
(scala 36 - 60)

$$\bar{x} = 51.8$$

Voti equivalenti:  
(scala 60 - 100)

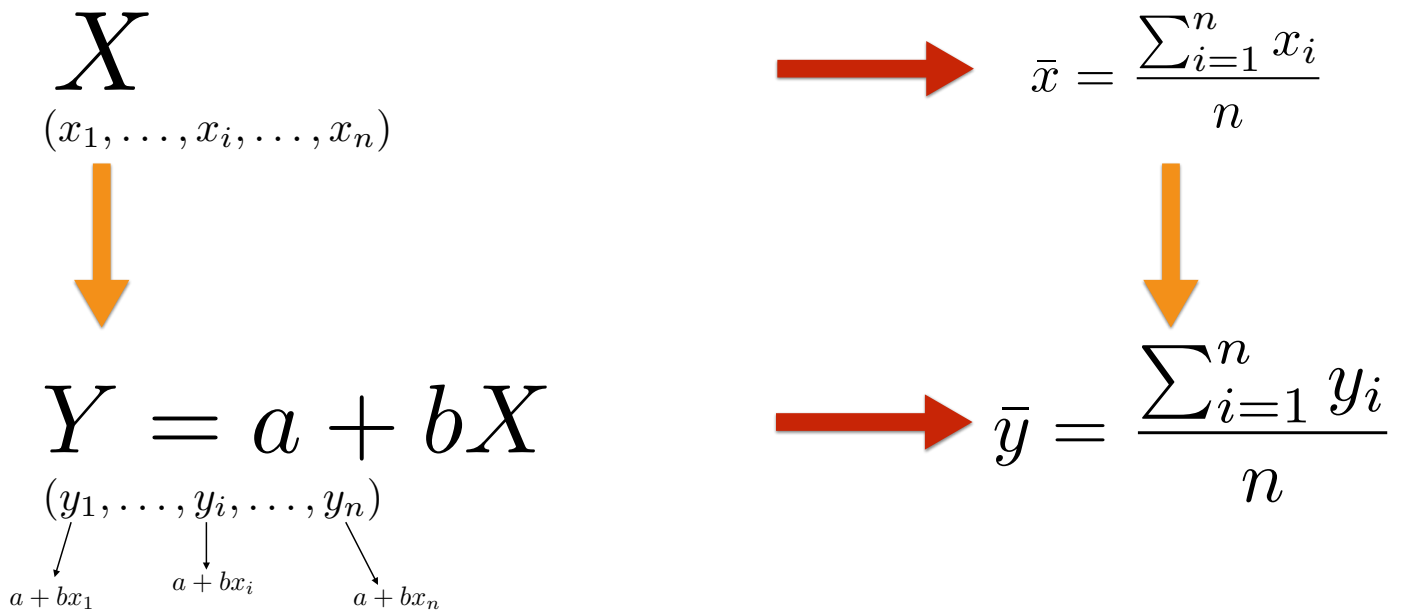
$$\bar{x} = 86.3$$

|                          |               |        |       |               |
|--------------------------|---------------|--------|-------|---------------|
| 50                       | 55            | 36     | 60    | 58            |
| Maria                    | Giovanni      | Simona | Laura | Giuseppe      |
| $\frac{50}{6} \times 10$ |               |        |       |               |
| 83. $\bar{3}$            | 91. $\bar{7}$ | 60     | 100   | 96. $\bar{7}$ |
| Maria                    | Giovanni      | Simona | Laura | Giuseppe      |

$$\sum_{i=1}^n k x_i = k \sum_{i=1}^n x_i \longrightarrow \frac{51.8}{6} \times 10 = 86.3$$

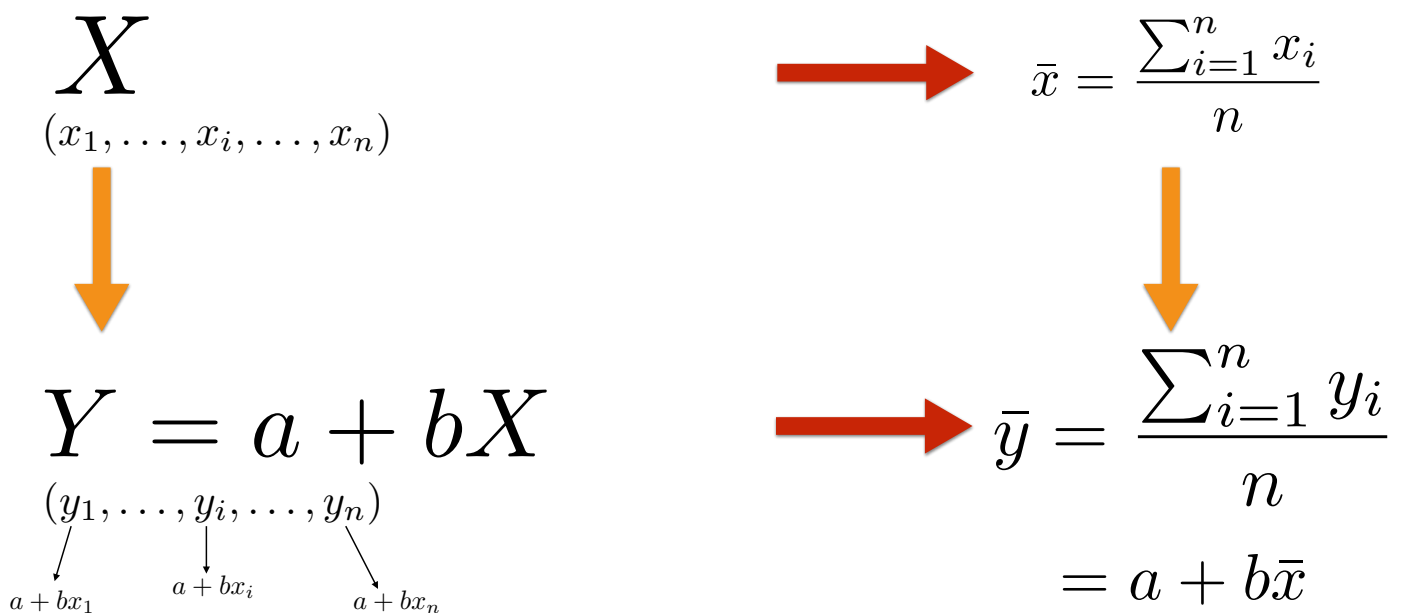
## Media aritmetica

Linearità (comportamento rispetto a trasformazioni lineari)



## Media aritmetica

Linearità (comportamento rispetto a trasformazioni lineari)



# Media aritmetica

Linearità (comportamento rispetto a trasformazioni lineari)

$$\begin{cases} a = -\bar{x} \\ b = 1 \end{cases}$$



$$Y = a + bX = -\bar{x} + X = X - \bar{x}$$



$$(y_1 = a + bx_1, \dots, y_i = a + bx_i, \dots, y_n = a + bx_n)$$



$$(y_1 = x_1 - \bar{x}, \dots, y_i = x_i - \bar{x}, \dots, y_n = x_n - \bar{x})$$



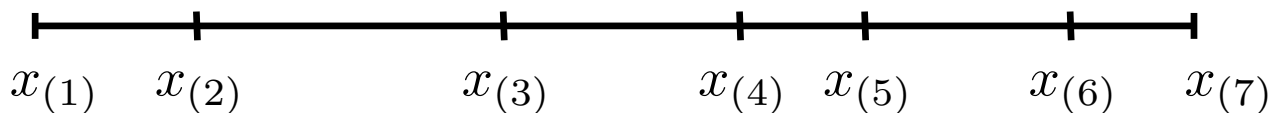
$$\bar{y} = ?$$

# Mediana

Criterio di ottimo



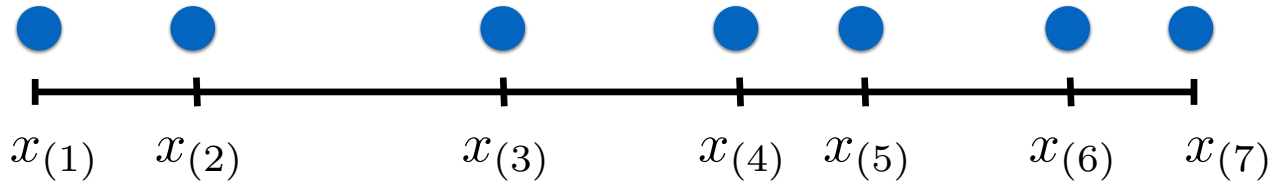
$$Me = \arg \min_c \sum_{i=1}^n |x_i - c|$$



To be continued..

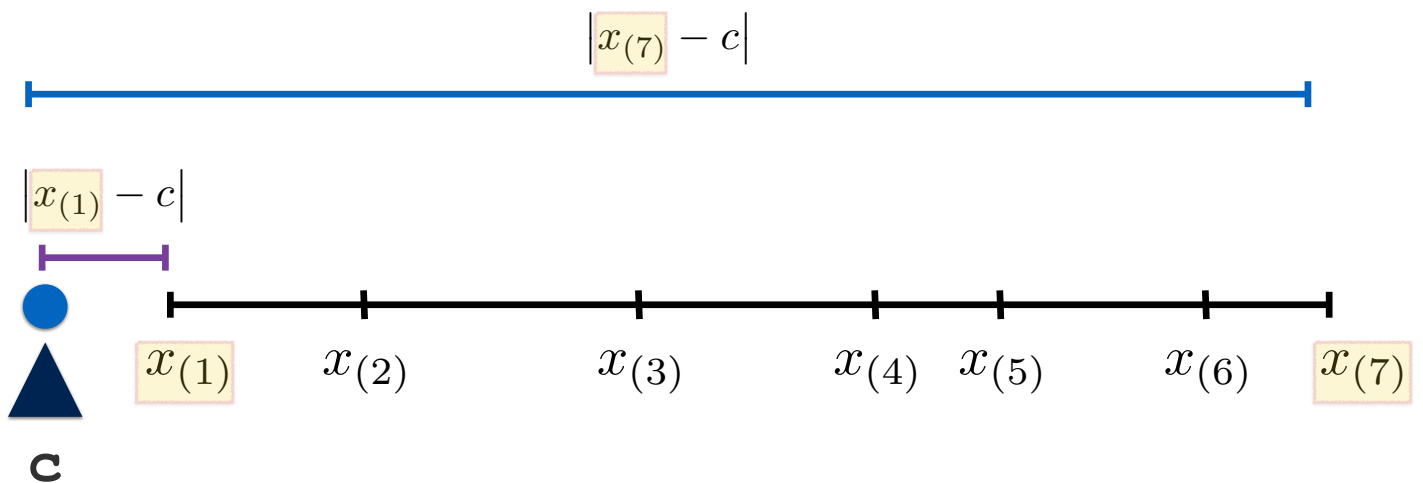
## Mediana

Criterio di ottimo   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



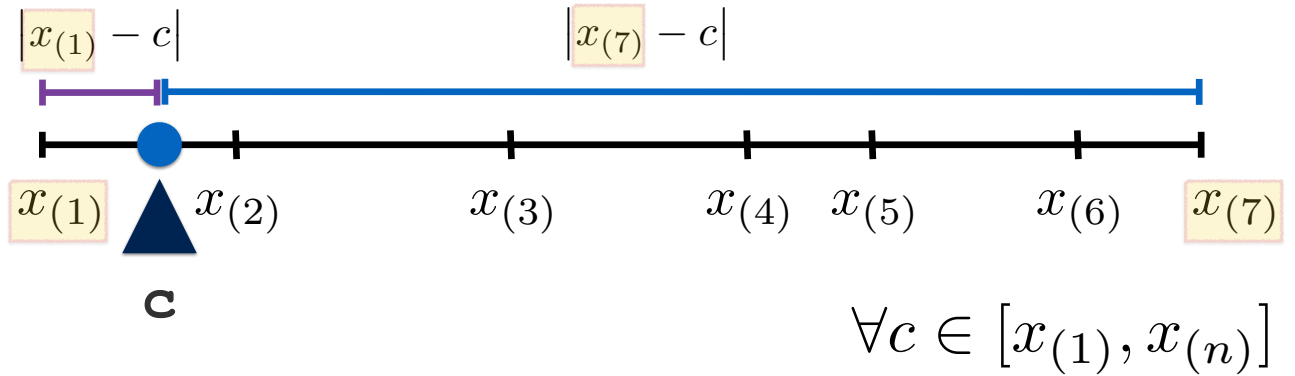
## Mediana

Criterio di ottimo   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



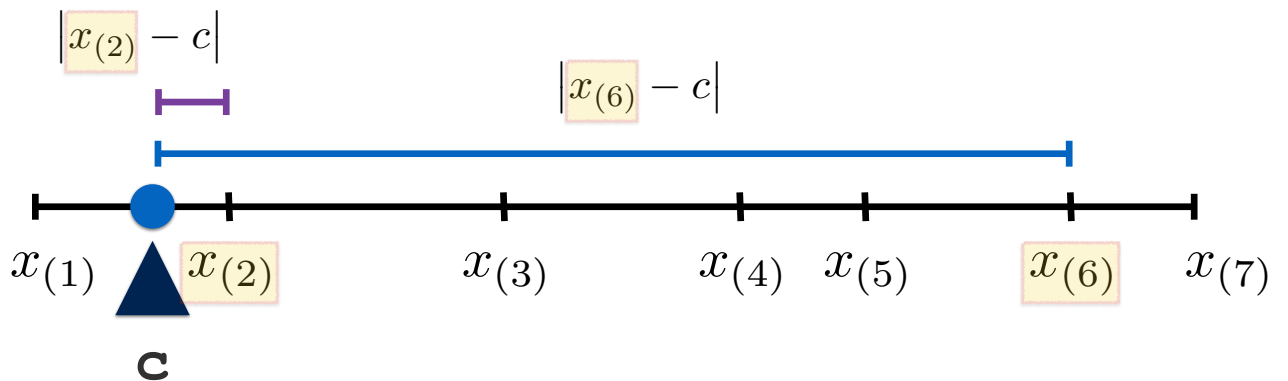
## Mediana

Criterio di ottimo  $\longrightarrow$   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



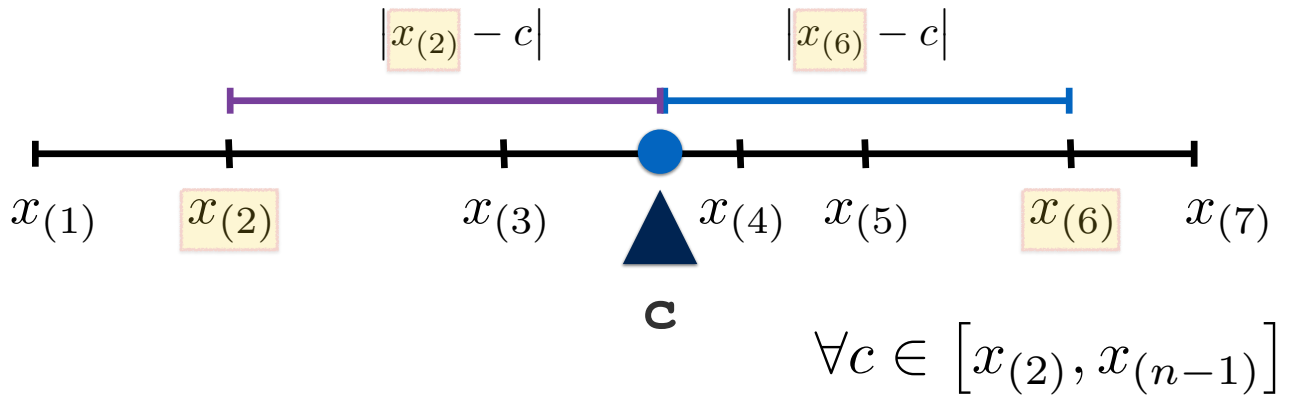
## Mediana

Criterio di ottimo  $\longrightarrow$   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



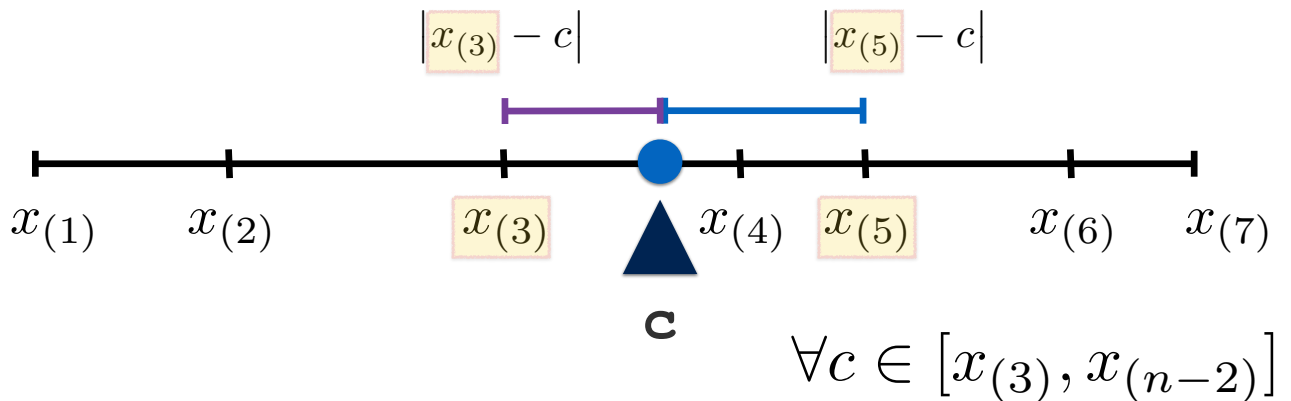
## Mediana

Criterio di ottimo   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



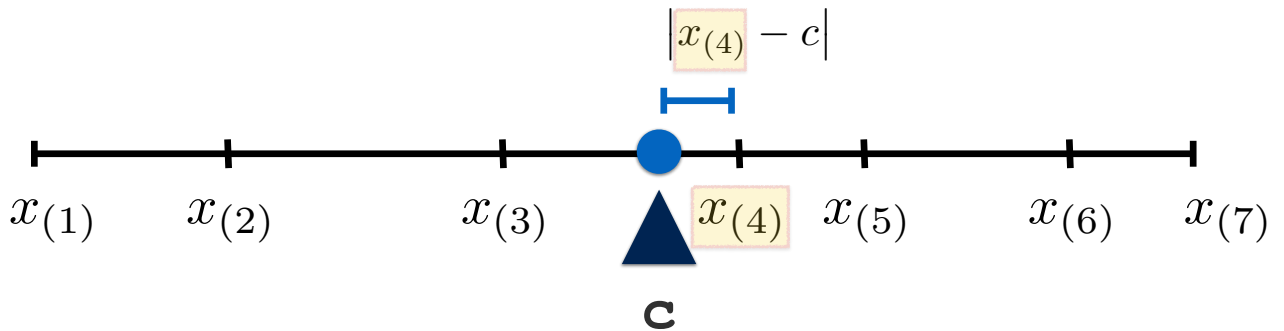
## Mediana

Criterio di ottimo   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



## Mediana

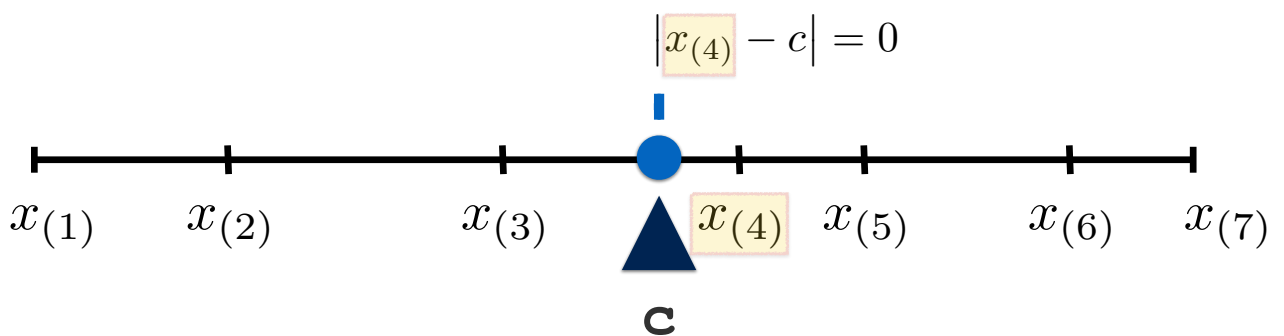
Criterio di ottimo   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$



## Mediana

Criterio di ottimo   $Me = \arg \min_c \sum_{i=1}^n |x_i - c|$

$Me = x_{(\frac{n+1}{2})}$  



# Mediana

Criterio di ottimo



$$Me = \arg \min_c \sum_{i=1}^n |x_i - c|$$



$$\forall c \in \left[ x_{(\frac{n}{2})}, x_{(\frac{n+1}{2})} \right]$$

