

## Misure di posizione

### INDICI DI POSIZIONE


*(measures of location or central tendency)*

1. MODA
2. MEDIA
3. MEDIANA

## MODA

*E' la scelta fatta dalla maggioranza della popolazione, lo stile che "tutti" seguono*

**in statistica non è diverso**





Si definisce moda di un insieme di dati o di una distribuzione di frequenza la modalità, il valore (o l'intervallo di classe) della variabile a cui corrisponde la **massima frequenza**.

esempio: (50 neonati)


modalità $x_i$	frequenza assoluta $n_i$	frequenza relativa $p_i$	frequenza relativa percentuale $p_i$ (%)
normale	35	0.70	70%
forcipe	1	0.02	2%
cesareo	14	0.28	28%
TOTALE	50	1.00	100%

MODA o classe modale


## ESERCIZIO

I dati seguenti si riferiscono all'abitudine al fumo in un campione di 168 soggetti senza bronchite cronica di età 20-44 anni:



	frequenza assoluta
-----	
non fumatore	74
ex fumatore	37
moderato fumatore	34
forte fumatore	23
-----	
Totale	168

Determinate la moda della distribuzione.



## SOLUZIONE

	frequenza assoluta	assoluta cumulata
non fumatore	74	74
ex fumatore	37	111
moderato fumatore	34	145
forte fumatore	23	168
Totale	168	

CALCOLO DELLA MODA

massima frequenza assoluta



MODA:  
non fumatore



## MEDIANA

- Il valore centrale di una serie **ORDINATA** di dati
- Le osservazioni vengono separate dal valore mediano in due parti numericamente uguali

es. sulla mediana

campione di 5 unità  
variabile d'interesse = altezza



165 cm



150 cm



180 cm



50 cm



155 cm

1. ordino le unità secondo un ordine crescente di altezza

es. sulla mediana

campione di 5 unità  
variabile d'interesse = altezza



50 cm



150 cm



155 cm

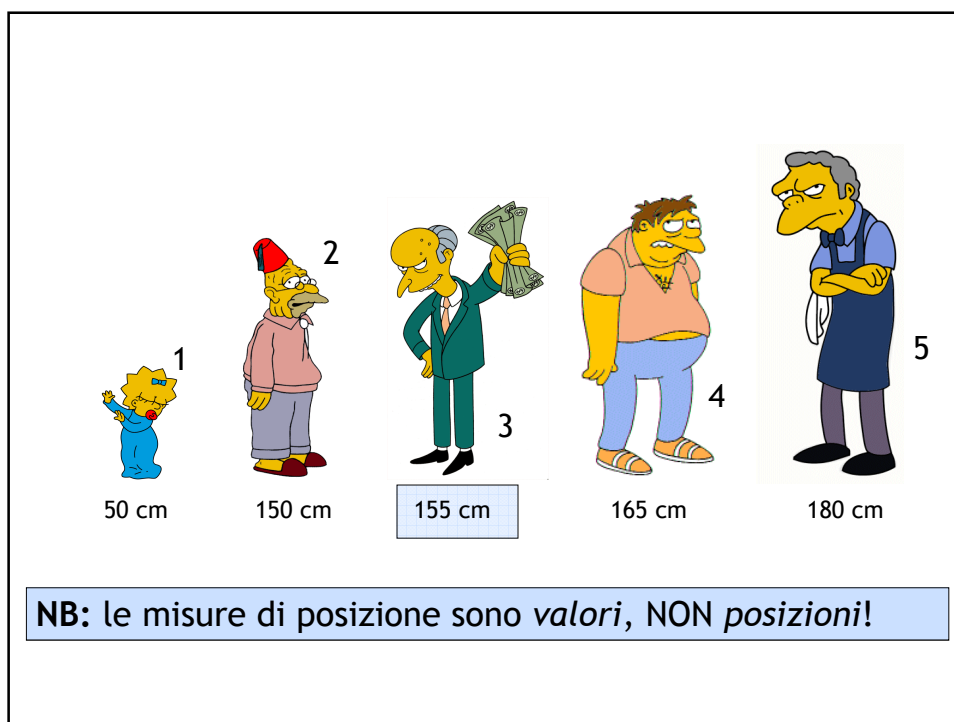
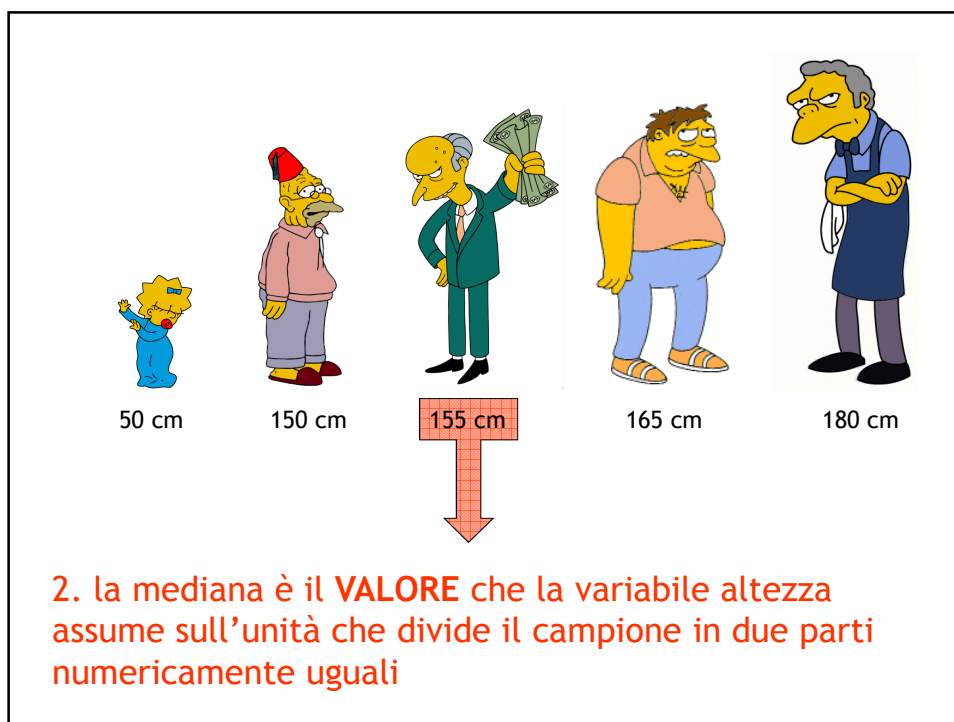


165 cm



180 cm

2. identifico l'unità centrale nella serie ordinata di dati



### Frequenza cardiaca in 9 pazienti con arresto respiratorio

167	150	125	120	150
150	40	136	120	

Ordino i dati:

				136				
40	120	120	125	136	150	150	150	167
1	2	3	4	5	6	7	8	9

→ mediana  
 → posizione

• n° dispari di osservazioni

Mediana  $\Rightarrow$  valore nella posizione

$$\frac{N + 1}{2} \Rightarrow \frac{9 + 1}{2}$$

### Volume plasmatico di 8 maschi adulti sani (dati ordinati)

2,62; 2,75; 2,76; 2,86;  
3,05; 3,12; 3,37; 3,49.

Mediana  $\Rightarrow$   $\frac{2,86 + 3,05}{2} = 2,96$

• n° pari di osservazioni

MEDIANA  $\Rightarrow$  media tra i valori nelle posizioni

$$\frac{N}{2} \text{ e } \frac{N + 1}{2} \Rightarrow \frac{8}{2} \text{ e } \frac{8 + 1}{2} \rightarrow 5$$

↓  
4

## MEDIA ARITMETICA

La media aritmetica di un insieme di osservazioni è pari alla somma dei valori diviso il numero totale delle osservazioni

*Formalmente:* siano  $(x_1, x_2, \dots, x_n)$  le osservazioni della variabile  $X$  su un campione di  $n$  unità statistiche, allora

$$\bar{x} = \sum_{i=1}^n x_i / n = (x_1 + x_2 + \dots + x_n) / n$$

**esempio:**  
(8 osservazioni)

	5	16	13	27	11	5	13	13
	↑	↑	↑	↑	↑	↑	↑	↑
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$

$$\bar{x} = (5+16+13+27+11+5+13+13)/8 = 103/8 = 12.9$$



### Volume plasmatico di 8 maschi adulti sani:

2,75;      2,86;      3,37;      2,76;      2,62;  
3,43;      3,05;      3,12 litri

### Media

$$\frac{2,75 + 2,86 + 3,37 + 2,76 + \dots}{8} = 3,001$$

$$\text{Media} = \bar{x} = \frac{\sum x_i}{n}$$

$$\sum x_i = x_1 + x_2 + x_3 + x_4 + \dots + x_n$$

Misure di posizione

Esempio:

La durata dell'apnea (in sec) di 9 soggetti è:  
30,15,23,46,64,126,180,623,780

La media è 209.6 sec

La mediana 64 sec

NB

○I valori più grandi influenzano fortemente la media!

○La mediana mostra meglio le caratteristiche della distribuzione

## Esercizio

- Un gruppo di pazienti presentano i seguenti livelli di glucosio nel sangue: 62, 78, 79, 80, 82, 82, 83, 85, 87, 91, 96, 97, 97, 97, 101, 120, 135, 180, 270, 400.

Calcolate media, moda e mediana per questo gruppo di dati.

Media= 120.1

Moda= 97

Mediana= 93.5



**Distribuzione del n° di precedenti gravidanze di un gruppo di donne tra i 30 e i 34 anni in una clinica prenatale.**

N° di precedenti gravidanze						
$x_i$	0	1	2	3	4	totale
$f_i$	18	27	31	19	5	100

0, 0, 0, 0, 0, 0 .... 1, 1, 1, .... 2, 2, 2, ....

$$\sum x_i \cdot f_i = 0 \cdot 18 + 1 \cdot 27 + 2 \cdot 31 + 3 \cdot 19 + 4 \cdot 5 = 166$$

$$n = \sum f_i = 100$$

$$\bar{x} = \frac{166}{100} = 1,66 = \frac{\sum x_i \cdot f_i}{\sum f_i}$$

**Moda = 2**

## MEDIA ARITMETICA PONDERATA - II

$$\bar{x} = \frac{\bar{x}_1 n_1 + \bar{x}_2 n_2 + \dots + \bar{x}_k n_k}{n_1 + n_2 + \dots + n_k}$$



k = numero di classi

$\bar{x}_i$  = media aritmetica nella classe i-esima

$n_i$  = numerosità della classe i-esima

$\bar{x}$  = media aritmetica complessiva

**esempio:** *valore medio dell'altezza nei maschi e nelle femmine matricole della Facoltà di Medicina (A.A. 95/96)*

Sesso	$n_i$	$\bar{x}_i$
<b>maschi</b>	34	177
<b>femmine</b>	91	166.1
<b>Totale</b>	125	

$$\bar{x} = \frac{177 \cdot 34 + 166.1 \cdot 91}{125} = 169.1 \text{ cm}$$

### Esempio-Calcolo della media pesata

- Nella I settimana di aprile sono nati **10** bimbi, peso medio **3,42**; nella seconda sett. di aprile sono nati **3** bimbi, peso medio **2,83**.
- Qual è il peso medio dei bimbi nati nelle prime due settimane di aprile?
- ?  $(3,42+2,83)/2=3,1$  ?? peso medio dei bambini nelle prime due settimane di aprile
- Otteniamo i dati originali di tutti i 13 bimbi:
- 2,9 3,5 3,8 3 2,9 4,2 3,3 3,4 3,2 4
- 2,8 2,7 3

### Esempio-Calcolo della media pesata-*continua*

- La media calcolata sui dati in dettaglio risulta **3,3** diversa da 3,1
- **3,3** è naturalmente quella vera, come ottenerla se non si conoscono i singoli dati?

$$\frac{3,42 \cdot 10 + 2,83 \cdot 3}{10 + 3} = 3,3 = \bar{x}$$

## Scala

	Nominale	Ordinale	ad Intervallo	di Rapporto
Media	/	/	→	→
Moda	→	→	→	→
Mediana	/	→	→	→

## CONFRONTO TRA LE MISURE DI POSIZIONE PER UNA VARIABILE QUANTITATIVA

### MODA

Buona misura quando un valore ha una frequenza relativa molto elevata



### MEDIAN

Buona misura con distribuzioni asimmetriche (es. tempo di sopravvivenza)

A

### MEDIA

ARITMETICA  
Buona misura con distribuzioni simmetriche (es. molti parametri biologici)

Facile da trattare matematicamente

Utilizza tutta l'informazione contenuta nei dati

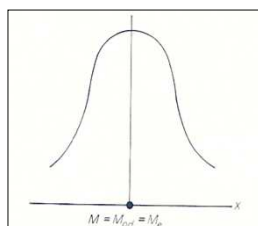


Dipende dal raggruppamento arbitrario dei dati

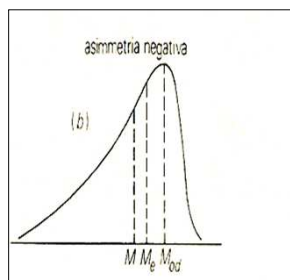
Varia molto da campione a campione

Difficile da trattare matematicamente

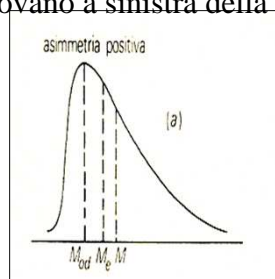
E' inaffidabile in caso di distribuzioni asimmetriche



Distribuzione **simmetrica**:  
media, moda e mediana COINCIDONO



**Asimmetrica a sinistra**: media e mediana  
si trovano a sinistra della moda



**Asimmetrica a destra**: media e mediana  
si trovano a destra della moda