INDICI STATISTICI

MEDIA, MODA, MEDIANA, VARIANZA

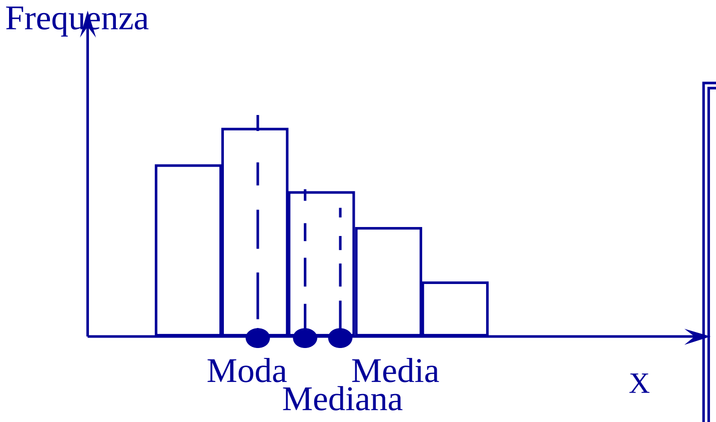
OBIETTIVI SPECIFICI: Comprendere il significato di valore medio. Comprendere il significato di variabilità di un carattere. Saper scegliere il valore medio adatto ai diversi tipi di carattere e saper confrontare valori medi diversi. Saper confrontare diverse distribuzioni in funzione del valore medio e di variabilità.

Principali Indici Statistici

INDICE

- **di posizione**
  - MODA
  - MEDIANA
  - MEDIA
- **di dispersione**
  - SCARTO QUADRATICO MEDIO
  - VARIANZA
  - RANGE
- **di forma**
  - ASIMMETRIA (SKEWNESS)
  - CURTOSI (KURTOSIS)

INDICI DI POSIZIONE



Consentono di valutare l'ordine di grandezza delle manifestazioni e servono per localizzare la distribuzione, ovvero individuare attorno a quale valore del carattere si accentra la distribuzione stessa.

Sono espressi nella stessa unità di misura con cui si estrinseca il fenomeno

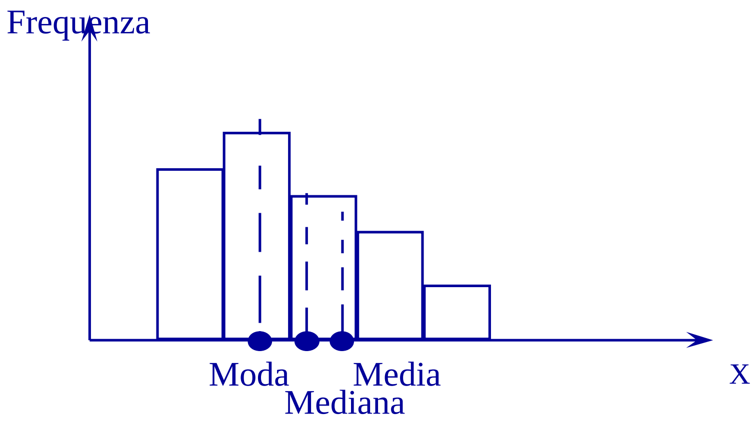
MODA

E' definita come il valore che ha la frequenza più alta.

MEDIANA

E' quel valore al di sotto del quale cadono la metà dei valori campionari.

INDICI DI POSIZIONE



MEDIA Aritmetica

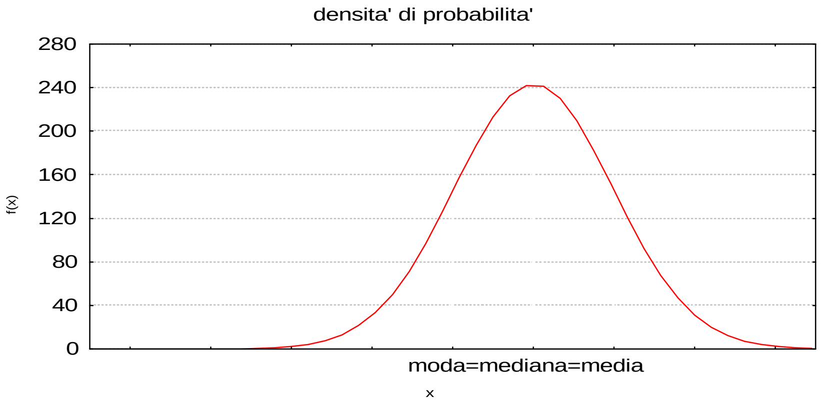
$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

E' quel valore che corrisponde alla somma di tutti i valori diviso il numero dei valori stessi.

Rappresenta il valore che sostituito a ciascun x_i lascia invariata la intensità totale (somma)

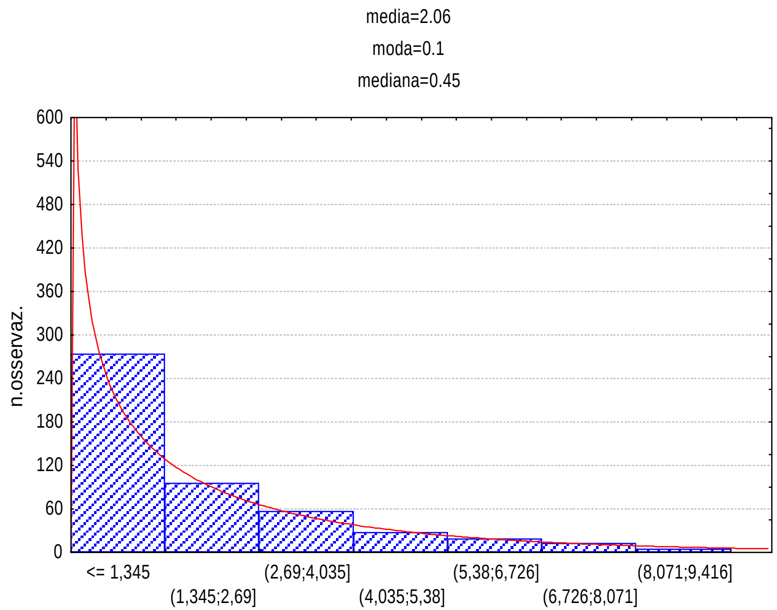
dove:
X_i = esito i-ma misura - n = numero dei dati (taglia del campione)

MODA E MEDIA sono indici di posizione, poiché la loro variazione sposta appunto la posizione della curva (verso destra o verso sinistra) in funzione del segno della variazione.



NELLA CURVA NORMALE (**simmetrica**)

MEDIA= MODA = MEDIANA

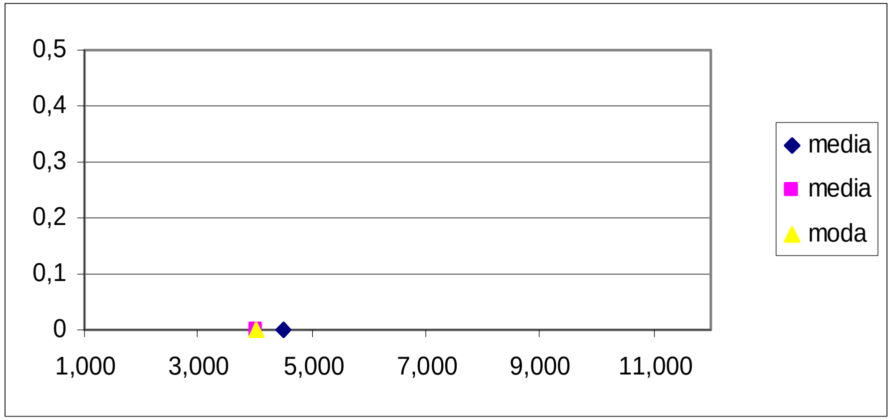


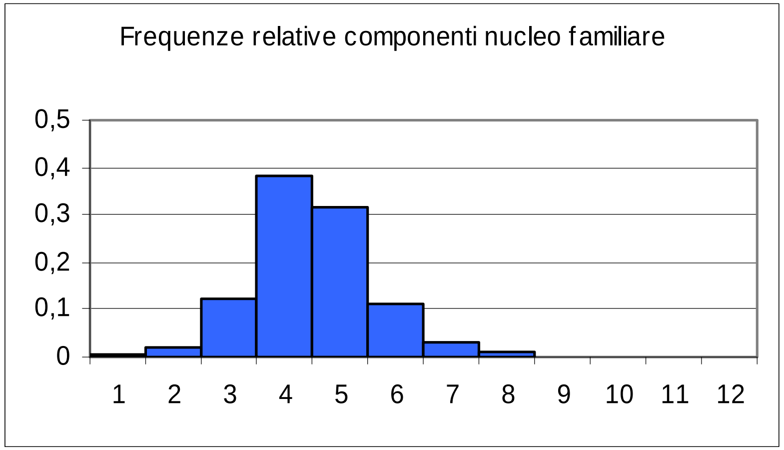
ESEMPIO 1

Numero di componenti familiari

studenti.xls utilizzo strumento di analisi: **statistica descrittiva**

COMPON	
Mean	4,513
Median	4
Mode	4





Media = 4.5 > 4 = mediana=moda

Ci sono più osservazioni alla sinistra della media

La distribuzione è più concentrata alla sinistra della media
MA
la coda è più lunga a destra

ESEMPIO 2

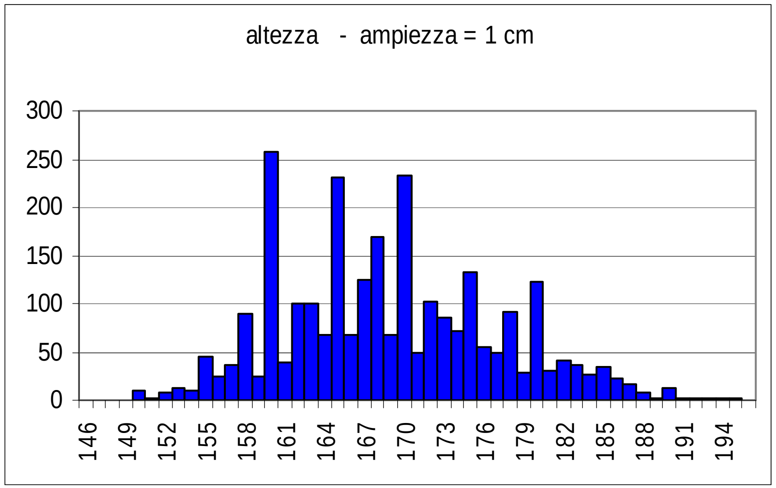
altezza

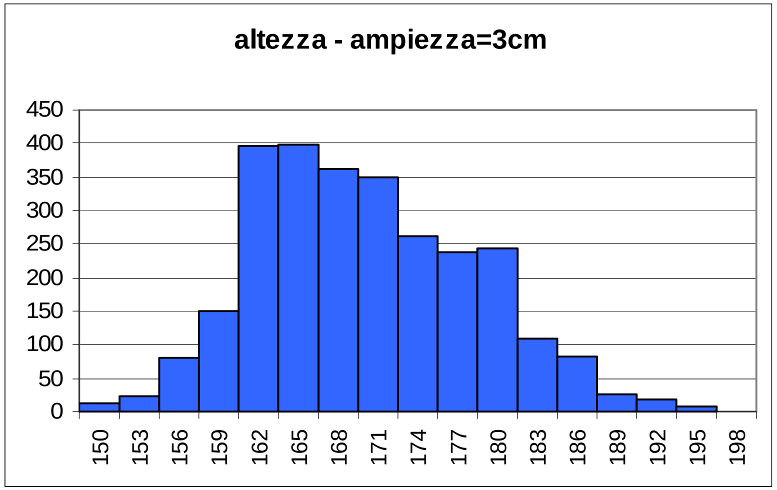
ALTEZZA	
Mean	169,011961
Median	168
Mode	160

media = 169 > 168 = mediana>Moda=160

Ci sono più osservazioni alla sinistra della media

La distribuzione è più concentrata alla sinistra della media
e
la coda è più lunga a destra





studenti.xls utilizzo strumento di analisi: **statistica descrittiva**

MISURE DI DISPERSIONE

$x_{max} - x_{min}$ **RANGE (Campo di variazione)**

$$\frac{1}{n}\sum_{i=1}^{n}\left|X_i - \overline{X}\right|$$ **SCARTO MEDIO ASSOLUTO**

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$ **VARIANZA (Media degli Scarti al Quadrato)**

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$ **VARIANZA <u>CAMPIONARIA</u> (Calcolata da Excel)**

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$ **DEVIAZIONE STANDARD <u>CAMPIONARIA</u>**

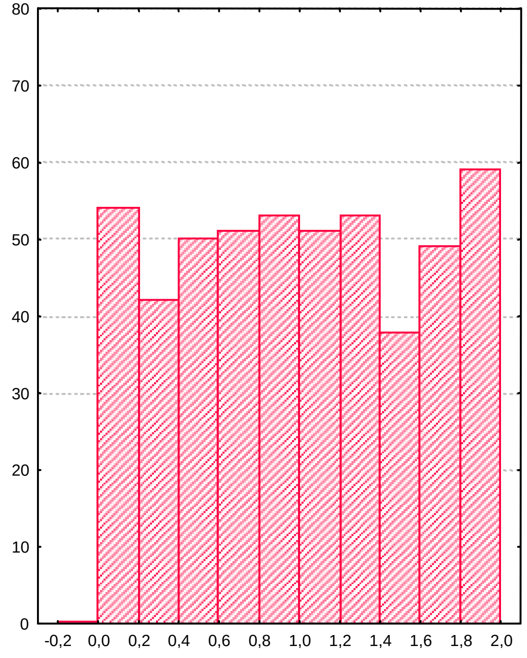
$$CV = \frac{s}{\left|\overline{X}\right|}$$ **COEFFICIENTE DI VARIAZIONE**

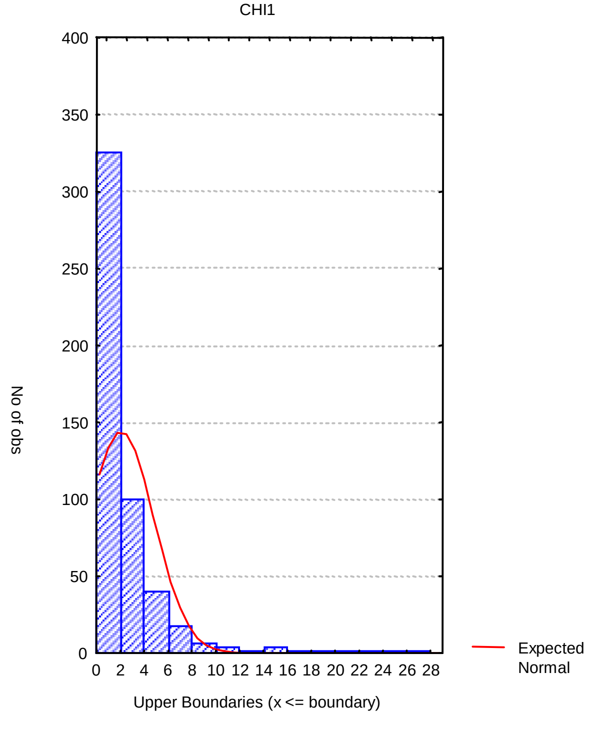
l'importanza della varianza

Media uguale

Deviazione Standard Diversa

Media uguale (2) e Varianza diversa 0,34 e 7,68 rispettivamente





ESEMPI 1-2

Componenti familiari e altezza

COMPON	
Mean	4,513
Median	4
Mode	4
Standard Deviati	1,131

ALTEZZA	
Mean	169,011961
Median	168
Mode	160
Standard Dev	8,23848231

Calcolare
1.range
2. Scostamento medio assoluto
3. varianza
4.scarto quadratico medio
5. Coefficiente di variazione

ESEMPIO 3

ETA'	
Media	21,851
Mediana	20,8
Moda	20,8
Deviazione standard	3,550
Varianza campionari	12,599
Curtosi	30,259
Asimmetria	4,463
range	44,4
Minimo	19,1
Massimo	63,5
taglia campione	2759

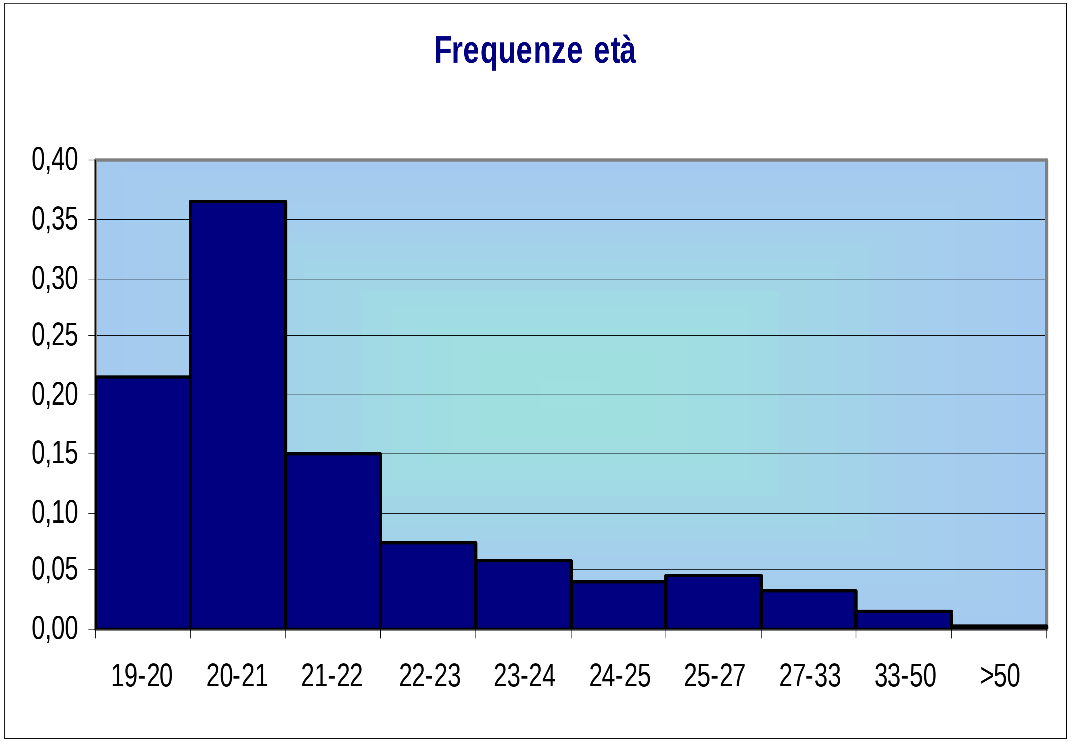
Media > mediana = = moda

Più osservazioni a sinistra della media

Dev. Stand. piccola rispetto al range

Non c'è molta dispersione

Ci si aspetta una distribuzione, con una lunga coda a destra e la maggior parte della distribuzione concentrata a sinistra con un picco intorno alla media=mediana.



I dati raggruppati

...e se non abbiamo a disposizione i dati grezzi, ma solo le distribuzioni di frequenza? Come determiniamo gli indici di sintesi?

NUMERO DI FIGLI x_i	FREQUENZA f_i
0	5
1	6
2	15
3	13
4	4
5	3
6	3
9	1
Totale	50

Tabella: Distribuzione di frequenza del numero di figli di 50 famiglie di una comunità

Caso discreto

CLASSE DI VALORI	FREQUENZA CLASSE f_i
10-19	4
20-29	66
30-39	47
40-49	36
50-59	12
60-69	4
Totale	169

Tabella: Distribuzione delle età di 169 soggetti

Caso raggruppamento per classi

Caso discreto : LA MEDIA PONDERATA

$$\overline{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i}$$

Somma degli elementi del campione

= n

taglia del campione

ESEMPIO:

NUMERO DI FIGLI x_i	FREQUENZA f_i
0	5
1	6
2	15
3	13
4	4
5	3
6	3
9	1
Totale	50

Tabella: Distribuzione di frequenza del numero di figli di 50 famiglie di una comunità

$$\overline{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i} = \frac{0\times 5 + 1\times 6 + \mathbf{K} + 9\times 1}{5+6+\ldots+1} = 2.66$$

LA MEDIA PER DATI RAGGRUPPATI IN CLASSI

m_i valore centrale della classe i-ma

f_i Frequenza della classe i-ma

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i}$$

Somma degli elementi del campione

= n

taglia del campione

ESEMPIO:

CLASSE DI VALORI	VALORE CENTRALE CLASSE **m_i**	FREQUENZA CLASSE **f_i**	**$m_i f_i$**
10-19	14.5	4	58.0
20-29	24.5	66	1617.0
30-39	35.5	47	1621.5
40-49	44.5	36	1602.0
50-59	54.5	12	654.0
60-69	64.5	4	258.0
Totale		169	5810.5

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i} = \frac{5810.5}{169} = 34.48$$

LA VARIANZA PER DATI RAGGRUPPATI IN CLASSI

m_i valore centrale della classe i-ma

f_i Frequenza della classe i-ma

$$s^2 = \frac{\sum_{i=1}^{k}(m_i - \overline{x})^2 f_i}{\sum_{i=1}^{k} f_i - 1} \quad = n - 1$$

ESEMPIO:

CLASSE DI VALORI	VALORE CENTRALE CLASSE m_i	FREQUENZA CLASSE f_i	$(m_i - \overline{x})$	$(m_i - x)^2$	$(m_i - x)^2 f_i$
10-19	14.5	4	-19.88	395.2144	1580.8576
20-29	24.5	66	-9.88	97.6144	6442.5504
30-39	35.5	47	.12	.0144	.6768
40-49	44.5	36	10.18	102.4144	3686.9184
50-59	54.5	12	20.12	404.2144	4857.7728
60-69	64.5	4	30.12	907.2144	3628.8576
Totale		169			20197.6336

$$s^2 = \frac{\sum_{i=1}^{k}(m_i - \overline{x})^2 f_i}{\sum_{i=1}^{k} f_i - 1} = -\frac{20197.6336}{168} = 120.224$$