

 Università di Ferrara <small>fondata nel 1391</small>	
<p>Elementi di statistica medica</p> <p>STATISTICA DESCRITTIVA</p>	

<p>STATISTICA DESCRITTIVA</p>	
<ul style="list-style-type: none"> ■ È quella branca della statistica che ha il fine di descrivere un fenomeno. ■ Deve quindi sintetizzare tramite pochi valori(INDICI o INDICATORI) un vasto numero di misure. 	

<p>GLI INDICI</p>	
<p>Possono essere di:</p> <ul style="list-style-type: none"> ■ Posizione ■ Dispersione ■ Forma 	

<p>INDICI DI DISPERSIONE</p>	
<ul style="list-style-type: none"> ■ I valori medi sono indici importanti per la descrizione sintetica di un fenomeno statistico. ■ Hanno però il limite di non darci alcuna informazione sulla distribuzione dei dati. ■ In statistica è possibile valutare in modo sintetico la distribuzione dei dati mediante gli indici di dispersione (detti anche di variabilità) 	

INDICI DI DISPERSIONE

- Estremi
- Campo di variazione (range)
- Scarto dalla media
- Varianza
- Deviazione standard (scarto quadratico medio)

ESTREMI

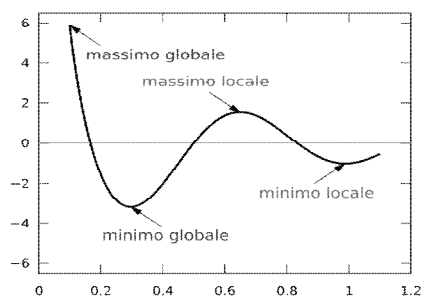
Sono costituiti dall'elemento con valore più basso e dall'elemento con valore più alto.

ESEMPIO:

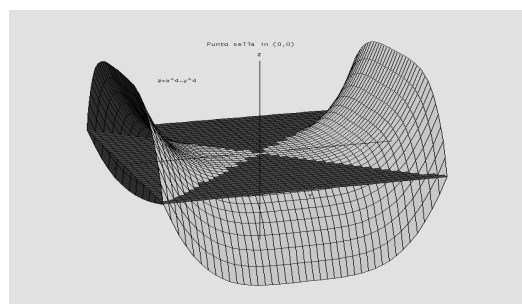
Nella serie 2 3 1 -1 5 8 0

- abbiamo come estremo inferiore (xmin) il valore -1;
- abbiamo come estremo superiore (xmax) il valore 8

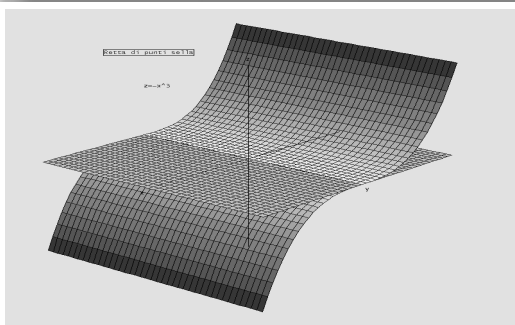
ESTREMI



ESTREMI



ESTREMI



CAMPO DI VARIAZIONE (RANGE)

- E' il più semplice degli indici di dispersione.
- Si calcola facendo la differenza tra gli estremi.
- Campo variazione = $x_{\max} - x_{\min}$
- Rappresenta l'ampiezza dell'intervallo dei dati.

1° ESEMPIO

Si prenda in considerazione la seguente serie: 1 2 3 6 9 10 15

- Il valore più alto è 15, il più basso 1.
- Il range è dato dalla differenza tra i due valori.
- $R = 15 - 1 = 14$

2° ESEMPIO

Si prenda in considerazione la seguente serie: -11 -2 3 9 10 18

- Il valore più alto è 18, il più basso -11
- Il range è dato dalla differenza tra i due valori:
 $R = 18 - (-11) = 18 + 11 = 29$

CAMPO DI VARIAZIONE	
	<p>Il campo di variazione dà informazioni sulla distribuzione dei dati:</p> <ul style="list-style-type: none"> ■ più R è piccolo più i dati sono concentrati; ■ più R è grande più i dati sono dispersi; ■ R è espresso nella stessa unità di misura dei dati. <p>Tuttavia, R tiene conto solo dei dati estremi della distribuzione e non di tutti i dati, pertanto distribuzioni diverse ma con gli stessi valori estremi hanno range uguali.</p>

SCARTO MEDIO	
	<p>Un altro modo per calcolare la variabilità dei dati (tenendo conto di tutti i dati) consiste nel calcolare la distanza, in valore assoluto, di tutti i dati dalla media e fare la media aritmetica di tali distanze.</p> <p>Scarto medio = Distanza media dei dati dalla media</p> <p>ESEMPIO: Si prenda in considerazione la seguente serie: 1 7 9 14 19</p> <ul style="list-style-type: none"> ■ Per ottenere lo scarto medio devo innanzitutto calcolare la media: Media: $50/5 = 10$ ■ Calcolo la distanza, in valore assoluto, dalla media per ogni valore della serie: 9 3 1 4 9 ■ Faccio la media dei valori ottenuti: $26/5 = 5,2$ ■ Il valore ottenuto è lo scarto medio

SCARTO MEDIO	
	<p>Lo scarto medio dalla media dà informazioni sulla distribuzione dei dati:</p> <ul style="list-style-type: none"> ■ più SM è piccolo più i dati sono concentrati; ■ più SM è grande più i dati sono dispersi; ■ SM è espresso nella stessa unità di misura dei dati; ■ non ha l'inconveniente del "Campo di variazione" in quanto SM tiene conto di tutti i dati della distribuzione.

VARIANZA E DEVIATIONE STANDARD	
	<ul style="list-style-type: none"> ■ Sono gli indici di dispersione che tengono conto della distribuzione di tutti i dati.

VARIANZA

- Rappresenta la media aritmetica dei quadrati delle distanze dei dati dalla media M.
- Essa è data dalla formula:

$$\text{Varianza} = \frac{\sum(x - M)^2}{N}$$

ESEMPIO

Calcolare la varianza della seguente serie:

2 3 5 7 8 11

- Calcolo la media: $36/6 = 6$
- Calcolo i quadrati delle distanze dei dati dalla media:
 $4 \times 4; 3 \times 3; 1 \times 1; 1 \times 1; 2 \times 2; 5 \times 5$
- Faccio la somma dei quadrati ottenuti:
 $16+9+1+1+4+25 = 56$
- Calcolo la media dei quadrati e ottengo la Varianza:
 $56/6 = 9,333$

DEVIAZIONE STANDARD

- È uguale alla radice quadrata della varianza, secondo la formula:

$$\text{Deviazione standard} = \sqrt{\frac{\sum(x-M)^2}{N}}$$

ESEMPIO

Calcolare la deviazione standard della seguente serie:

2 5 6 7 12 16

- Calcolo la media: $48/6 = 8$
- Calcolo i quadrati delle distanze dei dati dalla media:
 $6 \times 6; 3 \times 3; 2 \times 2; 1 \times 1; 4 \times 4; 8 \times 8$
- Faccio la somma dei quadrati ottenuti:
 $36 + 9 + 4 + 1 + 16 + 64 = 130$
- Calcolo la media dei quadrati e ottengo la Varianza:
 $130/6 = 21,667$
- Faccio la radice quadrata del risultato ottenuto per avere la Deviazione Standard
 $\sqrt{21,667} = 4,655$

VARIANZA E DEVIAZIONE STANDARD

La varianza (Var.) e lo scarto quadratico medio (Dev.st.) danno informazioni sulla distribuzione dei dati:

- più Var. e Dev.st. sono piccoli, più i dati sono concentrati;
- più Var. e Dev.st. sono grandi, più i dati sono dispersi.

Entrambi gli indici tengono conto di tutti i dati della distribuzione.

- La varianza è espressa mediante il quadrato dell'unità di misura dei dati, la deviazione standard nella stessa unità di misura dei dati.

FORMA

Le misure di forma sono indici sintetici utilizzati per evidenziare particolarità nella forma della distribuzione.

Vengono presi in considerazione:

- l'asimmetria
- la curtosi

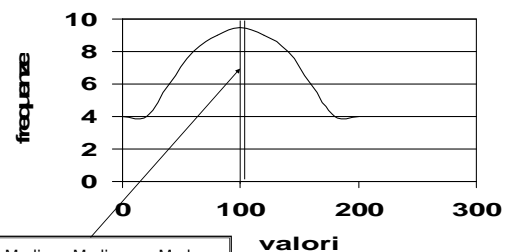
ASIMMETRIA

Una distribuzione è simmetrica quando la sua curva di frequenza presenta un asse di simmetria.

- In una distribuzione simmetrica media, mediana e moda sono coincidenti.
- In una distribuzione asimmetrica media, mediana e moda non sono più coincidenti e la differenza tra la media e la moda può essere considerata una misura della asimmetria.

ASIMMETRIA

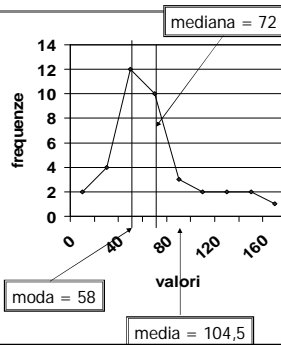
distribuzione normale



ASIMMETRIA DESTRA

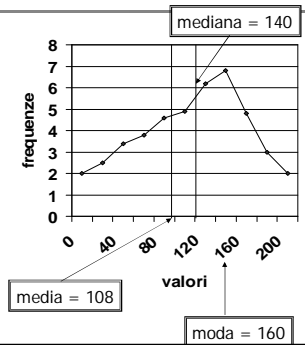
La distribuzione è asimmetrica quando non presenta nessun asse di simmetria.

- Si ha un'asimmetria positiva o destra quando il ramo destro della curva è più lungo di quello sinistro.
- In questo caso si ha:
moda < mediana < media



ASIMMETRIA SINISTRA

- Si ha un'asimmetria negativa o sinistra quando il ramo sinistro della curva è più lungo di quello destro.
- In questo caso si ha:
media < mediana < moda



CURTOSI

Se una distribuzione è simmetrica o quasi simmetrica allora può essere più o meno appuntita o più o meno appiattita rispetto alla distribuzione normale (o di Gauss).

- Se la curva è più appuntita si dice curva Leptocurtica;
- se più appiattita si dice curva Platicurtica.
- È utilizzato il coefficiente di curtosi di Pearson – $K=3$ curva normale

CURTOSI

