

STATISTICA
DESCRITTIVA

Statistica

La Statistica è la scienza che studia i fenomeni collettivi o di massa. Ci insegna a individuare i modi in cui un fenomeno si manifesta e a descriverlo sinteticamente. Ci insegna a trattare i dati raccolti e ad estrarre informazioni, sintesi, stime e previsioni da essi.

“ laddove la realtà si esprime in modo complesso, al punto che l’occhio fatica a leggerne con chiarezza i molteplici aspetti e le interrelazioni, la statistica offre gli strumenti per mettere a punto un appropriato “paio di occhiali”: un utile supporto che, come è buona norma, va comunque adottato solo dopo aver letto e compreso le appropriate prescrizioni d’uso...”

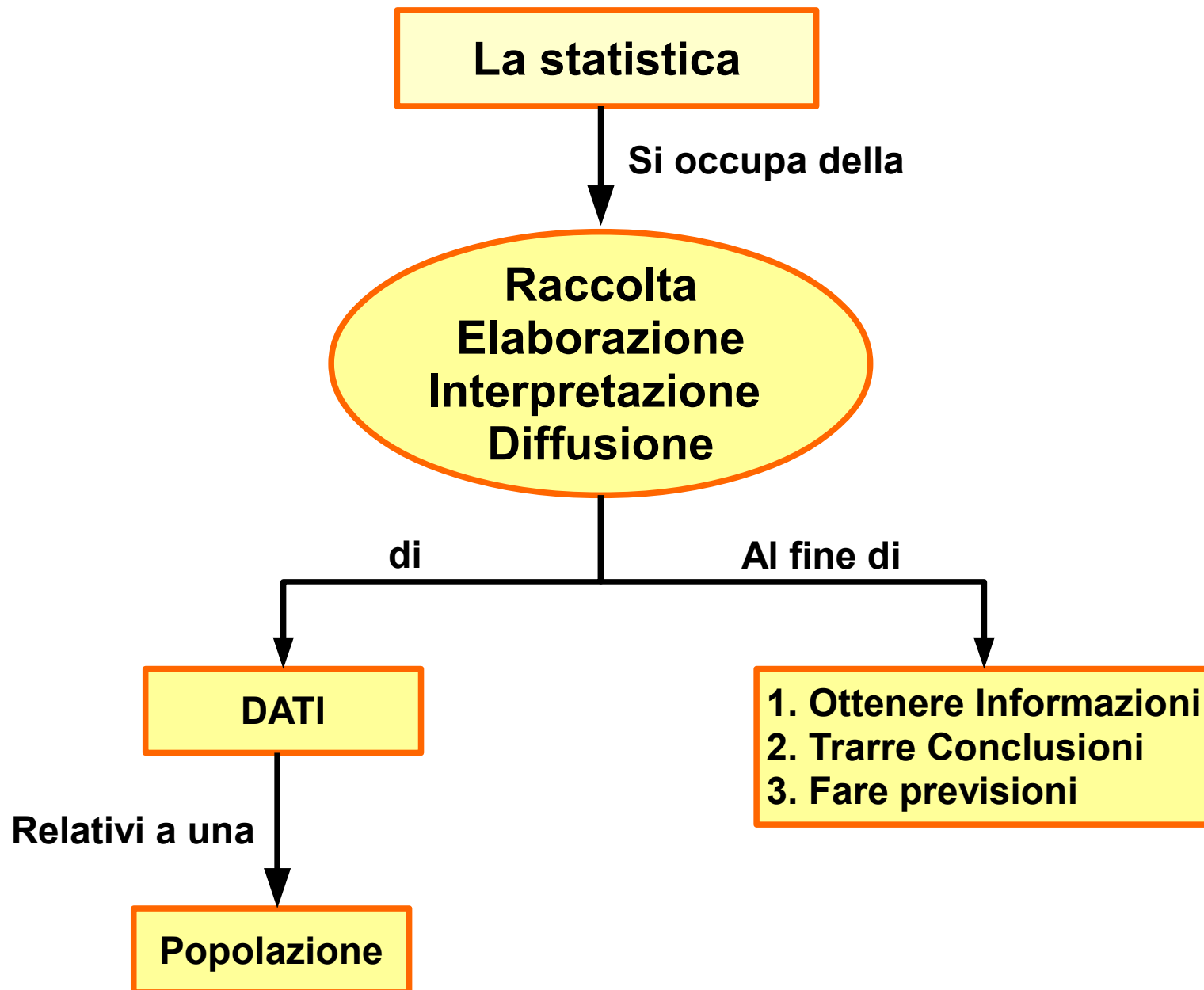
Blangiardo G.C. (2001), “Statistica, sondaggi e buon senso”

VIDEO STATISTICA

Video “The joy of stats”

Video “Population Pyramids”

Cosa è la statistica ?



Il Linguaggio della Statistica

POPOLAZIONE ED UNITA' STATISTICA

L'insieme degli individui oggetto di una indagine statistica si chiama **popolazione**.
Ciascun elemento della popolazione viene chiamato **unità statistica**.

CARATTERE

Si chiama **carattere** la proprietà che è oggetto di studio di un'indagine statistica

MODALITA'

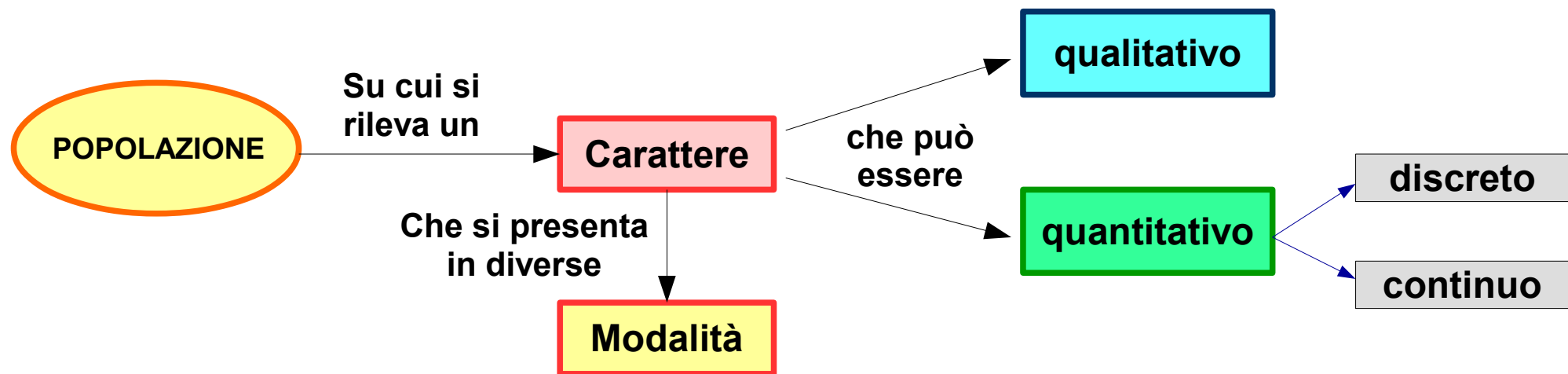
Si chiama **modalità** ciascuna delle varianti con cui un carattere può presentarsi.
Le modalità osservate sono i dati.

CARATTERI QUANTITATIVI E CARATTERI QUALITATIVI

Un carattere le cui modalità sono espresse da numeri è detto **carattere quantitativo** (o **variabile**).
Un carattere le cui modalità non sono numeriche è detto **carattere qualitativo**.

VARIABILE DISCRETA E VARIABILE CONTINUA

Una variabile è **discreta** se può assumere solo un numero finito di valori; si dice **Continua** quando può assumere tutti i valori di un determinato intervallo.



Fenomeno studiato	Popolazione	Carattere	Modalità	Tipo Carattere
Il mezzo di trasporto utilizzato per recarsi a scuola dagli studenti dell'istituto "ITT Fedi-Fermi"	Tutti gli studenti della scuola "ITT Fedi Fermi".	Mezzo di trasporto	Macchina, autobus; a piedi; motorino;	Qualitativo
Il peso degli studenti di una classe	Gli studenti della classe	Peso	55,3 ; 65,8; 60,2;	Quantitativo continuo
L'anno di nascita degli iscritti in una palestra	Tutti gli iscritti in palestra	L'anno di nascita	1972, 1990, 2000,	Quantitativo discreto

Le frequenze

FREQUENZA ASSOLUTA: Il numero di volte in cui una modalità di un carattere si presenta all'interno della popolazione

FREQUENZA RELATIVA: Il rapporto fra la frequenza assoluta ed il numero di elementi di una popolazione

FREQUENZA PERCENTUALE: La frequenza percentuale è data dalla frequenza relativa moltiplicata per 100

FREQUENZA CUMULATA: Somma di tutte le frequenze minori o uguali a quella considerata

OSSERVAZIONE: le frequenze assolute di due distribuzioni di dati non sono confrontabili perché, in generale, si riferiscono ad numero differente di elementi di una popolazione. Questo è il motivo per cui si devono usare le frequenze relative per confrontare due distribuzioni.

Indici di posizione e variabilità

MODA

Il dato o i dati che hanno la massima frequenza

VOTO	Frequenze assolute
5	4
6	8
7	4
8	2
9	1
Totale	19

La moda di questa distribuzione di dati è **6**

MEDIANA:

Ordinati gli n numeri in senso crescente la mediana è il numero che occupa la posizione centrale se n è dispari; se n è pari è la media aritmetica dei due numeri che occupano le posizioni centrali.

VOTO	Frequenze assolute
5	4
6	8
7	4
8	2
9	1
Totale	19

5 5 5 5 6 6 6 6 6 6 6 6 7 7 7 7 8 8 9

La mediana di questa distribuzione di dati è **6**.
Il valore centrale si trova alla 10° posizione.

MEDIA ARITMETICA:

Può essere calcolata solo per i caratteri quantitativi. Si ottiene sommando tutti i valori osservati e Dividendo per il numero delle unità statistiche.

VOTO	Frequenze assolute
5	4
6	8
7	4
8	2
9	1
Totale	19

La media di questa distribuzione di dati è **6,3**

$$\frac{5+5+5+5+6+6+6+6+6+6+6+6+7+7+7+7+8+8+9}{19} = \frac{(5 \cdot 4) + (6 \cdot 8) + (7 \cdot 4) + (8 \cdot 2) + (9 \cdot 1)}{19} = 6,3$$

ESEMPIO

I voti del compito di matematica dei 26 studenti della classe prima sono stati:

3 6 7 6 6 6 8 9 4 5 4 7 6 8 8 7 4 4 6 5 8 5 5 9 5 4

Costruiamo la tabella delle frequenze:

VOTO	Frequenza Assoluta	Frequenza Relativa	Frequenza Percentuale	Frequenza Cumulata
3	1	1/26 ; 0,0384	3,84%	1
4	5	5/26 ; 0,1923	19,23%	6
5	5	5/26 ; 0,1923	19,23%	11
6	6	6/26 ; 0,2307	23,07%	17
7	3	3/26 ; 0,1153	11,53%	20
8	4	4/26 ; 0,1538	15,38%	24
9	2	2/26 ; 0,0769	7,69%	26
TOT	26	1		

MODA = 6 → E' il voto che ha la maggiore frequenza assoluta

MEDIANA = 6 → dopo aver ordinato i dati devo guardare il voto presente alla 13° e 14° posizione e fare la media di questi valori. In questo caso sono entrambi uguale a 6.

$$\mathbf{MEDIA} = \frac{(3 \cdot 1) + (4 \cdot 5) + (5 \cdot 5) + (6 \cdot 6) + (7 \cdot 3) + (8 \cdot 4) + (9 \cdot 2)}{26} = 5,96$$

Oss: Per calcolare la media ogni modalità va moltiplicata per la sua frequenza.

Indici di posizione e variabilità

VARIANZA

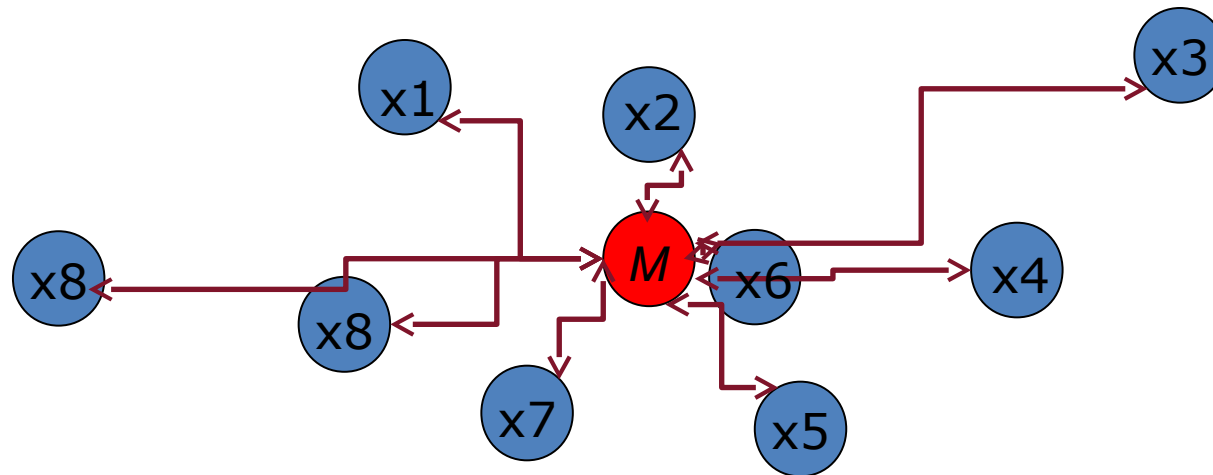
$$\frac{(x_1 - M)^2 + (x_2 - M)^2 + \dots + (x_n - M)^2}{n} = \sigma^2$$

(M è la media $M = \frac{x_1 + x_2 + \dots + x_n}{n}$)

SCARTO QUADRATICO MEDIO

$$\sigma = \sqrt{\frac{(x_1 - M)^2 + (x_2 - M)^2 + \dots + (x_n - M)^2}{n}}$$

La **VARIANZA** e lo **Scarto Quadratico Medio** danno indicazioni su quanto i valori si **discostano dalla media** e quindi di come sono distribuiti.



ESEMPIO Nella tabella sono riportati i voti di tre studenti in matematica

MARCO	6	5	6	7	MEDIA=6
LUCA	6	6	6	6	MEDIA=6
LUIGI	4	4	8	8	MEDIA=6

Gli studenti hanno la stessa media ma la distribuzione dei voti è molto differente !!

Calcoliamo lo scarto quadratico medio

	Voti di MARCO	Differenze dalla media	Quadrati delle differenze
	6	$(6-6)=0$	0
	5	$(5-6)=-1$	1
	6	$(6-6)=0$	0
	7	$(7-6)=1$	1
Somma	24	0	2
n=4		Varianza	0,5
		Scarto quadratico medio	0,71

Voti di LUCA	Differenze dalla media	Quadrati delle differenze	Voti di LUIGI	Differenze dalla media	Quadrati delle differenze
6	0	0	4	$(4-6)=-2$	4
6	0	0	4	$(4-6)=-2$	4
6	0	0	8	$(8-6)=2$	4
6	0	0	8	$(8-6)=2$	4
24	0	0	24	0	16
	Varianza	0		Varianza	4
	Scarto quadratico medio	0		Scarto quadratico medio	2

Luigi è il più Variabile !!!