

# ELABORAZIONE STATISTICA DEI DATI

La statistica studia la variabilità di un dato evento, valutandone tutte le caratteristiche.

I dati possono essere:

- **Qualitativi** - esprimono una qualità, come ad esempio il sesso, la professione, il titolo di studio.
- **Quantitativi** - esprimono una quantità data da valori numerici come ad esempio l'età, la statura, il peso, i voti di un esame.

Noi studieremo solo di dati quantitativi. La fase più importante di un'indagine statistica è l'**elaborazione dei dati quantitativi**. Tale elaborazione prevede 4 tappe:

## 1. RAGGRUPPAMENTO "SEMPLICE" O "IN CLASSI"

La prima cosa da organizzare in un'indagine statistica è la formazione del gruppo.

Il raggruppamento semplice ha un unico valore che denota il gruppo, mentre quello in classi è formato da più valori uniti secondo i seguenti criteri:

- individuare il campo di variazione, cioè il valore massimo e quello minimo che indica l'insieme di tutti i valori;
- stabilire l'ampiezza del raggruppamento, cioè quanti valori includere in una classe e deve essere la stessa quantità per tutti i dati.

**Es semplice:** In una classe di 20 alunni durante l'estate in 7 hanno letto 1 libro, in 4 hanno letto 2 libri, in 5 hanno letto 3 libri, in 3 hanno letto 4 libri e 1 solo ha letto 5 libri.

n° libri	alunni (frequenze assolute)
1	7
2	4
3	5
4	3
5	1

**Es in classi:** In una classe di 20 alunni, ognuno viene misurato in cm e si ottiene la seguente misurazione:

160,161;162;170,171,175,180,177,163,165,179,168,169,172,183,185,181,173,175,167.

(il raggruppamento è necessario perché ciascun dato è diverso dall'altro e non si riuscirebbe a creare una tabella da studiare)

altezze cm	alunni (frequenze assolute)
da 160 a 165	5
da 166 a 170	4
da 171 a 175	5
da 176 a 180	3
da 181 a 185	3

## 2. CALCOLO DELLE FREQUENZE

Bisogna ordinare i dati in base al numero di volte che si ripetono calcolando:

- **FREQUENZA ASSOLUTA:** il numero di volte che si verifica un dato statistico (fenomeno) a prescindere dal numero totale delle prove.
- **FREQUENZA RELATIVA:** è il rapporto tra la frequenza assoluta e il numero di prove eseguite. Si esprime con una frazione o viene misurata con un numero decimale compreso tra 0 e 1.

$$f_r = \frac{n^\circ \text{casi}}{n^\circ \text{totali}} = \frac{f_a}{n}$$

- **FREQUENZA PERCENTUALE:** è la frequenza relativa moltiplicata per 100. Viene misurata con un numero decimale compreso fra 0 e 100

$$f_r \% = f_r \cdot 100 = \frac{f_a}{n} \cdot 100$$

**Esempio:** elenco dei voti del compito in classe di matematica su un campione di 25 alunni

voti	frequenze assolute	frequenze relative	frequenze percentuali
9	1	1/25	4%
8	3	3/25	12%
7	5	5/25	20%
6	7	7/25	24%
5	4	4/25	16%
4	5	5/25	24%

### 3. CALCOLO DEGLI INDICI DI POSIZIONE (LE 3M)

Sono i **valori** così detti di **posizione** o anche **valori medi** o **valori centrali**:

\* **la media aritmetica** - Per calcolare la media aritmetica si deve calcolare la somma dei valori numerici, contare il numero degli elementi dell'insieme e dividere la somma per il numero degli elementi dell'insieme.

$$m = \frac{\sum n^{\circ} \text{classe} \cdot f_a}{n_{\text{tot}}}$$

**Esempio:** voto medio del compito in classe di matematica su un campione con 25 alunni:

$$m = \frac{(9 \cdot 1) + (8 \cdot 3) + (7 \cdot 5) + (6 \cdot 7) + (5 \cdot 4) + (4 \cdot 5)}{25} = \frac{9 + 24 + 35 + 42 + 20 + 20}{25} = \frac{150}{25} = 6,0$$

\* **la mediana** - calcolabile su **dati ordinati** in senso crescente o decrescente. (Vanno scritti cioè tutti in fila l'uno accanto all'altro)  
In una serie di dati ordinati, la mediana è il valore centrale.

Esistono due situazioni:

1. se la quantità di dati è pari (avremmo 2 valori centrali), va fatta la media aritmetica dei due valori;
2. se il numero dei dati è dispari il valore centrale è già unico e non va fatto nessun calcolo.

**Esempio:**

quantità **dispari** di dati

	A	B	C	D	E	F
1	Alunni	Sesso	Altezza (cm)	9 elementi ordinati		
2	Giorgia	F	140	4 elementi prima		
3	Luca	M	143			
4	Carlo	M	147			
5	Agnese	F	151			
6	Lucia	F	154	mediana in quinta posizione		
7	Elena	F	160	4 elementi dopo		
8	Federica	F	164			
9	Andrea	M	165			
10	Gianni	M	168			

quantità **pari** di dati

	A	B	C	D	E
1	Alunni	Sesso	Altezza (cm)		
2	Giorgia	F	140	mediana (151+154) : 2 = 152,5	
3	Luca	M	143		
4	Marta	F	145		
5	Carlo	M	147		
6	Agnese	F	151		
7	Lucia	F	154		
8	Elena	F	160		
9	Federica	F	164		
10	Andrea	M	165		
11	Gianni	M	168		

\* **la moda** - E' la classe che presenta il maggior numero di frequenze assolute. Non va calcolata ma riconosciuta tra i valori.

**Esempio:** voto medio del compito in classe di matematica su un campione con 25 alunni (voto con più frequenze 6)

### 3. RAPPRESENTAZIONE GRAFICA DEI DATI

Per agevolare la lettura e l'interpretazione dei dati, si ricorre alle rappresentazioni grafiche. Sono utilizzati:

- **ISTOGRAMMI** - sono grafici formati da tanti rettangoli quanti sono le classi di raggruppamento e le cui aree sono proporzionali alle frequenze di ogni classe. Siccome l'intervallo delle classi è sempre lo stesso, i rettangoli avranno tra loro uguale base. Ciò che varia è l'altezza del rettangolo che indicherà la frequenza della classe. Sull'asse y va la frequenza assoluta e sull'asse x il nome delle classi.

