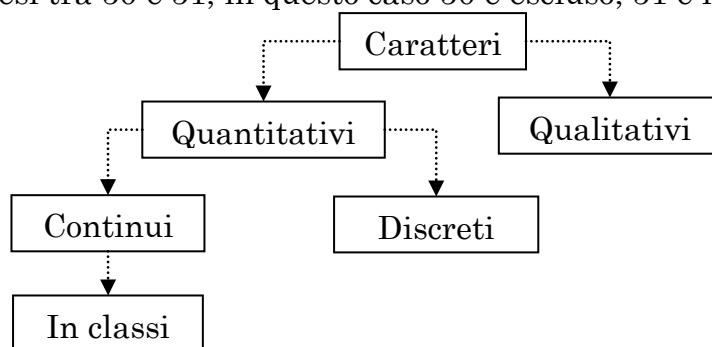


## APPUNTI DI STATISTICA

- La **statistica** è la scienza che si occupa della raccolta e della elaborazione di dati.
- La **popolazione** è l'insieme degli individui o degli elementi sui quali si vuole indagare (insieme delle persone che vivono in Italia, le aziende agrarie del Lazio, le lampadine prodotte da una certa azienda, le precipitazioni atmosferiche giornaliere in una certa località, le autovetture in circolazione in Italia, le estrazioni di numeri nel gioco del Lotto).
- L'**unità statistica** è ciascun elemento della popolazione statistica.
- Il **carattere** di una indagine statistica è il tipo di dati sul quale l'indagine si effettua: altezza, numeri di figli, titolo di studio, stato civile, fatturato.
- Un **carattere qualitativo** esprime una qualità, cioè valori non numerici come il mezzo di trasporto usato, il titolo di studio, il colore degli occhi. Può essere ordinabile (titolo di studio) o non ordinabile (colore degli occhi).
- Un **carattere quantitativo** esprime una quantità, cioè un valore numerico: numero di figli, età, altezza, peso, fatturato di un'azienda.
- I dati quantitativi si distinguono in **discreti**, se il valore numerico può essere espresso solo da un numero intero (numero di scarpa, numero di figli, anni di età) o **continui**, se il valore numerico può variare anche di piccole quantità (peso, altezza, superficie di un appartamento).
- Se i dati sono continui è opportuno organizzarli in **classi** cioè intervalli di valori tra i quali oscilla un gruppo (peso compreso tra 50 e 51 Kg, si può indicare come  $50 < p \leq 51$ , raccoglie tutti i valori compresi tra 50 e 51, in questo caso 50 è escluso, 51 è incluso).



- La **rilevazione** dei dati statistici consiste nel raccogliere le informazioni relative a dei caratteri oggetto di indagine e nel raggruppare i dati ottenuti.
- Il **censimento** è una indagine statistica che riguarda tutti gli elementi della popolazione
- Il **rilevamento per campione** riguarda un sottoinsieme della popolazione scelto in maniera opportuna in modo che sia rappresentativo dell'intera popolazione.
- La **frequenza assoluta** è il numero di volte con cui si rileva lo stesso valore; la **frequenza relativa** si ottiene facendo il rapporto tra la frequenza assoluta e il numero totale delle rilevazioni, può variare da 0 a 1; la **frequenza percentuale** si ottiene moltiplicando per 100 la frequenza relativa e aggiungendo il simbolo di percentuale %; le **frequenze cumulate** si ottengono sommando la frequenza percentuale di ciascun valore a quelle dei precedenti.
- La **moda** di un insieme di dati è il valore, o classe di valori, che si presenta con maggiore frequenza.
- La **media aritmetica** si ottiene sommando tutti i dati e dividendo il risultato per il numero di dati. La **media aritmetica ponderata** si ottiene moltiplicando i dati per il numero di volte in cui compaiono, si sommano tutti i prodotti ottenuti e si divide il risultato per il numero totale di dati.
- La **mediana** si ottiene ordinando i dati in modo crescente (o decrescente) e prendendo il valore che occupa il posto centrale, nel caso in cui i dati siano in numero pari avremmo due posizioni al centro e per trovare la mediana sommiamo i due valori centrali e dividiamo per due il risultato.

- Il **campo di variazione** di un insieme di dati statistici è la differenza tra il valore più grande e quello più piccolo.

Esempio. Uno studente ha preso i seguenti voti in pagella: 6, 6, 5, 7, 5, 6, 8, 4. La media

aritmetica è data da  $\frac{6+6+5+7+5+6+8+4}{8} = 5,875$ . La stessa media si può calcolare come

media aritmetica ponderata  $\frac{3 \times 6 + 2 \times 5 + 7 + 8 + 4}{8} = \frac{18 + 10 + 7 + 8 + 4}{8} = 5,875$ .

I grafici ci aiutano a interpretare i dati più rapidamente rispetto a una tabella.

- **Istogramma** è un grafico costituito da rettangoli di uguale base ma di altezza variabile proporzionale alla frequenza assoluta.
- **Areogramma** è un grafico costituito da un cerchio suddiviso in settori circolari la cui ampiezza è proporzionale alla frequenza con cui il valore corrispondente compare;
- **Ideogramma** è un grafico in cui si utilizzano immagini stilizzate che richiamano la natura dei dati di cui ci si occupa (es. spighe di grano per la produzione di frumento, damigiane per la produzione di vino) la grandezza del disegno o il numero di volte che viene rappresentato è proporzionale alla frequenza con cui è stato rilevato il carattere oggetto di indagine.
- **Cartogramma** è una rappresentazione grafica che fa uso di carte geografiche sulle quali vengono riportati i dati rilevati dall'indagine.
- **Diagramma cartesiano** è un grafico in cui i punti hanno come coordinata orizzontale (ascissa) il valore rilevato e come coordinata verticale (ordinata) la sua frequenza assoluta; questi punti si possono unire a formare la **poligonale delle frequenze** (una linea spezzata), nel caso in cui le rilevazioni sono molto fitte e numerose la poligonale assume una forma sempre più arrotondata e nel caso in cui l'andamento della variabile fosse normale, si avrebbe una **forma a campana** detta anche **gaussiana**.
- **ISTAT** è l'Istituto nazionale di statistica un ente di ricerca pubblico italiano che effettua censimenti e indagini campionarie sulle famiglie e sulla società italiana.

### Esempio

25 ragazzi di una classe hanno provato a pesare il libro di matematica ottenendo i seguenti risultati in Kg: 1,225; 1,200; 1,250; 1,200; 1,225; 1,225; 1,200; 1,150; 1,200; 1,225; 1,250; 1,250; 1,225; 1,200; 1,250; 1,250; 1,225; 1,225; 1,200; 1,200; 1,225; 1,225; 1,200; 1,225; 1,200

Peso	Frequenza assoluta	Frequenza relativa	Frequenza Percentuale	Frequenze cumulate	
1,150	1	1/25	4%	4%	Moda 1,225 Media aritmetica 1,218 Mediana 1,225 Campo di variazione 0,100
1,175	0	0/25	0%	4%	
1,200	9	9/25	36%	40%	
1,225	10	10/25	40%	80%	
1,250	5	5/25	20%	100%	

Istogramma delle frequenze

