

---

## Perché studiare statistica?

---

Le decisioni in qualsiasi campo vengono prese pur non avendo certezze a sostegno, per questo si cerca di ottenere più informazioni possibili in modo da prevedere il più possibile le conseguenze.

Per ottenere informazioni solitamente si effettua un campionamento, tuttavia non sempre si analizza la popolazione completa in quanto avremmo costi proibitivi o una complessità troppo elevata e per questo ci limitiamo a una parte della popolazione. La statistica non può prescindere dal contesto.

La **popolazione** è l'insieme e di tutte le unità oggetto di studio.

**Unità statistica** → elemento della popolazione, ci sono 2 tipologie:

- Semplici
- composte

**Variabile** → caratteristica rilevata sulle unità statistiche

**Modalità** → valori assunti da una variabile

**Dati** → risultato delle rilevazioni

Il **campione** è il sottoinsieme delle unità osservate della popolazione e la sua dimensione viene indicata con  $n$

L'obiettivo del campionamento è quello di ottenere maggiori informazioni in modo da poter effettuare delle scelte più affidabili.

**Campionamento casuale semplice** → metodo in cui ogni elemento del campionamento viene scelto rigorosamente a caso ed ha la stessa probabilità di essere scelto degli altri. Il campione scelto con questa modalità viene definito **campione casuale**.

L'utilizzo del campione è quello di venire analizzato in modo da poter trarre conclusioni sull'intera popolazione.

**Parametro** → caratteristica specifica della popolazione

**Statistica** → caratteristica tipica del campione

### Passi del metodo statistico

- **Disegno** → come otteniamo i dati che servono a rispondere agli obiettivi della ricerca
- **Descrizione** → come possiamo descrivere e sintetizzare i dati raccolti
- **Inferenza** → come possiamo generalizzare i risultati, fare previsioni e prendere decisioni

La scelta dei dati avviene a partire da un contesto, vengono astratti grazie ad una metodologia statistica e poi in base al contesto vengono interpretati.

Esistono diverse tipologie di statistica:

- **Statistica descrittiva** → comprende metodi grafici e numerici che sono usati per sintetizzare ed elaborare i dati in modo da trasformarli in informazioni e descriverne le caratteristiche.
- **Statistica inferenziale** → fornisce le basi per le previsioni e per le stime che consentono di trasformare le informazioni in conoscenza. Questa tipologia di statistica è un processo, non un semplice risultato numerico, composto da stime, verifiche di ipotesi, analisi di relazioni e previsioni. Ha lo scopo di generalizzare il risultato all'intera popolazione.

Problemi che possono scaturire dai dati

#### Dati errati

- A livello concettuale → ambiguità nella definizione del fenomeno oggetto di studio
- Arrotondamenti
- Errori di inserimento di dati

#### Dati mancanti

- Strutturali → dati necessariamente mancanti
- Non strutturali → dati che dovrebbero essere presenti, ma non lo sono per:
  - Motivi accidentali
  - Motivi legati allo stesso

#### Dati incompleti

I dati strutturali mancanti non sono un problema, mentre quelli non strutturali lo sono in quanto:

- Riducono il contenuto informativo
- Distorcono il risultato

Ci sono distorsioni se la mancanza del dato è legata al valore che si sarebbe osservato.

Tipi di variabili

- **Qualitative** → misurano delle qualità, queste variabili vengono codificate in modo da poter essere analizzate più facilmente. Possono essere:
  - Nominali → non ordinabili
  - Ordinali → ordinabili
- **Quantitative** → misurano caratteristiche numeriche. Possono essere:
  - Discrete → insieme finito o numerabile
  - Continue → insieme infinito non numerabile

I dati possono essere raccolti attraverso un esperimento o uno studio osservazionale, la differenza principale è che nel primo caso il ricercatore interviene manipolando la realtà mentre nel secondo ciò non si verifica.