

Département d'italien

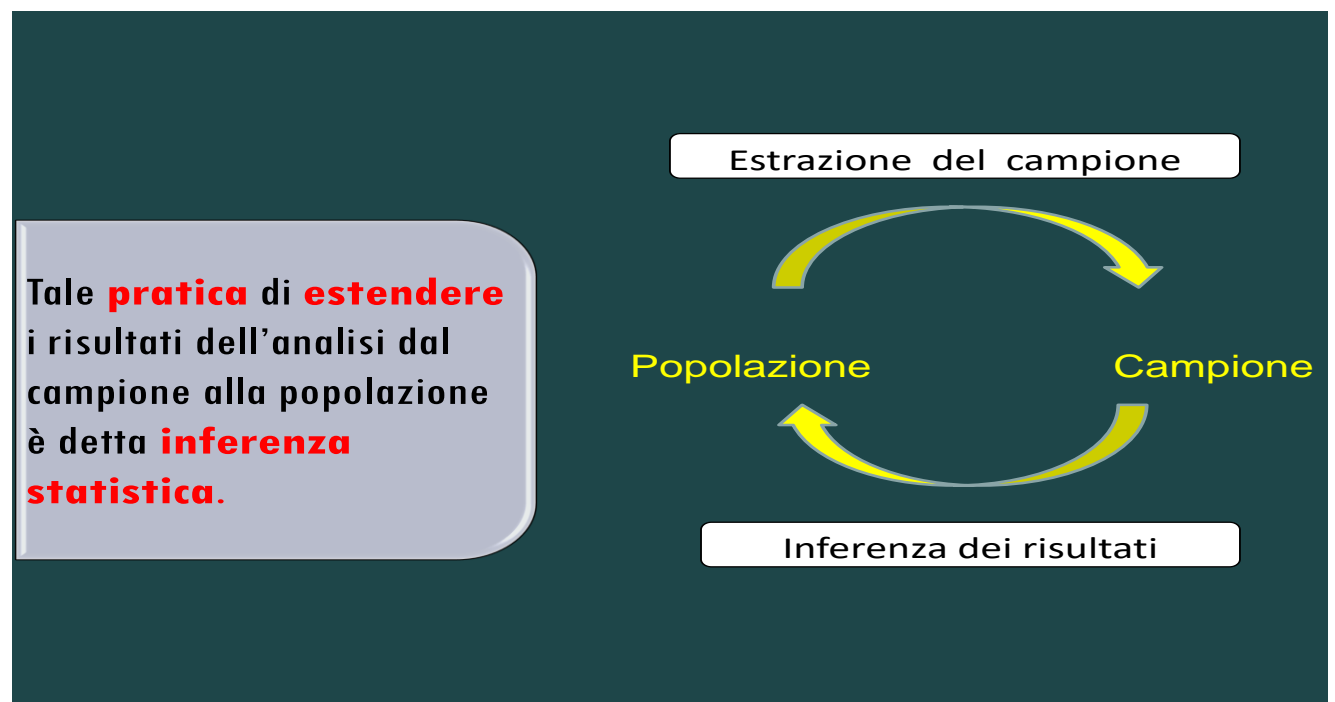
Master I : Module de initiation à la recherche

Cours réalisé par : Dr. Hachouf Amina

## 4.2. Campione e campionamento

- **Campione:** è qualsiasi **sotto-insieme** dei **membri** della **popolazione** oggetto della ricerca. È una porzione della popolazione oggetto di indagine che il ricercatore andrà a studiare con la speranza di ottenere da essa delle informazioni estendibili all'intera popolazione.

La qualità e la rappresentatività di un campione non è data dalla sua ampiezza, ma è data dal modo in cui esso viene estratto



**Perché si possa compiere inferenza statistica è necessario che il campione sia casuale.**

- **tutti i membri della popolazione da cui il campione viene estratto hanno la stessa probabilità di entrare a farne parte.**

Se si va all'uscita di un supermercato e si intervistano le prime 100 persone che passano di lì, non si otterrà un **campione casuale** dei consumatori italiani:

le **probabilità** di entrare a **far parte** di quel **campione** non sono uguali per tutti i **consumatori**:

- alcuni potrebbero non frequentare quel supermercato,
  - altri potrebbero aver scelto di non recarvisi quel giorno.
- 
- Per ottenere un campione casuale di una popolazione si deve allora necessariamente avere una lista completa dei membri di quella data popolazione ed estrarre quelli che entrano nel campione mediante un procedimento che assegni la stessa probabilità di essere estratto a ciascun membro.
  - La natura casuale o meno di un campione dipende integralmente dal **procedimento di estrazione** e non ha niente a che vedere con il suo risultato.

Ecco un esempio concreto: se abbiamo un'urna

- ne estraiamo delle palline, senza guardarvi dentro e senza rendere le palline identificabili al tatto
- il campione è perfettamente casuale qualunque sia il risultato: anche se l'urna contiene 20 palline rosa e 100 verdi ed estraiamo 20 palline rosa e nessuna verde.

- **Un campione si dice “casuale” quando l'estrazione degli oggetti non è influenzata in alcun modo da alcun fattore che non sia il caso.**

Accanto al concetto di casualità spesso troviamo associato il concetto di rappresentatività.

Un campione si definisce rappresentativo se in esso sono presenti caratteristiche in porzioni uguali a quelle della popolazione.

Per quello che riguarda la rappresentatività è dunque necessario tenere bene in considerazione i seguenti punti:

- la rappresentatività non dipende dal procedimento ma dal risultato dell'estrazione;
- un campione rappresentativo rispetto a una proprietà non lo è rispetto alle altre;
- un campione rappresentativo rispetto a una proprietà non lo è rispetto alle altre;
- la rappresentatività va controllata, ed eventualmente attestata, separatamente per ogni proprietà;
- Poiché la rappresentatività si valuta confrontando la distribuzione di una caratteristica nel campione con la distribuzione della stessa caratteristica nella popolazione, **essa** si può giudicare unicamente per quelle proprietà di cui si conosce la distribuzione nella popolazione, cioè solo per le caratteristiche rilevate attraverso un censimento di quella popolazione.

### **NB. Casuale e rappresentativo**

Nell'estrazione casuale, le distorsioni nella rappresentatività non sono introdotte dal ricercatore, ma si producono per effetto del caso.

Ciò permette di applicare le formule elaborate dagli statistici per stabilire, adottando una serie di assunti:

il **grado di fiducia** che possiamo avere nel fatto che la distribuzione di una certa caratteristica nel campione sia rappresentativa della distribuzione della stessa caratteristica nella popolazione.

L'estrazione casuale fornisce anche la ragionevole garanzia che le distorsioni nella rappresentatività siano di **grandezza limitata**.

L'esempio dell'urna sopra citato è senza dubbio un esempio estremo.

Se un campione di ambito nazionale è estratto casualmente c'è una probabilità quasi nulla che tutti gli estratti:

- frequentino lo stesso supermercato,
- come invece accadrebbe se andassimo davanti ad un solo supermercato ad intervistare quelli che vi sono entrati con l'unica preoccupazione di garantire la rappresentatività rispetto al genere.

### **Motivi per utilizzare un campione**

Questa parte della popolazione "il campione" viene selezionata dal ricercatore in modo da diminuire, in termini di tempo e costi, l'acquisizione tutte le informazioni utili per la ricerca.

In questo senso, si può intendere il campione come la rappresentazione in piccolo di tutta la popolazione che ci permette di sintetizzare tutte le caratteristiche della popolazione originaria.

Ci sono cinque *motivi* principali *per utilizzare un campione* , come mostra il riquadro

#### **Motivi per utilizzare un campione**

- L'estrazione di un campione richiede meno tempo rispetto all'esame dell'intera popolazione
- **un campione è meno costoso**
- un campione è più pratico da gestire
- **un campione garantisce un elevato grado di approfondimento e accuratezza**
- spesso è una necessità perché l'esame dell'intera popolazione non è accessibile
- **Le valutazioni sulle caratteristiche della popolazione si devono perciò basare sulle informazioni contenute in un campione estratto dalla popolazione.**

Perché i risultati di una indagine campionaria siano generalizzabili è necessario valutare la validità (bontà) del campione. Tale valutazione può essere condotta ponendosi due quesiti:

1. Le conclusioni sono corrette per gli individui che compongono il campione?
2. Il campione rappresenta bene la popolazione da cui è stato estratto?

Un campione è valido (bontà di un campione) se è:

1. **eterogeneo**: se include tutte le caratteristiche e qualità diverse. Il campione deve essere diversificato al suo interno in modo da presentare una variabilità di caratteristiche collegate alle informazioni da rilevare (tutti studenti)
2. **rappresentativo**: ovvero se presenta, senza distorsioni, tutte le caratteristiche della popolazione di riferimento. La rappresentatività è garantita se la procedura di campionamento è casuale ossia se è regolata dalla legge caso/probabilità
3. **accurato**: ovvero se il grado di minimizzazione degli errori di copertura (lista della popolazione) è elevato ed il numero di non risposte (tutti i casi del campione sono raggiungibili) è basso
4. **ampio**: ovvero se include un numero elevato di casi. L'ampiezza è inversamente proporzionale agli errori di rilevazione