

Statistica Sociale

Amin Gino Fabbrucci Barbagli

Orari e contatti

Statistica Sociale - 15ore (3 CFU)

- _amingino.fabbruccibarbagli@phd.units.it
- **Lezioni:** Mercoledì, 9:00 -11:00
Aula 1, Via Lazzaretto vecchio 8.
Codice MS Teams: go1pg8y
Per le registrazioni delle lezioni
- Moodle: [**CORSO237SF - STATISTICA SOCIALE 2025**](#)
- Per il materiale (slides lezioni e altro)
- Ricevimento:
 - online su MS Teams o in presenza mercoledì 15-17 (inviare una mail il giorno prima)
- **Esame:** scritto di circa 1 ora con domande V/F sul programma di esame e brevi esercizi simili a quelli visti a lezione
- **Libro di testo:**
 - A. Agresti, "Metodi statistici di base e avanzati. Per le scienze sociali", 5/Ed. a cura di A. Petrucci e M. Porcu, Pearson (2022)
 - Slide delle lezioni

Argomenti ed obiettivi del corso

- Rilevazione dei fenomeni statistici: unità statistica, popolazione e campione.
- Distribuzione di frequenza e rappresentazione grafica di un carattere
- Sintesi della distribuzione di un carattere: indici di posizione e variabilità
- Analisi congiunta (bivariata) di due caratteri
- Associazione tra due caratteri

Alcuni suggerimenti:

1. Seguite attivamente: prendete appunti, fate domande ed intervenite (non esistono domande sbagliate e anche altri potrebbero avere lo stesso dubbio)
2. Rivedete gli appunti, svolgete gli esercizi e approfondite gli argomenti consultando il libro di testo. Confrontatevi tra voi
3. **Venite o chiedete un ricevimento**

La statistica

La **statistica** ci aiuta a comprendere ed interpretare i fenomeni del mondo che ci circonda

Informazioni politiche, indagini sulle opinioni, notizie...

... in qualità di esperti di scienze sociali è importante saper leggere ed interpretare report contenenti informazioni statistiche (su fenomeni sociali)

La **statistica** è la scienza *dei dati*: li **definisce**, li **raccoglie**, li **organizza**, li **analizza** e li **interpreta**

La **statistica** è quindi l'insieme delle metodologie finalizzate alla raccolta e all'analisi dei **dati**

1. **Progettare studi ed indagini**: pianificare la raccolta dei dati
2. **Descrivere i dati**: attraverso strumenti di sintesi in grafici, tabelle, indici -> **statistica descrittiva**
3. **Inferire**: formulare previsioni sui dati raccolti -> **statistica inferenziale**

La statistica

Richiede diverse fasi che possono essere tradotte in:

Osservazione-> analisi-> comprensione

Che realizziamo come:

Raccolta dei dati-> elaborazione-> trasformazione dei dati in informazione

I fenomeni di interesse della statistica sono i fenomeni statistici.

Una rilevazione è l'operazione di raccolta dei dati ai fini statistici.

La raccolta di informazioni è il cuore della scienza, attraverso essa vengono ottenute le **osservazioni** impiegate per l'analisi statistica

Le osservazioni raccolte su determinate caratteristiche di interesse diventano i nostri **dati**

Ad esempio, in uno studio condotto su 1000 individui (persone)

vengono osservate caratteristiche quali

- l'opinione politica
- l'affiliazione a un determinato partito politico
- gli orientamenti politici
- l'intensità con cui vengono seguite le cerimonie religiose
- gli anni di istruzione
- il reddito annuo
- lo stato civile
- genere

Unità statistica

È presso le unità statistiche che è possibile registrare ed osservare le manifestazioni di un fenomeno di interesse.

Essa è l'unità base su cui si osserva e si misura il fenomeno.

L'insieme delle unità statistiche è detto popolazione.

La **popolazione U** è costituita dal totale dei soggetti di interesse in uno studio.

Un **campione** è un sottoinsieme della popolazione di riferimento dello studio.

L'insieme delle unità che compongono la popolazione di riferimento è detto numerosità N.

Quindi N indicherà la dimensione dell'intera popolazione.

Infine, definiamo x come variabile (qualitativo o quantitativo).

Esercizio

Vuoi studiare le abitudini di lettura di **3.000 studenti**. Ne intervisti **150** chiedendo loro **quanti libri hanno letto**.

Quali sono la popolazione totale, il campione e la variabile?

Esercizio

Vuoi studiare le abitudini di lettura di **3.000 studenti** (popolazione totale). Ne intervisti **150** (campione) chiedendo loro **quanti libri hanno letto** (variabile x)

Esercizio

Individuare l'unità statistica?

Qual è il fenomeno?

Qual è la numerosità della popolazione N ?

Cosa rappresentano i 150 studenti intervistati?

Cosa rappresenta la variabile x ?

Esercizio

Individuare l'unità statistica?

Il singolo studente. L'unità statistica è l'unità base su cui si osserva e misura il fenomeno.

Qual è il fenomeno?

In questo caso, "leggere libri"

Qual è la numerosità della popolazione N?

$N = 3.000$

Cosa rappresentano i 150 studenti intervistati?

Il campione. Un campione è un sottoinsieme della popolazione di riferimento dello studio

Cosa rappresenta la variabile x?

Il numero di libri letti.

Definizioni di base

La **popolazione** (statistica) è l'insieme dei soggetti di interesse di uno studio

I **soggetti** (o unità statistiche) possono essere persone fisiche o anche famiglie, imprese...

Può anche essere una popolazione ipotetica che non esiste realmente (es automobili prodotte)

Un **campione** è un sottoinsieme della popolazione di riferimento

Censimento vs indagine campionaria per raccogliere informazioni (**dati**) su una popolazione

- Possibilità
- Costi

La **statistica descrittiva** è impiegata per sintetizzare le informazioni riferite ad una popolazione (spesso raccolte in un'indagine o presenti in un archivio amministrativo)

La **statistica inferenziale** è impiegata per ottenere previsioni su una popolazione sulla base di informazioni raccolte su un campione selezionato da essa

Un **parametro** è una sintesi numerica delle caratteristiche della popolazione

Una **statistica** è una sintesi numerica dei dati campionari

Esempio

La General Social Survey (GSS) è un'indagine campionaria biennale condotta dal National Opinion Research Center dell'Università di Chicago.

L'indagine coinvolge circa 3.000 adulti statunitensi.

Esempi di quesiti formulati sono:

"Saresti disposto a pagare prezzi più alti per i prodotti che acquisti se questo servisse a proteggere l'ambiente?"

"Ritieni che un bambino in età prescolare possa soffrire se la mamma lavora?"

Alla domanda "Credi nel paradiso?", l'86% del campione ha risposto "Sì"

Tramite metodi inferenziali, possiamo affermare che la percentuale di adulti statunitensi che crede nell'esistenza del paradiso è compresa tra l'84% e l'88%.

Possiamo concludere che una gran parte della popolazione statunitense crede nel paradiso.

Fasi di un'analisi statistica

Formulare una domanda tradurre un'esigenza conoscitiva in modo che sia suscettibile di una risposta in termini statistici

Individuare o raccogliere i dati è un ambito vasto, che va sotto il nome di disegno di indagine e campionamento

Organizzare e/o elaborare i dati: dalla massa di dati così com'è non è immediato estrarre le informazioni che servono, si possono però sintetizzare opportunamente e/o rappresentarli graficamente, in funzione delle informazioni cercate

Modellare: si formula (specifica) un modello probabilistico che possa spiegare i dati osservati in base alle ipotesi fatte sul fenomeno, si stima il modello usando i dati

Esercizio

Il sondaggio annuale sull'ambiente 2008 di Gallup indica che:

- il 28% degli intervistati ha apportato importanti cambiamenti alle proprie abitudini di vita e di acquisto negli ultimi cinque anni per proteggere l'ambiente,
- il 55% degli intervistati ha apportato modifiche minori
- il 17% non ha apportato modifiche

I risultati si basano su interviste telefoniche con 1102 persone che vivono negli Stati Uniti, dai 18 anni in su, condotte dal 6 al 9 marzo 2008.

- Qual è la popolazione di interesse?
- Descrivere il campione utilizzato
- Qual è l'oggetto dell'analisi?

Selezione casuale

I **metodi statistici inferenziali** fanno uso delle statistiche campionarie per fare previsioni sui parametri delle popolazioni

L'**utilità** dell'inferenza dipende da quanto bene il campione rappresenta la popolazione

- È importante ridurre la probabilità di selezionare campioni che per le loro caratteristiche possano **distorcere** la rappresentatività della popolazione portando ad errate conclusioni inferenziali sui valori dei parametri

Campione casuale (semplice) di n soggetti estratti da una popolazione è tale se ogni possibile campione di pari numerosità ha uguale probabilità di selezione

- n è la numerosità/dimensione campionaria
- Per selezionare un campione è necessario avere una **lista di campionamento**

Indagine campionaria: selezione di un campione dalla popolazione di riferimento; le informazioni sono raccolte tramite intervista diretta, telefonica o autocompilata (per es., online)

Esperimento: confrontare le risposte sotto diverse condizioni (trattamenti) su cui si ha controllo sperimentale; il piano sperimentale è il processo attraverso cui il ricercatore assegna i soggetti ai diversi trattamenti (in modo casuale)

Studio osservazionale: si osservano i valori delle variabili di interesse ma non si ha controllo sperimentale; non è possibile determinare i rapporti causa-effetto

