

Statistica Sociale

Susanna Zaccarin

13/03/2025

Orari e contatti

Statistica Sociale - 15ore (3 CFU)

- susanna.zaccarin@deams.units.it
- **Lezioni:** Giovedì, 15:00 -18:00
Aula 3, Via Lazzaretto vecchio 8.
MS Teams: [CD2024 237SF STATISTICA SOCIALE](#)
 - Per le registrazioni delle lezioni
- Moodle: [237SF - STATISTICA SOCIALE 2024](#)
 - Per il materiale (slides lezioni e altro)
- Ricevimento:
 - Mercoledì 15:00-17:00, stanza 2.15 – Edificio D (II semestre)
 - online su MS Teams (su richiesta)
- **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)
 - Lucidi delle lezioni

Calendario lezioni

Marzo 2025: (13), 20, 27

Aprile 2025: 3, 10

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

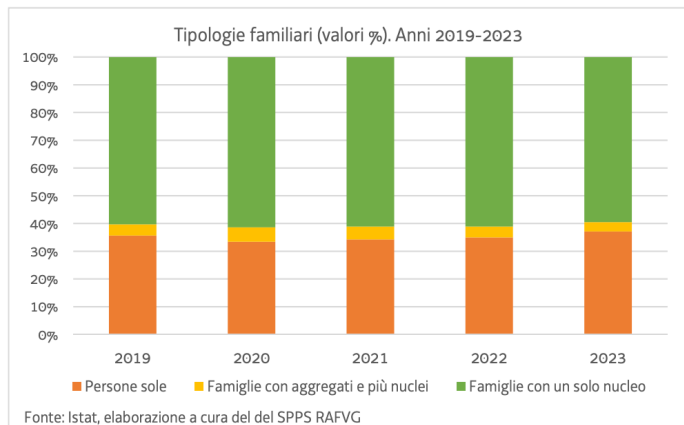
Informazioni politiche, indagini sulle opinioni, notizie...

... 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 l'ottenimento 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**

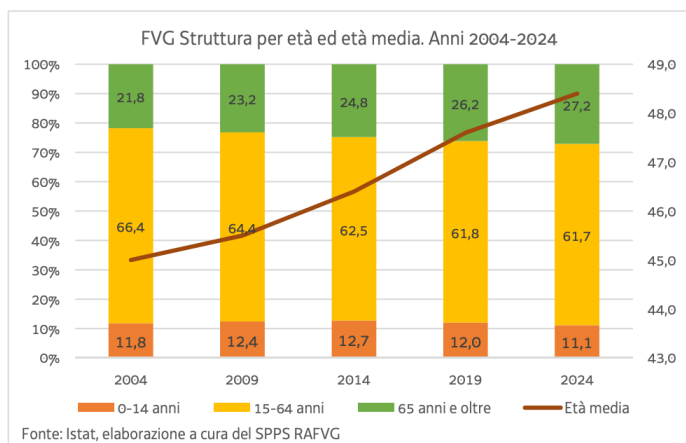


Esempio

L'EVOLUZIONE DELLA POPOLAZIONE REGIONALE E DELLE FAMIGLIE

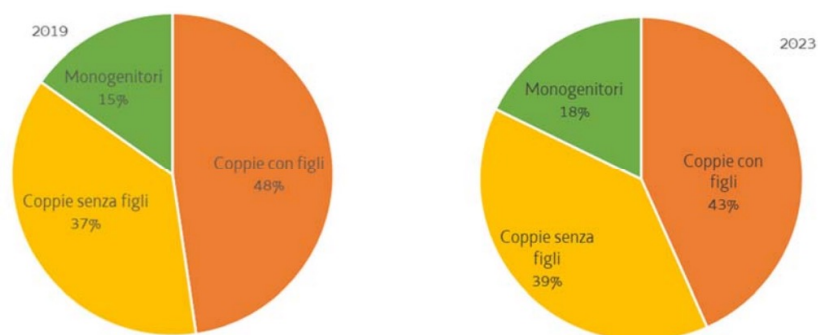
Analisi dei dati relativi alle caratteristiche della popolazione e all'evoluzione delle famiglie (Dicembre 2024)

SPPS: Servizio programmazione, pianificazione strategica, controllo di gestione e Statistica della Direzione generale della Regione Autonoma Friuli Venezia Giulia (Ufficio Statistica)



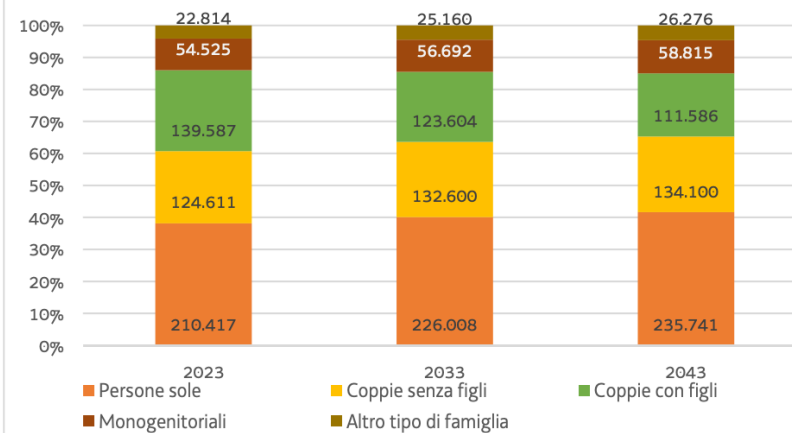
Esempio

Famiglie con un solo nucleo (valori %). Anni 2019 e 2023



Fonte: Istat, elaborazione a cura del SPPS RAFVG

Previsioni delle tipologie di famiglia (valori assoluti e %).
Anni 2023, 2033 e 2043



Fonte: Istat, elaborazione a cura del SPPS RAFVG

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.

E in Italia?

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

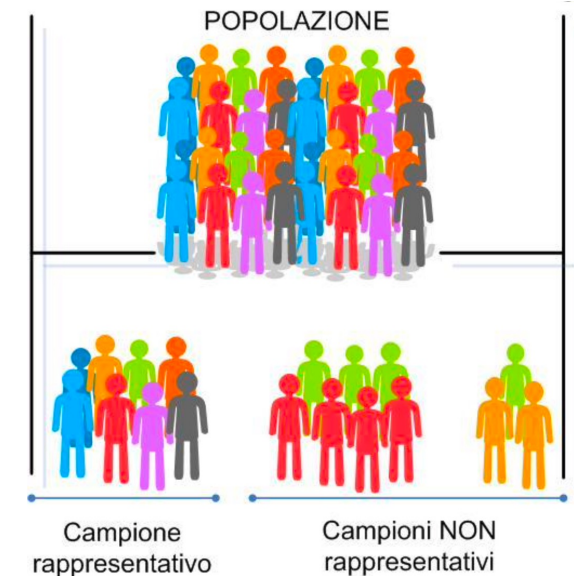
Campione casuale (semplice) di n soggetti estratti da una popolazione è tale se ogni possibile campione di pari numerosità ha uguale probabilità di selezione

- 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



Per casa (prossima lezione)

A. Consultare il sito dell'Istat <https://www.istat.it/it/archivio/91926>

farsi una idea delle caratteristiche dell'indagine:

- popolazione di interesse
- indagine campionaria o censimento
- argomenti/temi trattati
- modalità di raccolta dei dati
- ... (altre caratteristiche che vi paiono interessanti)