

# TERMINOLOGIA, NOTAZIONE, CONCETTI INTRODUTTIVI

## LEZIONE 2

### La Raccolta dei dati

#### (PROCEDIMENTO)

- Indagine statistica: studio che riguarda un collettivo di soggetti (es. una popolazione reale o fittizia)
  - l'indagine censuaria
  - prospettico
- L'esperimento
- Lo studio osservazionale
  - \* retrospettivo
  - \* prospettico

#### ESEMPIO: Censimento in Italia

informazioni esaustive / complete



permette una descrizione accurata di quanto rilevato

a meno di errori di rilevazioni, omissioni volontarie o no...

non c'è un costante aggiornamento dell'informazione



è però un procedimento costoso, lungo, condotto da persone poco esperte, su popolazioni sempre in movimento

#### ESEMPIO: Acque del bacino Maknassy: **CAMPIONAMENTO**

- Indagine per la valutazione delle acque sotterranee dei bacini idrografici e lo studio della mineralizzazione di queste acque

il sito di studio copre un'area di circa 1250 km nel sud della Tunisia

Non si possono analizzare TUTTE le acque della zona

si raccoglie una parte delle acque in alcuni punti

l'insieme delle rilevazioni si chiama **CAMPIONE STATISTICO**

**NB:** Il campione dei dati deve essere ben rappresentativo di tutte le acque / popolazione

**RISCHI:**

- selezionare pochi punti di rilevazione
- selezionare punti poco rappresentativi ⇒ darebbe un risultato fuorviante

- si fa una **MATRICE** di DATI (una tabella)

ci sono **TECNICHE** di **CAMPIONAMENTO**

## SCHEMI di CAMPIONAMENTO

- Campione casuale semplice
  - Campione casuale stratificato: si divide la popolazione in gruppi e dentro ogni gruppo si scelgono degli individui che faranno da campione
  - Campione casuale a grappolo: si suddivide la popolazione in maniera + grande (es. famiglie/suole, regioni, ...) e da queste se ne scelgono alcune a caso (insieme ai soggetti)
  - Campione casuale a due stadi: si fa un ulteriore campionamento di gruppi scelti nel campionamento a grappolo
- Si sceglie in base al tipo di indagine che si vuole fare

⊘ l'importante è che lo statistico scelga il **giusto campionamento di dati più rappresentativo!**

① • **Studio sperimentale**: si raccolgono i dati tramite uno sperimento. si fa un trattamento, e si osserva nel tempo

② • **Studio osservazionale**: c'è un'osservazione

PROSPETTICO

RETROSPETTIVO

↓  
Vengono selezionati i soggetti interessati dal fenomeno da indagare

↓  
È sempre uno studio osservazionale, NA: i soggetti vengono osservati in modo retrospettivo

↓  
Vengono seguiti nel tempo

↓  
si osservano dal passato

↓  
si raccolgono le informazioni

STATISTICAMENTE SIGNIFICATIVA: = se c'è un'associazione comprovata, che i nostri dati evidenziano, che possiamo considerare non frutto del caso ma è una testimonianza associativa testimoniata dai nostri dati.

- Molto + tempo
- Non è applicabile ai casi rari
- si ~~perde~~ riduce la possibilità di errore
- Segue da il paziente per molto tempo, quindi si può prevedere quello che succederà.

## "Tipi" di statistica

Una volta che i dati sono disponibili (dati da un determinato campione)  
l'analisi può procedere in due tipi:

### STATISTICA DESCRITTIVA

↓  
ci occuperemo di questo

**SCOPO:** descrivere il fenomeno per il quale sono stati raccolti i dati

individuare le relazioni tra due o più fenomeni

fa riferimento ai dati di un'intera popolazione (come il censimento)

Avviene prima della statistica inferenziale

### STATISTICA INFERENZIALE

↓

- cerca di descrivere le caratteristiche di un fenomeno in più in generale
- procede dal particolare per studiare il generale

↓  
**PROCEDIMENTO di INDUZIONE**

Le relazioni tra la popolazione e il campione si descrivono usando il **CALCOLO delle PROBABILITÀ**.

### Statistiche descrittive

- **OBIETTIVI**
  - introdurre la statistica quale disciplina per la raccolta e l'analisi dei dati
  - fornire gli strumenti tecnici di base utili per lo studio di uno o più fenomeni reali
  - acquisire le tecniche elementari dell'analisi empirica volta alla descrizione, sintesi e rappresentazione grafica dei dati
- **ATTENZIONE AD ALCUNI ASPETTI:**
  - accuratezza dei risultati
  - interpretazione critica dei risultati
  - qualche primo tentativo di previsione di un fenomeno al di fuori di quanto osservato
- **COME PROCEDEREMO:**
  - Teoria: concetti base
  - Esercizi
  - Attività di laboratorio
  - Indicazioni

# TERMINOLOGIA e ESEMPI

## VARIABILI e UNITÀ STATISTICA

- Il fenomeno di interesse che viene rilevato e studiato è comunemente chiamato **VARIABILE STATISTICA** o **CARATTERE STATISTICO**
- La statistica descrittiva considera che il fenomeno venga rilevato su una popolazione. Le componenti della popolazione sono i soggetti/le entità su ciascuna delle quali si misura/ci registra/si rileva il dato di interesse sono le **UNITÀ STATISTICHE** (oggetti su cui si fa le rilevazioni)

↓  
compongono la **POPOLAZIONE di RIFERIMENTO** per la rilevazione dei dati e l'analisi

**ATTENZIONE:** ad individuare l'unità statistica (≠ dalla variabile)

**MODALITÀ:** le diverse espressioni attraverso le quali una variabile si manifesta nella popolazione

↓  
Valori della variabile rilevate/rilevabili sulle unità statistiche

→ l'insieme delle modalità si chiamano **SUPPORTO** della **VARIABILE**

• In base alle modalità possiamo distinguere i tipi di variabili

### Variabili quantitative

**SUPPORTO**  
= l'insieme dei valori che la variabile può assumere

↓  
le modalità sono **VALORI NUMERICI**

Se i valori sono **NUMERI NATURALI**

↓  
la variabile è misurata su **SCALA DISCRETA**

es. membri di una famiglia

Se i valori sono **NUMERI REALI**

↓  
la variabile è misurata su **SCALA CONTINUA**

es. PIL/altezza

**NB:** a volte la variabile quantitativa si presenta con valori raccolti in classi

↓ (es. gruppo classe, ...)

### SCALA INTERVALLARE

### Variabili qualitative

↓  
le modalità sono **ESPRESSIONI VERBALI**

**SCALA SCONNESSA**  
o nominale

↓  
**NON** vi è ordine logico tra le modalità  
es. divisione in bianco/nero

**SCALA ORDINALE**  
o categoriale

↓  
vi è ordine logico tra le modalità

es. valutazione sufficiente/bravo



Dati i limiti di precisione delle misurazioni, le MOSAICITÀ possono apparire discrete, (MA) per lo statistico la natura della variabile NON cambia

la natura della variabile non va dedotta sull'osservazione, ma piuttosto sulla sua essenza

COSA SI FA?

- Si sceglie un livello di precisione legato all'ordine di grandezza del fenomeno da rilevare (es: statura  $\rightarrow$  cm / petali di fiori  $\rightarrow$  mm)
- Si accetta un implicito intervallo di tolleranza (invece di valori possibili in relazione a quella misurazione)