

# L'analisi statistica

# Classificazione

- Statistica descrittiva: rappresentazione mediante strumenti matematici di fenomeni reali
- Statistica inferenziale: induzione probabilistica circa la struttura incognita di una popolazione

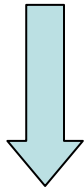
# Fasi dell'analisi statistica

1. Definizione degli obiettivi
2. Rilevazione (completa o parziale); generalmente effettuata su un campione della popolazione
3. Elaborazione dei dati: dati grezzi → dati derivati
4. Presentazione ed interpretazione dei dati
5. Applicazioni degli esiti dell'analisi

# Popolazione

Popolazione: non necessariamente esseri umani

Unità statistiche: componenti elementari della popolazione



1. Unità semplici: singola persona, abitazione,...
2. Unità composte: famiglia, edificio, ...
3. Unità complesse: rapporto di lavoro tra datore di lavoro e dipendenti

# Carattere statistico

Il carattere statistico consente di descrivere una popolazione

1. Carattere qualitativo: espresso mediante attributi
2. Carattere quantitativo: indicato mediante espressioni numeriche

# Tipi di variabile

- Continua: temperatura, statura, ...
- Discreta: componenti di una famiglia, ...

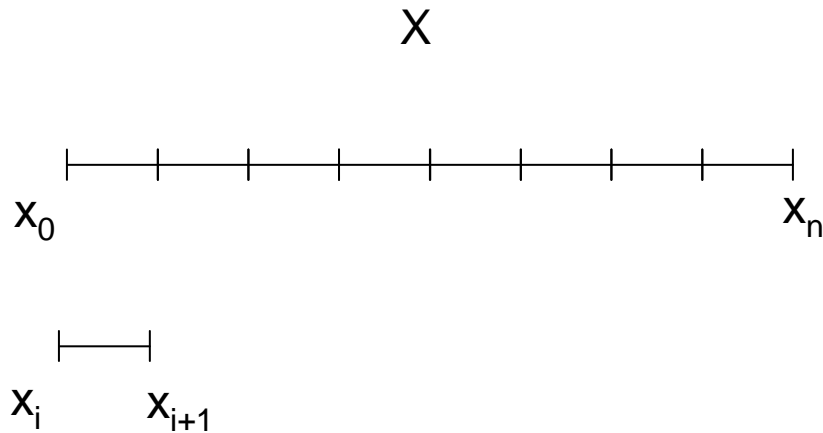
# Frequenze e intensità

- Frequenza assoluta: enumerazione  $n_i$
- Frequenza relativa:  $f_i = n_i/n$

$$0 \leq f_i \leq 1$$

# La classe di modalità

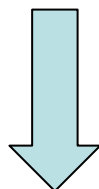
- Classe di modalità: intervallo di prefissata ampiezza in cui risulta suddiviso l'insieme delle modalità di un carattere quantitativo  $X$ .





# Valore centrale

Se l'insieme dei valori del carattere non è specificabile



La prima classe è aperta a sinistra, l'ultima è aperta a destra

Valore centrale:  $\bar{x}_i = \frac{x_i + x_{i+1}}{2}$

# Frequenze cumulate

Sia data una variabile statistica  $X$

$$x_1, x_2, \dots, x_i, \dots, x_{n-1}, x_n \quad x_1 \leq x_2 \leq \dots \leq x_i \leq \dots \leq x_{n-1} \leq x_n$$

Si definisce frequenza cumulata  $n_i'$  il numero di unità del carattere posseduto dalle prime  $i$  modalità

$$n_i' = \sum_{j=1}^i n_j$$

# Errori

Errore: differenza tra valore stimato ed il valore teorico

1. Errori sistematici: utilizzo di strumenti difettosi o modalità erranee di rilevazione
2. Errori casuali: fattori esterni.