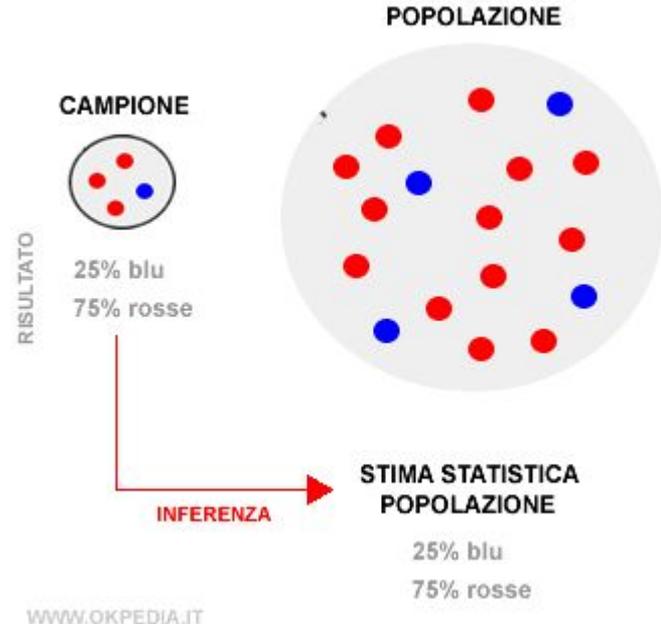
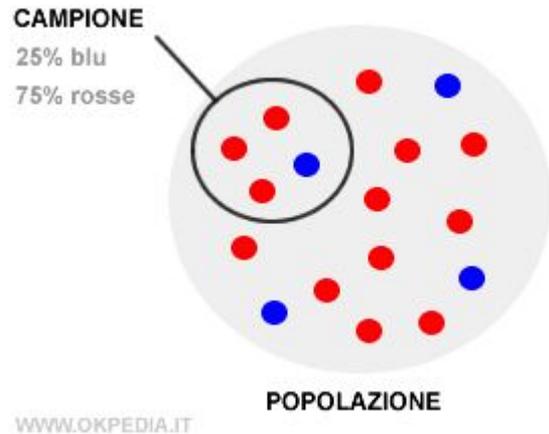


Analisi dei dati Statistica Inferenziale

Angelica Lo Duca
angelica.loduca@iit.cnr.it

Definizione*

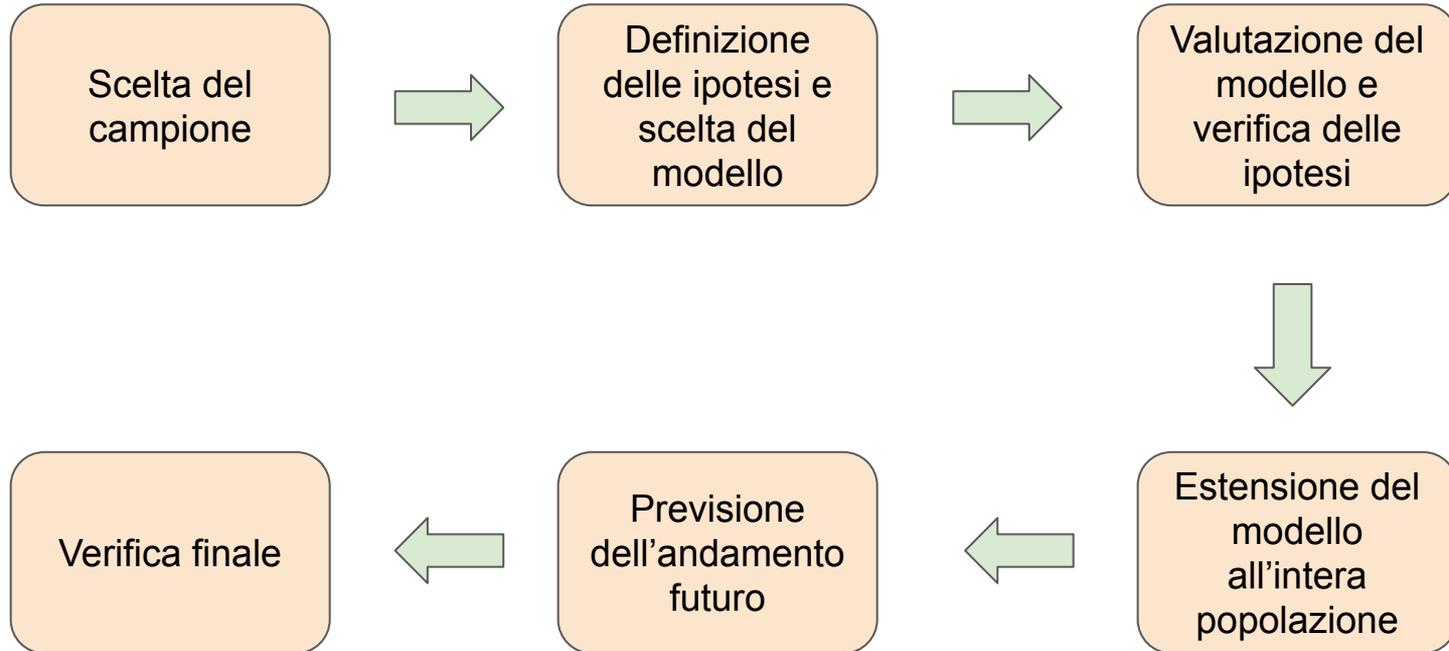
La Statistica Inferenziale si occupa di analizzare i dati ottenuti da un campione della popolazione per stimare un fenomeno statistico sull'intera popolazione di riferimento.



* Fonte: https://www.okpedia.it/statistica_inferenziale

Descriptive Statistics	Inferential Statistics
1. Works with a smaller data set	1. Works with a large data set
2. Process is simpler to do	2. Process is more complex as we have to decide on the best sampling techniques
3. Results obtained represent the entire data set	3. Results obtained represent a portion of the population, but can be used to deduce information about the entire population with some amount of uncertainty
4. Error involved is usually less	4. Error involved is usually more

Workflow



Scelta del campione

La scelta delle unità statistiche del campione può essere effettuata in modo casuale o seguendo un criterio oggettivo tale da garantire la rappresentatività del campione nei confronti dell'intera popolazione.

Essendo un'analisi condotta soltanto su una parte della popolazione (campione), i risultati di un'indagine statistica inferenziale hanno un margine di errore.

Definizione del modello

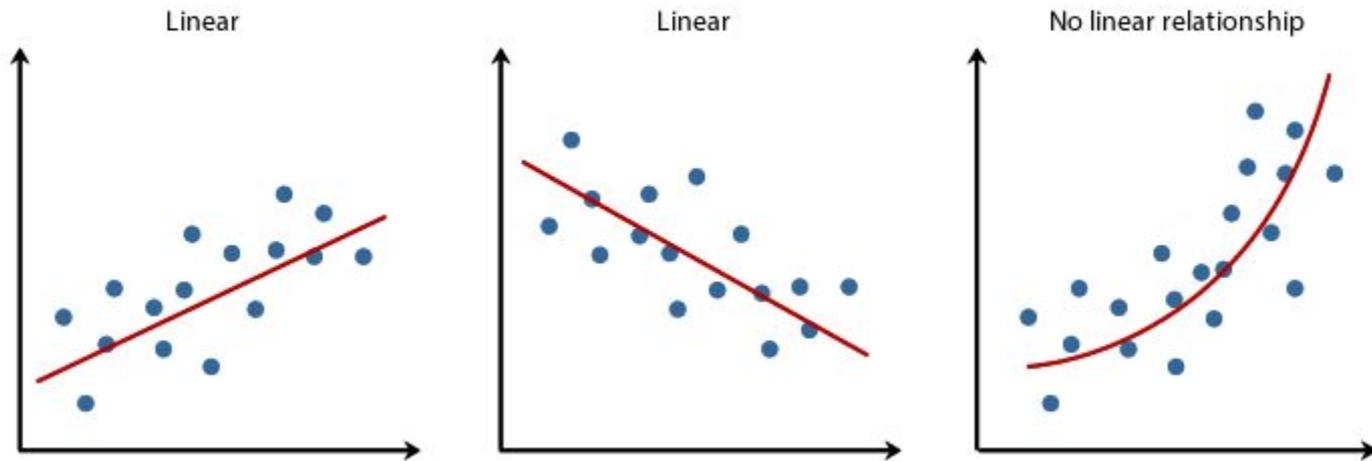
Grafico preliminare del campione per capire se esiste un modello descrittivo

Uso di modelli standard

- Regressione lineare
- Distribuzione normale
- ...

Adattamento dei modelli standard al caso particolare

Regressione Lineare



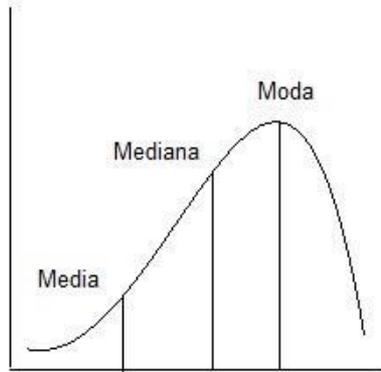
$$y = ax + b$$

L'obiettivo è determinare i coefficienti a e b

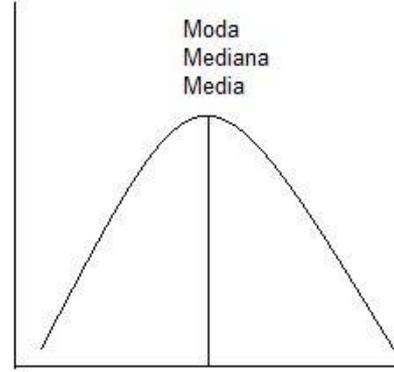
Distribuzione normale

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

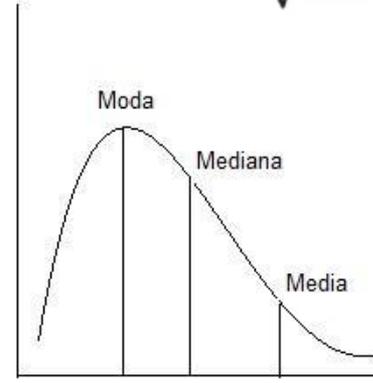
σ = devianza standard
 μ = media
 π = 3,14159
 e = 2,71828



Asimmetrica a sinistra



Distribuzione normale
(Simmetrica)



Asimmetrica a destra

Le probabilità più elevate di un fenomeno si concentrano intorno alla media. Le probabilità si riducono man mano che ci si allontana dal valore medio a destra o a sinistra.

<https://www.okpedia.it/asimmetria-statistica>

Valutazione del modello

Identificare l'errore tra il modello e i dati reali.

Esistono diverse metriche per la valutazione del modello

<https://www.datavedas.com/model-evaluation-regression-models/>

Riferimenti Bibliografici

D.R. Cox. [Principles of Statistical Inference](#)