

---

## Metodi statistici a supporto delle decisioni

### Finalità

La prima parte costituisce un solido approfondimento delle conoscenze introdotte nel corso di Statistica. Vengono introdotti nuovi tipi di test e affrontate problematiche più complesse e realistiche. La seconda parte è una panoramica delle tecniche gestionali e di analisi moderne che possono essere ricondotte alla statistica.

### Programma

PRIMA PARTE: STRUMENTI MULTIVARIATI DI BASE.

Richiami sulle variabili aleatorie e sull'inferenza statistica.

Classici test Z, T e F per il confronto dei parametri di due popolazioni normali.

Test di adattamento e di indipendenza (Fisher-Irwin, chi-quadro, tabelle di contingenza).

Regressione: determinazione dei coefficienti (modelli lineari, multilineari, linearizzabili; coefficiente di determinazione, analisi dei residui, minimi quadrati pesati); inferenza sui coefficienti (test T e F).

Analisi della varianza (a una via, a due vie e con interazioni).

SECONDA PARTE: TECNICHE DI EXPLORATIVE DATA ANALYSIS.

Rappresentazione di dati con elevata numerosità e dimensionalità (distribuzioni normali multivariate, matrice di correlazione, autovalori e autovettori).

Verifica del modello (funzioni kernel, test chi-quadro e di Kolmogorov-Smirnov).

Cluster analysis (distanze; hierarchical tree clustering, linkage; algoritmi k-means; algoritmi EM, misture, classificazioni bayesiane).

Factor analysis (principal component analysis, common factor analysis, riduzione delle variabili, interpretazione dei fattori, rotazioni dei fattori).

Discriminant function analysis (metodo lineare di Fisher, riduzione delle variabili).

Neural networks (multilayer perceptron).

Overfitting e overlearning: quando il modello interpreta il campione e non la popolazione.

Test non parametrici (dei segni, dei ranghi con segno, di Wilcoxon, per l'indipendenza dei dati).

Test parametrici bayesiani (cenni).

### Attività d'esercitazione

Per la sola prima parte: le esercitazioni si svolgono al calcolatore e illustrano l'utilizzo di un foglio elettronico per la risoluzione di esercizi selezionati e per chiarire meglio alcuni concetti teorici.

### Modalità d'esame

Per la prima parte: esame orale al calcolatore. Per la seconda parte: lavoro di approfondimento teorico oppure analisi su un set di dati. A seconda della complessità, tale lavoro potrà essere affrontato a gruppi.

### Propedeuticità

Statistica, Analisi AB, Analisi C.

### Testi consigliati

S. Ross - Probabilità e statistica per l'ingegneria e le scienze - Apogeo 2003

Hand, Mannila, Smyth - Principles of data mining.