

# UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

# **SYLLABUS DEL CORSO**

# Statistica Multivariata

2526-1-F8206B002-F8206B002-1

#### Obiettivi formativi

Il corso si pone come obiettivo quello di introdurre tecniche di analisi statistica multivatiata con finalità esplorativa e previsiva. Alla fine del corso lo studente sarà in grado di applicare le suddette tecniche a insiemi di dati raccolti in molteplici contesti, quali quello economico, sociale e delle scienze naturali, e che possono essere caratterizzati, per esempio, da elevata dimensionalità.

Il corso contribuisce a consolidare le conoscenze nell'ambito dei modelli probabilstici e della metodologia statistica e la capacità di applicare tali conoscenze, in coerenza con l'area di apprendimento di Statistica del corso di laurea magistrale in Scienze Statistiche ed Economiche.

## Contenuti sintetici

- · Distribuzione normale multivariata in statistica
- Metodi lineari per la classificazione
- Modelli per variabili categoriali
- Modelli per variabili ordinali
- Modelli per variabili a coppie o in cluster

## Programma esteso

#### Distribuzione normale multivariata in statistica

- definizione
- distanza di Mahalanobis

- proprietà
- stima di massima verosimiglianza
- distribuzione degli stimatori di massima verosimiglianza
- distribuzione Wishart
- strumenti per la verifica dell'assunzione di normalità multivariata
- · trasformazione di Box-Cox multivariata

#### Metodi lineari per la classificazione

- · classificazione e teoria delle decisioni
- · classificatore di Baves
- approccio naif: regressione lineare per la classificazione
- · funzioni discriminanti
- analisi discriminante lineare
- analogie tra analisi discriminante lineare e regressione lineare
- analisi discriminante quadratica
- approccio di Fisher all'analisi discriminante lineare
- metodo "naive Bayes" come caso speciale dell'analisi discriminante

#### Modelli per variabili categoriali

- · regressione logistica
- stima di massima verosimiglianza via Newton-Raphson
- · regressione logistica multinomiale
- · classificazione tramite regressione logistica
- parallelo tra analisi discriminante lineare e regressione logistica
- modelli alternativi per variabili binarie: modello probit, modello log-log complementare

#### Modelli per variabili ordinali

- modello logistico cumulato
- modelli alternativi per variabili ordinali: modello probit cumulato, modello log-log complementare cumulato
- modello proportional hazard di Cox
- modello logistico a categorie attigue

#### Modelli per variabili a coppie o in cluster

- approccio marginale e approccio condizionale
- modello logistico condizionale
- · verosimiglianza condizionale
- modelli a effetti casuali
- · modello di Rasch
- modello item-response
- modello logistico-normale
- proprietà condizionali e marginali del modello logistico-normale
- possibili generalizzazioni
- modelli a effetti casuali con distribuzione nonparametrica

# Prerequisiti

Analisi esplorativa, modelli statistici, calcolo delle probabilità, inferenza statistica, programmazione.

#### Metodi didattici

Lezioni frontali (teoria e esercizi) e sessioni di laboratorio.

Nello specifico, sono previste:

- 7 lezioni frontali da 3 ore (modalità erogativa nella prima parte e in modo interattivo nella parte successiva)
- 7 lezioni frontali da 2 ore (modalità erogativa)
- 6 attività di laboratorio da 2 (modalità interattiva)

E' previsto che l'80% siano erogate in presenza e il 20% da remoto.

## Modalità di verifica dell'apprendimento

L'esame è una prova scritta che si compone di domande di teoria, esercizi teorici e numerici da svolgere tramite l'utilizzo di R.

Non sono previste prove in itinere.

#### Testi di riferimento

Hastie, T., Tibshirani, R., Friedman, J.H. and Friedman, J.H., 2009. The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

Agresti, A., 2012. Categorical data analysis (Vol. 792). John Wiley & Sons.

## Periodo di erogazione dell'insegnamento

Secondo semestre. Primo ciclo.

## Lingua di insegnamento

Italiano

## **Sustainable Development Goals**

ISTRUZIONE DI QUALITÁ