



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Probabilità e Statistica per l'Informatica

2627-2-E3102Q114

Obiettivi

Obiettivi formativi (Descrittori di Dublino)

Al termine del corso, lo studente avrà raggiunto i seguenti risultati di apprendimento:

- 1. Conoscenza e capacità di comprensione.** Lo studente acquisirà le conoscenze di base nel campo del calcolo delle probabilità e della statistica matematica, comprendendone linguaggio, definizioni e principali risultati. Particolare attenzione sarà dedicata alle distribuzioni di probabilità più comuni e ai fondamenti della statistica descrittiva e inferenziale.
- 2. Capacità di applicare conoscenza e comprensione.** Lo studente sarà in grado di applicare i concetti teorici per analizzare fenomeni aleatori, condurre ragionamenti in condizioni di incertezza e descrivere sinteticamente insiemi di dati. Sarà inoltre in grado di stimare parametri e verificare ipotesi statistiche in contesti applicativi.
- 3. Autonomia di giudizio.** Lo studente svilupperà la capacità di valutare criticamente i modelli probabilistici e statistici utilizzati, scegliendo metodi adeguati per l'analisi e l'interpretazione dei dati.
- 4. Abilità comunicative.** Lo studente sarà in grado di comunicare in modo chiaro e rigoroso concetti e risultati relativi alla probabilità e alla statistica, sia in forma scritta che orale, utilizzando un linguaggio appropriato al contesto informatico.
- 5. Capacità di apprendimento.** Lo studente acquisirà le competenze necessarie per approfondire autonomamente lo studio della probabilità e della statistica, anche attraverso l'uso di strumenti software, con particolare riferimento al linguaggio di programmazione R.

Contenuti sintetici

La prima parte del corso, dedicata al **calcolo delle probabilità**, si apre con la **statistica descrittiva**, che permette di sintetizzare le informazioni salienti relative a un insieme di dati. Vengono quindi presentati gli **spazi di probabilità**, che sono alla base della descrizione matematica dei fenomeni aleatori, e le **variabili aleatorie**, che ne costituiscono il "linguaggio operativo".

La seconda parte del corso, dedicata alla **statistica matematica**, si apre con i **teoremi di convergenza** (legge dei grandi numeri e teorema limite centrale) che descrivono il comportamento di fenomeni aleatori con un grande numero di componenti. Vengono presentate le tecniche principali per la **stima di parametri** in un modello probabilistico e viene discussa la **verifica di ipotesi** e i corrispondenti test statistici, sia relativi a parametri incogniti in un modello aleatorio (*test parametrici*) sia relativi al confronto di distribuzioni (*test non parametrici*). Il corso si conclude con la **regressione lineare**, una tecnica importante per studiare il legame che può sussistere tra una variabile (output) e un insieme di variabili (input) in condizioni di incertezza.

Programma esteso

1. Statistica descrittiva

- Introduzione all'analisi dei dati
- Statistiche campionarie (media, mediana, quantili, varianza, correlazione)
- Rappresentazioni grafiche

2. Spazi di probabilità

- Fenomeni aleatori, spazi di probabilità ed eventi
- Proprietà di base della probabilità
- Probabilità condizionata
- Elementi di calcolo combinatorio
- Indipendenza di eventi

3. Variabili aleatorie

- Variabili aleatorie discrete
- Valore medio, momenti, varianza e covarianza
- Variabili aleatorie assolutamente continue
- Distribuzioni notevoli discrete e assolutamente continue
- Variabili aleatorie normali

4. Teoremi di convergenza

- Convergenza di variabili aleatorie e distribuzioni (cenni)
- Legge dei grandi numeri
- Teorema limite centrale

5. Stima di parametri

- Campioni e statistiche
- Stimatori (media e varianza campionarie)
- Intervalli di confidenza

6. Verifica di ipotesi

- Test per la verifica di un'ipotesi, errori di I e II specie

- Test parametrici per media e varianza
- Test non parametrici di buon adattamento e di indipendenza

7. Regressione lineare

- Introduzione alla regressione
- Inferenza statistica sui parametri
- Analisi dei residui

Prerequisiti

Le conoscenze, competenze e abilità impartite negli insegnamenti precedenti di matematica, in particolare di Analisi Matematica.

Modalità didattica

Il corso viene erogato in lingua Italiana in modalità *blended-learning*.

Le lezioni frontali sono articolate nel modo seguente:

- 16 lezioni teoriche da 2 ore svolte in modalità erogativa in presenza, in cui si fornisce la conoscenza di definizioni, risultati, dimostrazioni ed esempi rilevanti;
- 10 esercitazioni da 2 ore svolte in modalità erogativa in presenza, in cui si forniscono competenze e abilità necessarie per utilizzare le nozioni teoriche per la risoluzione di esercizi.

La parte *blended* del corso (laboratorio con R) consiste in esercitazioni con il software R svolte in modalità interattiva da remoto, in cui si acquisiscono le competenze e abilità necessarie per effettuare analisi statistiche con l'uso di R. Le esercitazioni sono organizzate in sette sessioni distinte (una per ciascun argomento del corso) che prevedono un impegno di circa 3 h 30 min a sessione.

Materiale didattico

Testo di riferimento:

- S. M. Ross (2014), Introduzione alla Statistica, seconda edizione, Apogeo Editore

Inoltre, viene reso disponibile il seguente materiale didattico:

- Appunti dei docenti
- Fogli di esercizi

Laboratorio con R

- video lezioni e appunti

Periodo di erogazione dell'insegnamento

Secondo semestre

Modalità di verifica del profitto e valutazione

L'esame è costituito da una prova scritta e da una eventuale prova orale e riceve un voto in trentesimi.

La prova scritta è costituita da due parti:

- una prima parte con domande a risposta chiusa, che contribuisce per un terzo al punteggio totale, in cui vengono valutate le conoscenze teoriche;
- una seconda parte con esercizi a risposta aperta, che contribuiscono per due terzi al punteggio totale, in cui vengono valutate le abilità pratiche.

È prevista una prova parziale a metà del corso. Il primo appello d'esame coincide con la seconda prova parziale ed è riservato a chi abbia superato la prima.

La prova orale è facoltativa (su richiesta dello studente e/o del docente) e può contribuire sia in maniera positiva che in maniera negativa al voto finale.

È previsto un progetto con *il software "R"*, da svolgere secondo le modalità descritte all'inizio del corso, dal valore di 2 punti.

Orario di ricevimento

Su appuntamento

Sustainable Development Goals

ISTRUZIONE DI QUALITÀ
