UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Models and Applications in Economics and Environmental Sciences

2627-2-E3304M015

Obiettivi formativi

Il corso fornisce una conoscenza dell'inferenza statistica e del modello lineare dal punto di vista teorico e mediante l'utilizzo del software R, bilanciando i principi teorici con gli aspetti pratici dell'analisi dei dati. Il corso introdurrà la stima puntuale, gli intervalli di confidenza e i test d'ipotesi, per poi approfondire il modello lineare, sia semplice che multiplo, applicando questi metodi a dati reali attraverso casi studio.

Contenuti sintetici

Inferenza statistica e modellizzazione

- Stima puntuale: popolazioni e campioni, distribuzioni campionarie e Teorema del Limite Centrale (TLC), stimatori puntuali.
- Intervalli di confidenza per medie e proporzioni.
- Verifica di ipotesi: elementi di un test di significatività, statistiche test e p-value, test per medie e proporzioni.
- Modelli lineari: regressione lineare semplice e multipla, strumenti diagnostici, inferenza nei modelli lineari normali, predittori categorici e analisi della varianza (ANOVA).

Programmazione e analisi dei dati con R

- Fondamenti di R, gestione dei dati e statistica descrittiva. Inferenza statistica e distribuzioni campionarie. Stima, interpretazione e utilizzo dei modelli lineari. Casi di studio in ambito ecologico.

Programma esteso

Inferenza statistica e modellizzazione

- Stima puntuale
    - Concetto di popolazione e campione
    - Distribuzioni campionarie della media campionaria, Teorema del Limite Centrale (richiami)
    - Stima puntuale
- Intervalli di confidenza
    - Intervalli di confidenza per medie e proporzioni
- Verifica delle ipotesi
    - Elementi di un test di significatività: assunzioni e ipotesi
    - Statistiche test e *p-value*
    - Test di significatività per medie e proporzioni
- Modelli lineari
    - Regressione lineare semplice e stima mediante il metodo dei minimi quadrati
    - Regressione lineare multipla
    - Strumenti diagnostici
    - Inferenza statistica nei modelli lineari normali
    - Variabili esplicative categoriche e analisi della varianza (ANOVA)

Programmazione e analisi dei dati con R

- Introduzione a R: fondamenti di programmazione, gestione dei dati e statistica descrittiva
- Inferenza statistica e distribuzioni campionarie: variabili casuali, stima puntuale, test d'ipotesi e intervalli di confidenza per i principali modelli probabilistici
- Modelli lineari e metodo dei minimi quadrati: stima dei modelli e interpretazione dei risultati
- Analisi dei dati: casi di studio in ecologia, economia e scienze ambientali

Prerequisiti

Si presuppone la conoscenza degli argomenti trattati nei corsi di Matematica e Statistica, in particolare analisi reale, statistica descrittiva e calcolo delle probabilità.

Metodi didattici

Le lezioni si svolgono sia in aula che in laboratorio, integrando i principi teorici con gli aspetti pratici dell'analisi dei dati e della programmazione in R. Le 78 ore di didattica sono così articolate:

- 48 ore di lezioni teoriche e pratiche (con esercitazioni), svolte in presenza. Alcune lezioni potranno essere erogate a distanza (da remoto) fino ad un massimo del 20% delle ore.
- 30 ore di attività di laboratorio, svolte in modalità interattiva e a distanza.

Modalità di verifica dell'apprendimento

L'esame consiste in una prova scritta, svolta in laboratorio, che comprende quesiti teorici, esercizi e prove pratiche

in R relativi a entrambe le parti del corso: inferenza statistica e modelli lineari, nonché alla loro implementazione in R. Per rispondere ad alcune delle domande, gli studenti dovranno scrivere ed eseguire codice in R.

Testi di riferimento

- Agresti, A. and Kateri, M. (2022), Foundations of Statistics for Data Scientists, CRC press.
- Le dispense dei docenti verranno rese disponibili.

Testo opzionale

- Albert, J. & M. Rizzo (2012). R by Example. Springer.

Periodo di erogazione dell'insegnamento

Primo semestre

Lingua di insegnamento

Inglese

Sustainable Development Goals

ISTRUZIONE DI QUALITÁ