



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Statistical Models and Applications in Economics and Environmental Sciences

2627-2-E3304M015

Learning objectives

The course provides training in statistical inference and linear modelling using R, balancing theoretical principles with the practical aspects of data analysis. It covers point estimation, confidence intervals and hypothesis testing, then develops the linear model - simple and multiple - applying these methods to real data through case studies.

Contents

Statistical Inference and Modeling

- Point estimation: population and samples, sampling distributions and the Central Limit Theorem (CTL), point estimates.
- Confidence intervals for means and proportions
- Hypothesis testing: the elements of a significance test, test statistics and p-values, tests for means and proportions.
- Linear models: simple and multiple regression, diagnostic tools, inference for Normal linear models, categorical predictors and ANOVA.

Programming and Data Analysis with R

- R basics, data handling and descriptive statistics; inference and sampling distributions; fitting and interpreting linear models; ecological case studies.

Detailed program

Statistical Inference and Modeling

- Point estimation
 - Population and samples (recap)
 - Sampling distributions of sample means, Central Limit Theorem (recap)
 - Point estimation and the maximum likelihood method
- Confidence intervals:
 - Confidence intervals for means and proportions
 - Comparing two population means or proportions
- Hypothesis testing
 - The elements of a significance test: assumptions and hypotheses
 - Test statistics and p-values
 - Significance tests for means and proportions
- Linear models
 - Simple linear regression and its least squares fit
 - Multiple linear regression
 - Diagnostic tools
 - Statistical inference for Normal linear models
 - Categorical explanatory variables and ANalysis Of VAriance (ANOVA)

Programming and Data Analysis with R

- Introduction to R: basic programming, data handling and descriptive statistics
- Inference and sampling distributions: random variables, point estimation, testing and confidence intervals for notable models
- Linear models and least squares: model fitting and interpretation of the results
- Data analysis: case studies in ecology, economics, and environmental sciences

Prerequisites

Students are expected to have knowledge of the topics covered in the Mathematics and Statistics courses, in particular real analysis, descriptive statistics, and probability.

Teaching methods

Lessons are held both in the classroom and in the lab, integrating theoretical principles with the practical aspects of data analysis and programming in R. The 72 hours of teaching are organized as follows:

- 42 hours of theoretical and practical (exercises) lectures, in person. Some lectures may be conducted remotely, for up to 20% of the total hours.
- 30 hours of laboratory activities, conducted interactively and remotely.

Assessment methods

The exam is a written examination held in the lab, combining theoretical questions, exercises and exercises in R covering both parts of the course - statistical inference and linear models, together with their implementation in R. Students will be required to execute R code in order to answer some of the questions.

Textbooks and Reading Materials

- Agresti, A. and Kateri, M. (2022), Foundations of Statistics for Data Scientists, CRC press.
- Additional teaching material will be made available.

Optional textbook

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

Semester

First semester

Teaching language

English

Sustainable Development Goals

QUALITY EDUCATION
