UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Advanced Foundations of Statistics for AI

2425-1-F9102Q002

Aims

The aim of this course is to provide the theoretical foundations of Statistics and indicate how the theory sheds light on the properties of practical methods used in AI. The topics include estimation, prediction, testing, confidence sets, and the general approach to statistical modeling and learning, with some exposure to robust statistical inference.

Contents

The course consists of a theoretical part, with applications on data using the R software environment. Statistical Inference and statistical modeling are presented, with a view on modern methods for supervised and unsupervised learning widely employed in AI.

Detailed program

- Crash course: Quick review of basic probability theory, and on Statistical inference to recall Interval estimation and Hypothesis testing.
- Simple linear regression: least squares estimation method
- Multiple linear regression: least squares estimation method, model adequacy measures, sampling distribution of OLS estimators, hypothesis tests and confidence intervals for the regression coefficients, analysis of variance, outliers and influential observations, robust linear regression.
- Linear models for regression and for classification
- Lasso and Ridge regression, the selection of the model
- Robust statistical modeling
- Unsupervised learning: model-based clustering, latent variables, PCA and factor analysis.

Prerequisites

Statistics background: basis of probability theory and knowledge of the most relevant continuous and discrete random variables.
Mathematics background: linear algebra, matrix theory and advanced calculus.

Teaching form

40 hourd of frontal lessons and 24 hours of interactive exercise lessons.
Attendance to lectures and interactive exercises is highly recommended.

Textbook and teaching resource

- Devore, J. L. (2011). Probability and Statistics for Engineering and the Sciences. Cengage learning.
- James G., Witten D., Hastie T. and Tibshirani R. (2021). An Introduction to Statistical Learning, with applications in R (2nd edition). Springer Verlag.
- Hastie T. , Tibshirani R., Friedman J. (2021). The Elements of Statistical Learning (2nd edition). Springer Verlag.

Semester

First

Assessment method

The exam consists in a written exam with three open-ended questions to assess the student's knowledge and understanding of the subject, (1h30m) jointly with the development of an application in R on real data (1h). The exam is closed book. An additional oral discussion can be requested by the lecturer.

Office hours

Office: University of Milan-Bicocca,
Department of Statistics and Quantitative Methods, Via Bicocca degli Arcimboldi, 8
20125 MILANO
Building U7-Civitas, Floor 2, room 2063
Phone: +39-02-6448-3118
just drop me an email (francesca.greselin@unimib.it) and we will agree on how and when to meet

Sustainable Development Goals

QUALITY EDUCATION | GENDER EQUALITY | SUSTAINABLE CITIES AND COMMUNITIES