Syllabus

Course description

The course will provide the opportunity to tackle real world problems requiring advanced computational skills and visualisation techniques to complement statistical thinking. Students will practice proposing efficient solutions, and effectively communicating the results with stakeholders.

Content

- Modern statistical computing environments (e.g., R, Julia, and Python)
- Aids to efficiency and reproducibility (e.g., GitHub, Markdown)
- Data management, wrangling, and ethics
- Statistical graphics (grammar, good practices, applications, and examples)
- Kernel density estimation and smoothing
- Cross-validation
- EM algorithm and applications
- Resampling methods for uncertainty assessment (bootstrap, jackknife, cross-validation), with applications to regression, time series and dependent data
- Monte Carlo methods for sampling and numerical integration
- Introduction to Bayesian inference
- Markov chain Monte Carlo techniques (Gibbs sampler, Metropolis-Hastings algorithm, Hamiltonian Monte Carlo, convergence diagnostics)
- Decision trees for classification or Conformal prediction

Prerequisites

Required courses: Probability and statistics, Linear models

Learning outcomes

By the end of the course, the student must be able to:

- Plan complex visualisation and computational tasks
- Perform complex visualisation and computational tasks
- Implement reproducible computational solutions to statistical problems in modern environments and platforms
- Expound the main approaches used for problem solving

Transversal skills

- Take feedback (critique) and respond in an appropriate manner
- Demonstrate the capacity for critical thinking
- Identify the different roles that are involved in well-functioning teams and assume different roles, including leadership roles
- Write a scientific or technical report