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, Rstudio and Python)
- Overview of other software (e.g., MATLAB, Julia)
- Aids to efficiency and reproducibility (e.g., GitHub, Markdown, Jupyter)
- Data management, wrangling, and ethics
- Statistical graphics (grammar, good practices, applications, and examples)
- EM algorithm and applications
- Kernel density estimation and smoothing
- Resampling methods for uncertainty assessment (bootstrap, jackknife, cross-validation), with applications to regression, time series and dependent data
- Markov chain Monte Carlo techniques (Gibbs sampler, Metropolis-Hastings algorithm, Hamiltonian Monte Carlo, convergence diagnostics) and software (e.g., Stan)
- Other methods for Bayesian inference (e.g., importance sampling, INLA, AGHQ, …)

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

Transversal skills

- Take feedback (critique) and respond in an appropriate manner.
- Communicate effectively with professionals from other disciplines.
- Demonstrate the capacity for critical thinking
- Identify the different roles that are involved in well-functioning teams and assume different roles, including leadership roles.