MATH-517

Statistical computation and visualisation

Mhalla Linda

Cursus	Sem.	Type
Ing.-math	MA1, MA3	Opt.
Mathématicien	MA1, MA3	Opt.
Statistics	MA1, MA3	Obl.

Language of teaching	English
Credits	5
Withdrawal	Unauthorized
Session	Winter
Semester	Fall
Exam	During the semester
Workload	150h
Weeks	14
Hours	**4 weekly**
Courses	2 weekly
Exercises	2 weekly
Number of positions	**45**

It is not allowed to withdraw from this subject after the registration deadline.

Summary

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)

• Aids to efficiency and reproducibility (e.g., GitHub, Markdown, Jupyter)

•
Data management, wrangling, and ethics

• Statistical graphics (grammar, good practices, applications, and examples)
• Kernel density estimation and smoothing

• 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) and software (e.g., Stan)

Keywords

Bayesian inference, Data visualisation, Data wrangling, EM algorithm, MCMC, Resampling methods, Statistical computation.

Learning 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.

Teaching methods

Two lecture hours per week, two hours of exercises and support on mini-projects and assignments

Expected student activities

Students will work on individual assignments and mini-projects in teams

Assessment methods

Contrôle continue

Supervision

Office hours	No
Assistants	Yes
Forum	No

Resources

Bibliography

Wickham H. & Grolemund G. (2017) R for Data Science
Bootstrap Methods and their Application
An Introduction to Statistical Learning

Ressources en bibliothèque

- Bootstrap Methods and their Application / Davison

- R for Data Science / Wickham
- An Introduction to Statistical Learning / Gareth

Moodle Link

- https://go.epfl.ch/MATH-517

Prerequisite for

Applied Statistics (MATH-516)