2EL1750 – Advanced statistics

Instructors: Sarah Lemler
Department: DÉPARTEMENT MATHÉMATIQUES
Language of instruction: FRANCAIS
Campus: CAMPUS DE PARIS - SACLAY
Workload (HEE): 60
On-site hours (HPE): 35,00
Elective Category : Fundamental Sciences
Advanced level : Yes

Description
The Advanced Statistics course aims to present various statistical methods for the estimation, the prediction and for determining the properties of the proposed estimators. We will develop two themes:

- multivariate linear regression
- non-parametric statistics

The course covers both the theoretical aspects of these concepts but also proposes a practical application of the models and methods considered using TPs with R software on data sets from different domains.

To be downloaded before the first TP:

- the R software https://www.r-project.org/
- the RStudio interface https://www.rstudio.com/

Quarter number
SG8

Prerequisites (in terms of CS courses)

For this course the prerequisites are as follows:

You must be familiar with the concepts seen in Statistics and Learning in 1st year, in particular:

- estimators,
- confidence intervals,
- the tests,
- the first notions on the linear model

Syllabus

We will develop two themes:

- multivariate linear regression,
- non-parametric statistics.

Class components (lecture, labs, etc.)

The course will be done on the board (for the most theoretical part) or from projected slides when it will be necessary, for example, to present an application made from the R software.

There are 35 hours scheduled for this course, including 2 hours for the final exam, about 15 hours of TP/TD (can be modulated) and 18 hours of classes.

Grading

The course consists of two evaluations: a homework assignment (DM) to be written in pairs to be returned halfway through the course (the date will be specified during the first class) a 2-hour final exam (EX) covering the entire course. The score of this course will then be the average of the scores of the two previous evaluations ((1/3)*DM+(2/3)*EX) rounded to the nearest half point. Failing students can re-take the exam in a second session EX2 in the same way as the first session exam EX. The course's score after the second session is the one of the exam EX2 (without the DM).

Course support, bibliography

The 1st year course material of Statistics and learning written by Paul-Henry Cournède

Resources

At the end of each course, TDs or TPs exercises will be proposed to familiarize themselves with the theoretical concepts seen in class and put them into practice to answer concrete problems, possibly based on data sets.

Learning outcomes covered on the course

- Use of parametric and non-parametric statistical estimation techniques
- Validate a model and understand the limits of a statistical model
- Propose, implement and calibrate a predictive model
- Use the R software and interpret results

Description of the skills acquired at the end of the course

- Use of parametric and non-parametric statistical estimation techniques
- Validate a model and understand the limits of a statistical model
- Propose, implement and calibrate a predictive model
- Use the R software and interpret results