STATISTIQUE APPLIQUÉE AUX SCIENCES DE L'INGÉNIEUR

STATISTICAL ENGINEERING

Lecturers: **Céline HARTWEG-HELBERT, Alexandre SAIDI**

| Lecturers : 16.0 | TC : 0.0 | PW : 0.0 | Autonomy : 0.0 | Study : 12.0 | Project : 0.0 | Language : FR

Objectives

The objective of this course is to provide the classic tools of mathematical statistics which includes the choice of the probabilistic model, its estimation and its evaluation. We will be particularly interested in the linear model and its extensions in the context of high-dimensional statistical learning (LASSO, RIDGE, PCR PLS), the logistic model and tree-based models (CART, RF, Boosting etc.). The aim of this course is also to provide training in the manipulation of data and the practical implementation of the studied models. For this, a substantial part of the course is oriented towards the implementation of the different models using the R software through the study of a large number of examples.

Keywords : Linear and logistic regression. Model selection. Design of experiments. L1 L2 Penalized regression. Regression trees.

Programme	1) Linear regression. Validities and limitations of the method. Model selection. 2) Design of experiments: screening and response surface 3) Logistic regression 4) Elements of statistical learning in high dimension 5) Clustering PRACTICAL ACTIVITIES The three activities will be devoted to learning the techniques of regression models on the R software. Numerous data sets will be studied.
Learning outcomes	• Know how to recognize different classes of statistical learning problems. • Know how to implement basic models of statistical learning and validate their relevance. • Know how to propose learning methods adapted to the high dimension • Know how to use R.
Independent study	Objectifs : This activity is not concerned with framed autonomy activities outside personal work. Méhodes : This activity is not concerned with framed autonomy activities outside personal work.
Core texts	G. Saporta, *PROBABILITÉS, ANALYSE DES DONNÉES ET STATISTIQUE*, Technip, 2020 T Hastie, R Tibshirani, J Friedman *THE ELEMENTS OF STATISTICAL LEARNING: DATA MINING, INFERENCE, AND PREDICTION*, Springer, 2009
Assessment	Final mark = 60% Knowledge + 40% Know-how Know-how = 100% continuous assessment Knowledge= 100% final exam

| Lecturers : 16.0 | TC : 0.0 | PW : 0.0 | Autonomy : 0.0 | Study : 12.0 | Project : 0.0 | Language : FR