

## Seminar of Applied Statistics

### Objectives

Integrating the practice and theory of statistics to case studies. Solving real live problems by applying adequate statistical methods. Studying different topics in statistics in order to help students develop statistical thinking. Learning from data or turning data into knowledge from planning for the collection data and data management to exploratory data analysis, statistical data analysis, interpretation of statistical software outputs, presentation of results and composition of the report. Finally, illustrating a variety of topics in statistical theory.

### Contents

1. A case study is proposed to a group of two students. More precisely, each group has to solve a scientific or an economic problem. For solving it, students have :
  - to make investigations about the problem by reading different papers or books
  - to collect data
  - to analyse the data by applying the adequate statistical method
  - to summarize and interpret outputs (tables and graphs) provided by a statistical software (S-Plus or R)
  - to write a report on their findings
2. Groups will present their solutions, and supply their own handouts, outputs and materials
3. Different areas of statistics will be covered in the seminar as for example :
  - data and information visualisation
  - data mining (knowledge discovery in databases, large-scale data analysis)
  - business analytics and management statistics
  - quality control (industrial statistics)

### Evaluation

According curriculum 2009-2010 :

- CC : final grade according to the following weighting system : 80% for the report and 20% for the presentation.
- *Reexamination session (September) : a new projet will be given with the same system.*

### Textbooks

- Bishop, C. M. (2006). *Pattern Recognition and Machine learning*. Springer: Berlin
- Groebner, D. F., Shannon, P. W., Fry, P. C. & Smith, K. D. (2005). *Business Statistics, A Decision – Making approach* (6th Edition). Pearson International Edition: New Jersey
- Han, J. & Kamber, M. (2001). *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers: San Diego
- Harrel, F. E. Jr (2001). *Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer Verlag: New-York
- Hastie, T., Tibshirani, R. & Friedman, J. H. (2001). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Springer Series in Statistics: New-York
- Larose, D. T. & Vallaud, T. (2005). *Des données à la connaissance : Une introduction au data mining*. Vuibert: Paris
- Montgomery, D. C. (2005). *Introduction to Statistical Quality Control* (5th Edition). John Wiley & Sons: New York
- Moore, D. S. & McCabe, G. P. (2003). *Introduction to the Practice of Statistics* (4th Edition). W. H. Freeman & Co.: New York
- Nolan, D. & Speed, T (2001). *Stat Labs, Mathematical Statistics Through Applications*. Springer Texts in Statistics: New-York

## Characteristics

- 3 ECTS credits
- Compulsory course for master in statistics
- Spring semester
- Course : 2 hours
- Prerequisite : common basis in probability and statistics

## Teaching team

- *Prof. HES Jacques Zuber*  
Département des Technologies de l'Information et de la Communication (TIC)  
Haute École d'Ingénierie et de Gestion du Canton de Vaud  
1, route de Cheseaux  
CH - 1401 Yverdon-les-Bains  
☎ phone office : +41 (0)24 557 64 59 - ✉ email : [jacques.zuber@unine.ch](mailto:jacques.zuber@unine.ch)

## Exercises

Application of the methods using software **R**.