

- Faculté des sciences
- www.unine.ch/sciences

### Seminar of applied statistics (3ST2011)

Filières concernées	Nombre d'heures	Validation	Crédits ECTS
<b>Master en statistique</b>	<b>Séminaire: 2 ph</b>	Voir ci-dessous	3

ph=période hebdomadaire, pg=période globale, j=jour, dj=demi-jour, h=heure, min=minute

#### Période d'enseignement:

- Semestre Printemps

#### Equipe enseignante

Professor Jacques Zuber  
University of Neuchâtel  
jacques.zuber@unine.ch

#### Contenu

- Different areas of statistics will be covered in the seminar as for example:
  - data and information visualisation
  - machine learning techniques (decision trees, neural networks, support vector machines, bagging, random forests, boosting, ...)
  - data mining (knowledge discovery in databases)
  - big data analytics (methodological training in data science)
  - business analytics and management statistics
- A case study is proposed to a group of two students. 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
- Groups will present their solutions, and supply their own handouts, outputs and materials

#### Forme de l'évaluation

CA graded: Continuous assesment, final grade according to the following weighting system : 80% for the report and 20% for the presentation.

Reexamination next session (August-September): a new projet will be given with the same system.

#### Documentation

- Bishop, C. M. (2009). Pattern Recognition and Machine Learning. Springer: New York
- Groebner, D. F., Shannon, P. W. & Fry, P. C. (2017). Business Statistics, A Decision-Making Approach (10th Edition). Pearson International Edition: New Jersey
- Han, J., Kamber, M. & Pei J. (2011). Data Mining: Concepts and Techniques (3rd Edition). Morgan Kaufmann Publishers: San Diego
- Harrell, F. E. Jr (2015). Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer Series in Statistics: New York
- Hastie, T., Tibshirani, R. & Friedman, J. H. (2009). The Elements of Statistical Learning. Data Mining, Inference, and Prediction (2nd Edition). Springer Series in Statistics: New York
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. Springer Texts 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. (2012). Statistical Quality Control: A Modern Introduction (7th Edition). John Wiley & Sons: New York
- Moore, D. S., McCabe, G. P. & Craig, B. A. (2015). Introduction to the Practice of Statistics (8th Edition). W. H. Freeman & Co.: New York
- Nolan, D. & Speed, T (2001). Stat Labs, Mathematical Statistics Through Applications. Springer Texts in Statistics: New York

#### Pré-requis

common basis in probability and statistics

#### Forme de l'enseignement

- Faculté des sciences
- [www.unine.ch/sciences](http://www.unine.ch/sciences)

**Seminar of applied statistics (3ST2011)**

- 3 ECTS credits
- Compulsory course for master in statistics
- Spring semester
- Exercises : Application of the methods using software R.