■ Faculté des sciences économiques
■ www.unine.ch/seco

Multivariate Analysis

Characteristics

- 3 ECTS credits
- Compulsory course for master in statistics
- Autumn semester
- Course : 2 hours
- Evaluation : 2-hours written exam
- Prerequisite : basics in statistics and linear algebra

Teaching Team

- *Ludovic Lebart*
  Ecole Nationale Supérieure des Télécommunications
  46 rue Barrault, 75013, Paris, France
  ☎+33 1 45 81 75 59 - ✉ lebart@enst.fr

Objectives

This course is designed to broaden the student's understanding of the statistical processing of multivariate data. It should enable him or her to master both the theoretical background and the context of applications of multivariate analysis. At the end, the student should be able to apply the multivariate techniques presented in this course to his or her own research studies. The student should also be able to carry out real life applications with a critical appraisal of the results and conclusions. The main domains leading to multivariate data sets are: socio-economic surveys, biometrics, behavioural sciences, geographic data bases, demographic, marketing research, data mining, text mining.

Content

The course will start with a brief review of basics of statistics and matrix algebra. It will cover then the two broad types of methods in Exploratory Multivariate Analysis: principal axes techniques (singular value decomposition, principal component analysis, canonical analysis, simple and multiple correspondence analysis) and clustering techniques (hierarchical clustering, k-means and related methods, self organizing maps). Particular emphasis will be placed on the links between these exploratory tools and some classical predictive methods such as multiple regression, discriminant analysis, regression trees. The course also comprises a presentation of the assessment methods that involve resampling schemes such as the *bootstrap* method. Eventually, since hand calculations are virtually impossible in this advanced field, application examples will be performed with the aid of the software R and DTM (Data and Text Mining).

Textbooks

L. Lebart, A. Morineau, K. Warwick (1984) *Multivariate Descriptive Statistical Analysis*, Wiley, (Wiley Series in Probability and Mathematical Statistics), New York

L. Lebart, M. Piron, A. Morineau (2006) *Statistique Exploratoire Multidimensionnelle*, Dunod, 4ème édition, 480p *(in French)*