

MSA200, MULTIVARIATE STATISTICAL ANALYSIS, 7.5 credit points

Level: advanced

1. Authorisation.

The course plan has been authorised by the vice-dean of the Department of Mathematical Sciences on November 9, 2006, to be valid from July 1, 2007.

Educational field: Mathematical Sciences

2. Educational context.

The course is part of the Master Program in Mathematical Sciences. It is also open for students outside the program who meet the course prerequisites.

3. Prerequisites.

MSG500, Linear statistical models.

4. Goals and learning outcomes.

The goal of the course is to give the student familiarity with theory and applications of multivariate analysis. Students are to gain an understanding of the type of research questions and the type of data that have lead to the development of different statistical methods, they will become familiar with different computational tools needed to carry out multivariate analyses and with case studies which illustrate the interpretation of statistical results in the light of corresponding research questions.

5. Course description.

The course contains both statistical and computational theory for estimation and testing hypotheses in connection with multivariate analyses. It is centered around the following models as building blocks.

Models for univariate dependencies, which may be combined for some multivariate analyses, include

- logit regression,
- logistic regression,
- linear regressions including interactive and nonlinear components,
- principal component regression

- ridge regression.

Models for joint response dependencies include - multivariate linear regression
- subclasses of identified structural equation models, including factor analysis and models for panel studies
- log-linear models.

The following special computational tools are used for solving maximum-likelihood equations for problems with different types of exponential family distribution
- iterative proportional fitting
- cyclic fitting algorithms
- expectation-maximisation algorithms
- asymptotically efficient closed-form approximations.

Methods to check the quality and properties of a set of given data include
- checks for nonlinear and interactive effects,
- computation of determinants, condition numbers,
- triangular, orthogonal and singular value decompositions.

6. Literature.

Will be provided for each topic from lecture notes and publications.

7. Assessment.

There are repeated home assignments. A minimum of 50 percent of correct solutions is required in each to pass the course. The final grade is a sum score of the solutions reached in each home assignment.

8. Grades.

The grade levels are Fail (U), Pass (G), and High Pass (VG). A wish for an ECTS grade should be reported to the examiner at the beginning of the course.

9. Course evaluation.

In the middle of the course the teacher arranges a feedback discussion with the students and at the end of the course the students will be asked to answer a questionnaire. The results of the questionnaire will be processed by the teacher together with student representatives.