UNIVERSITY OF GOTHENBURG

DEPARTMENT OF MATHEMATICAL SCIENCES

MSG500 Linear Statistical Models, 7.5 credits

Linjära statistiska modeller, 7,5 högskolepoäng

First Cycle

Confirmation

This course syllabus was confirmed by Department of Mathematical Sciences on 2019-05-17 and was last revised on 2020-01-31 to be valid from 2020-08-31, autumn semester of 2020.

Field of education: Science 100%
Department: Department of Mathematical Sciences

Position in the educational system

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

The course can be part of the following programmes: 1) Mathematical Sciences, Master's Programme (N2MAT), 2) Applied Data Science Master's Programme (N2ADS) and 3) Bachelor's Programme in Mathematics (N1MAT)

Main field of studies	*Specialization*
Mathematical Statistics	G1F, First cycle, has less than 60 credits in first-cycle course/s as entry requirements

Entry requirements

For entrance to the course, knowledge corresponding to the courses MSG110 Probability Theory and MSG200 Statistical Inference is required

Learning outcomes

On successful completion of the course, the student should be able to

- explain the common mathematical structure of linear regression models and generalized linear models
- construct and use these models for data analysis using statistical inference and suitable software
- interpret the results and criticize the model limitations
- identify data analysis situations for which linear models apply naturally and to estimate and interpret parameters
- predict future observations and test hypotheses using suitable software such as R
- construct regression models that are suitable for the current data but can also generalize to future observations
- explain the model limitations, identify situations where the hypothesized model is not suitable for the given data, and possibly transform the data to increase the model predictive ability

Course content

The course covers the following topics:

- simple linear and multivariate linear models and underlying assumptions
- the bias/variance trade-of
- properties of least squares estimators
- identification of outliers and the use of residuals and other diagnostics to verify if model assumptions are met
- the use of categorical covariates in regression
- testing parameters using the t-test
- goodness of fit indices (R2 and adjusted R2)
- confidence and prediction intervals
- the multicollinearity problem, its identification and remedial measures
- Model selection via greedy algorithms (stepwise procedures) and the AIC
- Model selection via the partial F test
- Prediction error and cross validation
- Interaction between covariates
- an introduction to generalized linear models, the exponential family, and asymptotic properties of the maximum likelihood estimators
- testing procedures for generalized linear models

Form of teaching

Lectures; weekly (or almost weekly) mini-projects and presentations

Language of instruction: English

Assessment

Summary report of the weekly mini-projects; a final project report; a written exam. Attendance to the weekly presentations of mini-analyses is mandatory. See the course page for how to compensate for missed attendance.

Grades

The grading scale comprises: Pass with Distinction (VG), Pass (G) and Fail (U).
The grade levels are Fail (U), Pass (G), and High Pass (VG). Students who are contractually entitled to ECTS grades should inform the examiner about this no later than one week after the start of the course. Students without such entitlement will not be awarded ECTS grades. Grades will be converted into ECTS terminology according to a standard model approved by the University President.

Course evaluation

Written course evaluation will be performed at the end of the course. The results of the evaluation will be communicated to the students and will serve as a guide for the development of the course.