*Course syllabus*

# Linjär och logistisk regression
# Linear and Logistic Regression

## FMSN30, 7,5 credits, A (Second Cycle)

**Valid for:** 2023/24
**Faculty:** Faculty of Engineering, LTH
**Decided by:** PLED I
**Date of Decision:** 2023-04-14

## General Information

**Elective for:** BME4, D4, E4-ae, F4, F4-fm, F4-mai, I4, L4-fe, Pi4-fm, Pi4-biek, Pi4-bam, MMSR1, R4
**Language of instruction:** The course will be given in English

## Aim

Regression analysis deals with modelling how one characteristic (height, weight, price, concentration, etc) varies with one or several other characteristics (sex, living area, expenditures, temperature, etc). Linear regression is introduced in the basic course in mathematical statistics but here we expand with, e.g., "how do I check that the model fits the data", "what should I do i it doesn't fit", "how uncertain is it", and "how do I use it to draw conclusions about reality".

When perfoming a survey where people can awnser yes/no or little/just fine/much, or car/bicycle/bus or some other categorical alternative, you cannot use linear regression. Then you need logistic regression instead. This is the topic in the second half of the course.

If you have a data material suitable for analysis using linear or logistic regression, you may analyse it as part of the project.

## Learning outcomes

*Knowledge and understanding*
For a passing grade the student must

- Describe the differences between continuous and discrete data, and the resulting consequences for the choice of statistical model
- Give an account of the principles behind different estimation principles,

- Describe the statistical properties of such estimates as appear in regression analysis,
- Interpret regression relations in terms of conditional distributions,
- Explain the concepts odds and odds ratio, and describe their relation to probabilities and to logistic regression.

*Competences and skills*
For a passing grade the student must

- Formulate a multiple linear regression model for a concrete problem,
- Formulate a multiple logistic regression model for a concrete problem,
- Estimate the parameters in the regression model and interpret them,
- Examine the validity of the model and make suitable modifications of the model,
- Use the model resulting for prediction,
- Use some statistical computer program for analysis of regression data, and interpret the results,
- Present the analysis and conclusions of a practical problem in a written report and an oral presentation.

*Judgement and approach*
For a passing grade the student must

- Always check the prerequisites before stating a regression model,
- Evaluate the plausibility of a performed study,
- Reflect over the limitations of the chosen model and estimation method, as well as alternative solutions.

## Contents

Least squares and maximum-likelihood-method; odds ratios; Multiple and linear regression; Matrix formulation; Methods for model validation, residuals, outliers, influential observations, multi co-linearity, change of variables; Choice of regressors, F-test, likelihood-ratio-test; Confidence intervals and prediction. Introduction to: Correlated errors, Poisson regression as well as multinomial and ordinal logistic regression.

## Examination details

**Grading scale:** TH - (U,3,4,5) - (Fail, Three, Four, Five)
**Assessment:** Written and oral project presentation, peer assessment and oral exam.

The examiner, in consultation with Disability Support Services, may deviate from the regular form of examination in order to provide a permanently disabled student with a form of examination equivalent to that of a student without a disability.

**Parts**
**Code:** 0117. **Name:** Examination.
**Credits:** 3. **Grading scale:** TH. **Assessment:** Oral examination.
**Code:** 0217. **Name:** Project 1.
**Credits:** 1,5. **Grading scale:** UG. **Assessment:** Written project report and peer assessment **Contents:** Linear regression
**Code:** 0317. **Name:** Project 2.
**Credits:** 1,5. **Grading scale:** UG. **Assessment:** Written project report and peer assessment **Contents:** Logistic regression
**Code:** 0417. **Name:** Project 3.
**Credits:** 1. **Grading scale:** UG. **Assessment:** Oral project presentation **Contents:** Other regression models
**Code:** 0517. **Name:** Laboratory Work.
**Credits:** 0,5. **Grading scale:** UG. **Assessment:** Computer exercises

## Admission

**Admission requirements:**

- FMSF20 Mathematical Statistics, Basic Course or FMSF25 Mathematical Statistics - Complementary Project or FMSF32 Mathematical Statistics or FMSF45 Mathematical Statistics, Basic Course or FMSF50 Mathematical Statistics, Basic Course or FMSF55 Mathematical Statistics, Basic Course or FMSF70 Mathematical Statistics or FMSF75 Mathematical Statistics, Basic Course or FMSF80 Mathematical Statistics, Basic Course

**The number of participants is limited to:** No
**The course overlaps following course/s:** MASM22, FMSN40

## Reading list

- Rawlings, J.O., Pantula, S.G., Dickey, D.A.: Applied Regression Analysis - A Research Tool, 2ed. Springer, 1998, ISBN: 0-387-98454-2. Available as e-book.
- Alan Agresti: An introduction to categorical data analysis, 2nd ed. Wiley, 2007, ISBN: 978-0-471-22618-5. Available as e-book.

## Contact and other information

**Director of studies:** Johan Lindström, studierektor@matstat.lu.se
**Course administrator:** Susann Nordqvist, expedition@matstat.lu.se
**Course homepage:**
https://www.maths.lu.se/utbildning/civilingenjoersutbildning/matematisk-statistik-paa-civilingenjoersprogram/
**Further information:** The course is also given at the faculty of science with the code MASM22. Only one of the courses FMSN30 and FMSN40 may be included in a degree.