

# Statistics and Data Science II

Single subject and programme course

7.5 credits

Statistik och dataanalys II

771A18

Valid from: 2018 Autumn semester

**Determined by**

The Quality Board at the Faculty of Arts  
and Sciences

**Date determined**

2017-10-20

## Main field of study

Computational Social Science

## Course level

Second cycle

## Advancement level

A1X

## Entry requirements

A bachelor's degree or equivalent in social science, physical science, biological science, engineering, statistics or math. Additionally required at least 15 ECTS credits in one or several of the following subjects: statistics, computer science, mathematics or equivalent, and at least 7,5 ECTS credits at advanced level in one or several of the following subjects: statistics, computer science, mathematics or equivalent.

English corresponding to the level of English in Swedish upper secondary education (English 6/B).

## Intended learning outcomes

After completion of the course, the student should at an advanced level be able to:

- use statistical software to estimate appropriate linear regression models for cross-sectional and panel data and explain the statistical principles underlying these estimates;
- use statistical software to assemble appropriate data structures for estimating regression models and implementing robustness checks;
- interpret the parameters of linear regression models, produce predictions, and evaluate goodness of fit;
- describe the logic of causal inference and how it applies to regression models, distinguishing between causality and correlation;
- identify common threats to causal interpretation of linear models, and assess and justify modeling approaches for solving these threats;
- evaluate the validity and robustness of causal inferences under a variety of assumptions about how the data was generated.

## Course content

This course introduces the principles and practice of linear regression modeling. Underlying model assumptions are reviewed and scrutinized. In intensive computer laboratories, statistical tools for creating appropriate data structures and estimating models using real data are presented and guidance is provided in interpretation of model parameters. The remainder of the course focuses on causal inference and the potential outcomes framework. Panel data models and statistical tools for their estimation are presented, and their potential to improve causal inference are compared. Discussion is extended to consider natural experiments and instrumental variable approaches to causal inference. The sensitivity of estimates to violations of model assumptions are evaluated, with special attention given to methods centering on computer simulation.

## Teaching and working methods

The teaching consists of readings, lectures, seminars, and interactive computer labs. Homework and independent studies are a necessary complement to the course.

Language of instruction: English.

## Examination

The course is examined through written assignments, completed computer laboratories, and a final individual written assignment. Detailed information about the examination can be found in the course's study guide.

Students failing an exam covering either the entire course or part of the course twice are entitled to have a new examiner appointed for the reexamination.

Students who have passed an examination may not retake it in order to improve their grades.

## Grades

ECTS, EC

## Other information

Planning and implementation of a course must take its starting point in the wording of the syllabus. The course evaluation included in each course must therefore take up the question how well the course agrees with the syllabus.

The course is carried out in such a way that both men's and women's experience and knowledge is made visible and developed.

## Department

Institutionen för ekonomisk och industriell utveckling