FACULTY OFFICE FOR SCIENCE

UNIVERSITY OF
GOTHENBURG

NMAT302   Statistics using R , 5 credits

Statistik med R, 5 högskolepoäng

Third-cycle level / Forskarnivå

Confirmation

This syllabus was confirmed by the Department of Mathematical Sciences on 2019-03-27, and is valid from Spring semester 2019.

Responsible Department
Faculty Office for Science, Faculty of Science

Entry requirements

The course is open to all graduate students enrolled in doctoral or licentiate programmes at the Faculty of Science. If seats are available, we also welcome senior researchers within the Faculty of Scince and PhD students from other universities.

The course is designed as a second course in statistics, and you should already have had at least a basic undergraduate course in statistics of a size corresponding to 2-3 weeks of full time studies. Previous experience of programming is not required.

Learning outcomes

This course aims to train PhD students in using flexible statistical computing environments, explorative modelling based on appropriate graphics, and on model selection techniques. This is to help overcome the problem that there is an increasing gap between statistical knowledge and hands-on skills of the typical graduate students in sciences and the expected scientific data analysis and statistical understanding needed for their licentiate or PhD theses and in their future research careers.

Fulfilling this course, participants will know:

- When, where, and how to utilize basic statistical notions and basic standard methods.

- Explain and specify the potential of statistical methods and analysis tools that are available to use, and how to adapt or develop these further in connection to their specific data analysis problem areas.

- Utilize the computational statistical environment R to an appropriate level of skills and understanding in their daily research.

Course content

We will make it possible for students to work with data from their own research, when teaching methods relevant for that type of data. While the core course focuses primarily on hands-on skills for data analysis, optional exercises designed to deepen statistical understanding will also be available.

The full course is 5 hp, but if you choose to cut the course after two first teaching blocks you will get a shorter (3 hp) course variant.

R is the primary programming tool for statistical data analysis. It is a free statistical environment that has been created by the academic community. Besides statistics, it is also widely used in other empirical research fields and in industry. The experience is that students who are used to other languages (Matlab, Python etc.) quickly get used to R, and that experience in R will be useful even if the use of Matlab or Python would be required in the future.

Types of instruction

The course takes place in Mathematical Sciences facilities on Chalmers Campus Johanneberg in three blocks, each with three full day studies in class. Participants should add an equal amount of equally distributed self-study time.

Language of instruction
The course is given in English.

Grades

The grade Pass (G) or Fail (U) is given in this course.

Types of assessment

There will be two minor examination computer projects to be performed and handed in after the first block ideally before the start of the second block, one examination project to be performed between the second and the third project, and one final examination project with more flexibility started in the third block and handed in after a week. This last examination project may if found suitable be individually coupled to some of your own statistically based research questions or it may be coupled to one of the Block 3 presented areas.

The full course is 5 credits, but if you choose to cut the course after two first teaching blocks you will get a shorter (3 credits) course variant.

Course evaluation

An opportunity for course evaluation will be presented on the last day of the course.

Other information

Bring Computers

For simplicity of organization, and for general convenience, we require all participants to bring suitable portable computers to the lectures. We will make sure that there are enough electricity plug in possibilities in the lecture rooms that we use.

Further preliminary course details:

- R is a public license software and is downloadable over internet. https://www.r-project.org/
- There is a Wikipedia text with a lot of useful information: https://en.wikipedia.org/wiki/R_(programming_language)
- There are several books devoted to R-based introductory and secondary level statistics, and the early part of our course will be inspired by the book "Introductory Statistics with R" by the Danish Professor P. Dahlgaard, see: https://www.amazon.com/Introductory-Statistics-R-Computing/dp/0387790535

However there are many other options, some of which are published in e-book or downloadable electronic formats.