



Exchange programme Vrije Universiteit

Vrije Universiteit Amsterdam - Exchange programme Vrije Universiteit - 2022-2023

Exchange

Vrije Universiteit Amsterdam offers many English-taught courses in a variety of subjects, ranging from arts & culture and social sciences, neurosciences and computer science, to economics and business administration.

The International Office is responsible for course approval and course registration for exchange students. For details about course registration, requirements, credits, semesters and so on, please [visit the exchange programmes webpages](#).

Statistical Data Analysis

Course Code	X_401029
Credits	6.00
Period	P4+5
Course Level	300
Language Of Tuition	English
Faculty	Faculty of Science
Course Coordinator	dr. D. Dobler
Examiner	dr. D. Dobler
Teaching Staff	dr. D. Dobler
Teaching method(s)	Partial Exam, Lecture, Seminar

Course Objective

After the student has followed the course, she/he should be able to

- find a suitable model distribution for a dataset at hand, with the help of QQ-plots, symplots, histograms, box plots, goodness-of-fit tests, etc., e.g., a normal or exponential distribution; to estimate unknown (e.g. location and/or scale) parameters, and the underlying density function,
- decide which statistical method is preferred by taking characteristics of the dataset into account (e.g. when to use a nonparametric test), to draw conclusions on the population underlying the data, for example with the help of hypothesis tests and confidence intervals,
- apply tests for location parameters, stochastic order, or equality of distributions in two-sample problems, and be able to assess the asymptotic relative efficiency of tests,
- apply (in R) resampling methods such as the bootstrap or random permutation to find characteristics of a statistic, even if no model assumptions are made,
- analyze, with the help of rank-based correlation tests, chi-square tests for contingency tables, or multiple linear regression, the relationship between two or more variables in a given dataset. In the context of multiple linear regression, the student should be able to identify influential observations and select variables for the linear regression model,
- communicate with colleagues about statistical topics, i.e. solve the assignments in groups of two students, talk to other students, the teaching assistants, and the teacher of the course about the statistical subjects,
- create informative and scientifically appropriate reports, i.e. a good mixture of figures and written text, completeness of the report, yet conciseness, question-oriented reporting with the use of adequate language.

Course Content

This is an advanced level statistical data analysis course that builds on introductory course on probability and statistics, e.g. Statistics (X_400004) and Probability Theory (X_400622). The course introduces the students to several widely used statistical models and methods, and the students learn how to apply these tools to real data with the use of the statistical software package R. The following subjects are covered:

- summarizing data;
- investigating the distribution of data;
- density estimation;
- bootstrap;
- nonparametric methods;
- two-sample problems;
- contingency tables;

- multiple linear regression.

The course is a combination of theory (in the lectures) and practice (in the exercise classes) in such a way that the theory is explicitly linked to the practice of statistical data analysis.

Additional Information Teaching Methods

Lectures (13x2h; once per week), computer classes (13x2h; once per week).

Attendance is not mandatory but strongly recommended.

Method of Assessment

Homework assignments with the help of the programming language R and two written exams.

50% of the final grade consists of the average assignment grades, the other 50% of the final grade consists of the exam grade. Both of these grades have to be at least 5.5. Otherwise, the course is failed.

The exam grade equals either the average of the grades of both partial exams, that are written during the semester, or it equals the resit exam grade.

If the resit exam is written, the homework assignment grades still count towards the final course grade as explained above; there is no resit possibility for the assignments.

If both partial exams are passed with an average of at least 5.5, then doing the resit exam is not possible anymore.

Literature

Lecture notes.

Additional Information Target Audience

2BA, 2W, 2W-B, 3W, 3W-B, 3Ect.

Additional Information

Language of tuition: English

Recommended background knowledge

The required knowledge has been obtained if the students had previously passed the VU courses Statistics (X_400004) / Mathematical Statistics (XB_0049) and Probability Theory (X_400622) or equivalent courses.