# Exchange programme Vrije Universiteit Amsterdam

Vrije Universiteit Amsterdam - Exchange programme Vrije Universiteit Amsterdam - 2024-2025

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 |
| Period | P4+5 |
| Course Level | 300 |
| Language Of Tuition | English |
| Faculty | Faculty of Science |
| Course Coordinator | C. Damian |
| Examiner | C. Damian |
| Teaching Staff | C. Damian |
| Teaching method(s) | Written partial exam, Lecture, Seminar |

## Course Objective

After students have followed the course, they should be able to

- make use of QQ-plots, symplots, histograms, boxplots, goodness-of-fit tests, etc. to find a suitable model distribution for a dataset at hand;
- estimate unknown model parameters (e.g. location and/or scale) and underlying density function;
- decide which statistical method is preferred to draw conclusions about the population underlying the data, for example using hypothesis tests and confidence intervals, while taking characteristics of the dataset into account (e.g. knowing when to use a nonparametric test);
- 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;
- using R, apply resampling methods such as the bootstrap or random permutation to find characteristics of a statistic even if no model assumptions are made;
- analyze the relationship between two (or more) variables in a given dataset with the help of rank-based correlation tests, chi-square tests for contingency tables, or multiple linear regression. In the context of multiple linear regression, students should be able to select variables for the linear regression model and identify potentially influential observations;
- communicate with colleagues about statistical topics, e.g. by solving the assignments in groups of two students and by talking to other students, teaching assistants, and the teacher of the course;
- create informative and scientifically appropriate reports which include relevant text, graphics and (if necessary) code. Reports should use appropriate language and be concise yet complete, as well as tidy, self-explanatory and (if applicable) reproducible.

## Course Content

This is an advanced level statistical data analysis course that builds on introductory courses 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 R. The following topics 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 (lectures) and practice (exercise classes), with the aim to directly link the theory 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

- 50% of the final grade consists of an average of homework assignment grades, the other 50% consists of the exam grade. Both of these components 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 the grade of the resit exam;
- 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;
- Homework assignments should be solved with the help of the programming language R.

## Literature

Lecture notes.

## Additional Information Target Audience

Business Analytics Year 2, Mathematics Year 2 or 3

## 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.