

Statistics

Winter 2024/25 – Syllabus

Important information:

- We will communicate any organizational changes via the StudOn website for this course. Make sure to register on [StudOn](#)!
- Please do not underestimate the fact that many topics in the course build on one another. To keep up throughout the semester, you need to start learning from the beginning and stay on track!

Lecture: Wednesday, 9:45 - 11:15, LG H5

Exercise session: Friday, 11:30 – 13:00, LG H6

R tutorial: Friday, 13:15 – 14:45, LG 0.422

Instructor:

Jonas Dovern

Office: LG 4.169

E-Mail: jonas.dovern@fau.de

Office hours: by appointment

Assistant

Gohar Grigoryan

Office: LG 4.172

E-Mail: gohar.g.grigoryan@fau.de

Office hours: by appointment

Overview: We have to deal with statistics and probabilities every day. Is a certain drug effectively curing headache? How long, on average, do I wait for the subway? Can women drive better than men? Of course, statistical methods are also key in business and economics and an essential tool that everybody active in these fields should know about. Does a new web design attract more buyers than the old one? How many clients should a company survey to obtain reliable evidence about the interests of their entire clientele? What is the average household income in Germany? Which country has a particularly high income inequality? Does it take less than 6 months, on average, for an unemployed person to find a new job?

This course offers an introduction to statistical methods that you can use to i) describe and plot data, ii) plan a data collection project, iii) derive general statements from data samples, iv) analyze random processes, and v) analyze big data sets by means of machine learning algorithms.

Since an increasing number of jobs in both the private sector and economic research requires knowledge about how to handle data for empirical analyses, this course does also cover the practical implementation of statistical methods using the software R.

Students who successfully master this course, are able to i) describe data sets using appropriate measures, ii) select appropriate types of plots to visualize data sets, iii) estimate distribution parameters, iv) conduct and interpret statistical hypothesis tests, v) use inductive statistics as a foundation for their empirical work and to critically reflect about

statistical results in general, vi) explain the basic functioning of selected machine learning algorithms, and vii) to implement the discussed methods using the software R.

Grading: Grading is based on a single-choice exam (90 minutes). You can improve your grade by successfully participating in six quizzes during the semester.

Exam date: At the end of the semester. The date is determined and communicated later by the examination office.

Course outline:

Descriptive statistics:

- Why do we need statistics?
- Which types of data are there?
- How can we describe and plot the distribution of data?
- How do we measure relationships between two variables?

Theory of probability:

- What is a probability?
- How can we calculate conditional probabilities?
- What is a random variable and how do we describe it mathematically?
- Which important models for discrete and continuous random variables exist?

Inductive statistics:

- What can we infer from a random data sample?
- How can we estimate unknown distribution parameters?
- How can we test if a parameter has a certain value?
- How can we test if two samples are similar?

Basics of machine learning:

- What is machine learning?
- How can we use machine learning to group a sample in a data-driven way?
- How can we use machine learning to predict random variables?

Textbook: Newbold, P., W. L. Carlson, and B. M. Thorne (2022), *Statistics for Business and Economics - Global Edition*, 10th edition, Pearson.

Software R: We will devote a substantial part of the exercise sessions to the implementation of statistical methods using the software R. There will also be a weekly R tutorial that you can attend to get help with any problems you have with the software or the R problem sets.

We do not expect any prior knowledge with R or programming in general. Instead, we offer the [online course "Basic Introduction to R/RStudio"](#) via StudOn. You need to study this course on your own during the first two weeks of the semester (and then hand in the course certificate via StudOn). This online course also explains how you can install the free software packages R and RStudio (an IDE for R) on your own computer.

Course requirements: Course participants are strongly advised to ...

- **Attend.** You can only fully benefit from this course if you attend both lectures and exercises regularly because the course content is cumulative, meaning that later topics rely heavily on stuff covered in the previous weeks.
- **Prepare.** Do the assigned readings before the lecture and come to class prepared to discuss them and to ask questions that you have. Similarly, work on the problem sets before we discuss them in class.
- **Follow the website.** We will make course material available through the course website on StudOn. We will also make announcements using this platform.
- **Code.** Your long-term learning gains will be much, much higher if you regularly work on the R assignments, which ask to implement the material that we cover in the lectures. Do not underestimate how much actually applying an approach helps understanding it!