STOCKHOLMS UNIVERSITET
Statistiska institutionen
Autumn term 2025

Dan Hedlin
26 June 2025

Course Description
Selected Statistical Methods with Applications, 7.5 ECTS, undergraduate level, ST3201

Course contents

The course aims to broaden students' knowledge in statistics and give inspiration when choosing topic for the bachelor thesis.

The course comprises three selected topics: survival analysis, missing data methods and text analysis. For each topic, the course relates theory to practice by combining lectures, practical exercises and programming. A central element is the students' independent work with a number of problems (case studies).

Expected learning outcomes

After completing the course, students should be able to:

- describe definitions and central concepts,
- explain relevant theory,
- discuss pros and cons of different methodological considerations,
- solve relevant theoretical problems,
- apply the selected methods to various problems,
- perform computation and analyses in R,
- present methods and applicable problem solutions in writing and orally,
- interpret, evaluate and critically review results with regard to relevant scientific aspects.

Teaching format

Teaching consists of 12 lectures (L1 - L12), three computer sessions (C1 - C3) and three seminars (S1 – S3).

Attendance at the seminars is compulsory. For students who have been absent from one or more seminars an extra seminar will be organised. The course coordinator or the teacher responsible for the part will inform about the date for this seminar via Athena. See further details under Examination and assessment criteria.

Submission of home assignments before seminars and oral presentations at seminars are mandatory elements of the course.

The course is given on campus, in English.

Stockholms universitet

STOCKHOLMS UNIVERSITET
Statistiska institutionen
Autumn term 2025

Dan Hedlin
26 June 2025

Examiner, course coordinator	***Reception hours***	***E-mail***
Dan Hedlin	On agreement	dan.hedlin@stat.su.se

The course covers the following three parts/topics.

Part 1: Missing data methods

This part provides an introduction to handling missing data. Topics include: missing data theory and single and multiple imputation methods. Additionally, in the computer session students will learn how to implement the imputation techniques using R.

Teaching consists of four lectures (L1 – L4), one computer-session (C1) and one seminar session (S1) according to the schedule.

Literature: van Buuren, S. (2018). Flexible imputation of missing data, 2nd ed. Boca Raton, Florida: CRC Press. Available online via the library of Stockholm University.
Some research papers and other material provided by instructor will be made available at Athena.

Responsible for part 1: Dan Hedlin, email: dan.hedlin@stat.su.se. Consultation hours by appointment.

Part 2: Analysis of survival data

This part provides introduction to models and methods used in the analysis of survival (duration) data - with applications in the social sciences. Relevant R-procedures will be also covered in the computer session. Preliminary topics that will be covered include: censoring and other special features of survival data; functions of survival time; Kaplan-Meier and life-table estimation of survival functions; log-rank test for comparison of survival functions; Cox proportional hazards models; parametric survival models.
For successful completion of this part students should be able to: describe and explain basic concepts, functions, and distributions for survival data; compute and compare survival functions for different groups; and model associations between survival functions and explanatory variables using R.
Teaching consists of 4 lectures (L5 – L8), one computer-session (C2) and one seminar session (S2) according to the schedule. Lecture notes and other relevant material can be handed-out (or made available in Athena) in connection with the lectures and computer session.

Stockholms universitet

STOCKHOLMS UNIVERSITET
Statistiska institutionen
Autumn term 2025

Dan Hedlin
26 June 2025

Literature: Moore, D. F. (2016), *Applied Survival Analysis Using R*. Springer. Available online via SU-library.

Responsible for part 2: Gebrenegus Ghilagaber, email: Gebre@stat.su.se.
Consultation hours: in connection with lectures or by appointment.

Part 3: Text mining

Nowadays, there are vast quantities of unstructured textual information available, for example, from emails, medical journals, social media activities, movie recommendations and web server logs. It is of utmost importance (e.g. for companies) to be able to analyze such information and make it quantifiable, in order to see trends and remain competitive.

This part provides an introduction to quantitative methods for analyzing text. Students will learn how to retrieve the text from the original source, process the text data, analyze and summarize the results of text mining experiments.

For successful completion of this part students should be able to: describe and explain basic concepts related to text mining, apply text mining methods to practical problems using software R.

Teaching consists of four lectures (L9 – L12), one computer-session (C3) and one seminar session (S3) according to the schedule. Lecture notes and other relevant materials will be available in Athena in connection with the lectures and computer session.

Literature: research papers and other material provided by instructor will be available via Athena.

Responsible for part 3: Edgar Bueno, email: edgar.bueno@stat.su.se. Consultation hours by appointment.

Examination and assessment criteria

The course is examined through three individual assignments. To pass the course, all four of the following requirements need to be met for each part of the course:

1. The written report is submitted at the latest at 15.00 hrs the day before the seminar
2. a minimum of 14 points for the written report
3. a minimum of 3 points for the oral presentation at the seminar
4. attendance at the seminars

Stockholms universitet

STOCKHOLMS UNIVERSITET
Statistiska institutionen
Autumn term 2025

Dan Hedlin
26 June 2025

A student may fail if she or he is unable to give satisfactory answers to written or oral questions about any of the written. Should the examiner wish to ask further questions about any assignment orally, a time will be agreed on with the student.

The sum of the points from the three individual assignments defines the final grade. Maximum 100 points. Part 1: 34 points (28 for written report, 6 for oral presentation), part 2: 33 points (27 for written report, 6 for oral presentation), part 3: 33 points (27 for written report, 6 for oral presentation).

Grading of the course is done according to a seven-point scale related to the specified learning outcomes: A = Excellent, B = Very good, C = Good, D = Satisfactory, E = Adequate, Fx = Inadequate, F = Very inadequate.

No points from the assignments achieved in the autumn term 2025 can be transferred to the next time the course will be given.

Students who fail any of requirements 1-4 above are given one opportunity for re-exam normally seven days after the scheduled seminar. However, then at most 20 points are given for this assignment. The next opportunity after that will take place in spring term 2026. Then all three assignments need to be retaken.

The grading criteria are as follows.
A: Excellent (90-100 points). The student has in a really well-structured and excellent manner solved the pre-specified statistical problem that reflects the course material using software R. Furthermore, the student has also demonstrated the ability to solve problems that have not explicitly been explored in the course material. The student was able to draw correct conclusions from the statistical analyses, and clearly present the obtained results (in a written report and orally).

B: Very good (80-89 points). The student has in a well-structured and correct manner solved the pre-specified statistical problem that reflects the course material using software R. The student has demonstrated the ability to solve problems that were partly explored in the course material. The student was able to draw correct conclusions from the statistical analyses, and clearly present the obtained results (in a written report and orally).

C: Good (70-79 points). The student has correctly solved the pre-specified statistical problem that reflects the course material and that was directly explored in the course material. The student was able to use software R for performing statistical analyses, to draw correct conclusions, interpret and discuss the obtained results (in a written report and orally).

D: Satisfactory (60-69 points). The student has mostly correctly solved the pre-specified statistical problem that reflects the course material and that was directly explored in the course material. The student was able to use the software R for conducting the statistical

analysis, to draw mostly correct conclusions from this and to interpret the obtained results (in a written report and orally).

E: Adequate (51-59 points). The student has for the most part correctly solved the pre-specified statistical problem that reflects the course material and that was directly explored in the course material. The student was able to use the software R for conducting the statistical analyses, to draw in most cases largely correct conclusions and to interpret the obtained results in an adequate manner (in a written report and orally).

Fx: Inadequate (26-50 points). The student has met some but not all requirements for an E grade.

F: Very inadequate (0-25 point). The student cannot correctly apply statistical methods that have been considered in the course

Use of AI

The use of AI tools is permitted as an aid during the learning process but not to produce material for any kind of examination.

- Any type of plagiarism is prohibited, which includes AI-generated text.
- The use of AI tools for the improvement of an originally self-written text is not permitted.
- Text matching software and AI-generated text detectors are used by the department.

Read "Guidelines for disciplinary matters at Stockholm University"

https://www.su.se/medabarteta/organisation-styrning/styrdokument-regelboken/utbildning/regler-och-handläggningsordning-för-disciplinärenden-1.605869