Stockholm University

Department of Statistics
Fall 2026

Statistics and Data Analysis for Computer and Systems Sciences 15 ECTS credits, ST1601 Course description

Course contents

The course provides knowledge of classical and modern statistical methods for data analysis as well as its theoretical foundations. Central is understanding of the entire process of analysis from data sources and data collection, data management, estimation, inference, prediction and practical applications. Great emphasis is placed on practical data handling, visualization and analysis through programming in R. Throughout, emphasis is placed on using a critical approach when using statistical methods and interpreting results.

The course covers:

- data collection methods and data sources. Different data types such as numerical and categorical. Graphical and numerical descriptions of data;
- regression analysis: models with one and more explanatory variables, assumptions, estimation, inference, prediction, model evaluation. Time series analysis and forecasting. The connection to modern data analysis methods such as machine learning is addressed;
- probability theory: basic concepts, probability models, discrete and continuous random variables, probability distributions, expected value and variance, covariance and correlation, different standard distributions, linear combinations of several random variables, sampling distributions and the central limit theorem;
- statistical inference, point and interval estimation, hypothesis testing, p-values and prediction and introduction to likelihood.

Learning goals

- Account for different data collection methods and data sources;
- Explain basic statistical concepts;
- Process data, describe data numerically and graphically and perform analyzes through basic programming in R;
- Formulate statistical models and solve basic problems in probability theory and inference;

- Perform regression and time series analysis and select the appropriate model for a few different types of cases;
- Interpret, evaluate and critically review results with regard to relevant scientific aspects.

Examination and grading

Students will be assessed based on the course's learning goals. The formal assessment consists of four parts, shown in the table below.

Code	Type	Part	Credits	First	Second
11ST	Exam	Part 1	4.5 hp	November 27, 14-18	TBA
12SI	Home Assignment	Part 1	3 hp	November 23, 23.59	December 7, 23.59
13ST	Exam	Part 2	6 hp	January 14, 08–13	TBA
14SI	Home Assignment	Part 2	1.5 hp	January 11, 23.59	January 25, 23.59

- Exam part 1 (11ST) and Exam part 2 (13ST) are individual written, proctored exams and are graded according to the following seven-point scale:

  A: Excellent,

  B: Very good,

  C: Good,

  D: Satisfactory,

  E: Adequate,

  Fx: Fail, inadequate,

  F: Fail, totally inadequate.

  These exams will each be offered twice.

- Home Assignment part 1 (12ST) and Home Assignment part 2 (14ST) are group assignments that are graded according to a two-point scale where **G** is a passing grade and **U** is a failing grade.
- Each examination is graded separately and independently. This means that if you pass a test, you are not required to re-take that examination should you fail to pass any of the others. E.g. if you have passed both home assignments but fail to pass the exams, you will retain the corresponding credit points and will not be required to do them again; you are only required to do the parts you have not yet passed.
- A minimum grade of E on the each exam and a grade of G on each home assignment is required for a passing grade on the course. Students who have not earned a passing grade on all four tests will not receive a final grade.
- Combined grade for the entire course is determined according to the following matrix:

		Exam Part 2				
		A	B	C	D	E
Exam Part 1	A	A	B	B	C	C
	B	A	B	C	C	D
	C	B	B	C	D	D
	D	B	C	C	D	E
	E	C	C	D	D	E

Rules for Exams (11ST and 13ST)

The following applies to both exams:

- During the examination no forms of collaboration or discussion are allowed. Calculators will be provided in the exam halls; you are not allowed to bring your own calculator. Information about the calculators can be found here. Formula sheets will be provided by the department during the exam.
- If necessary, special support can be allowed upon request to the department's student counsellor (studivagledare@stat.su.se) and after approval by the examiner. Contact student counsellor in good time before the exam, preferably no later than three weeks before the exam takes place.
- Students must register for the exams/re-exams via Ladok. Registration must normally be made no later than 10 days before the exam, but in some cases the deadline is earlier than that. Check in Ladok which deadline applies to each individual exam. Registration opens 30 days before the exam. If you have registered correctly, an anonymous code is generated directly on the screen. If you have problems with your registration, contact the Student Office (expedition@stat.su.se) before the registration deadline. If you fail to register, you cannot take the exam!
- Students who get at least grade E on an exam are not allowed to retake it for obtaining a higher grade.
- Both Fx and F on an exam are failing grades and require re-examination. Supplementary assignments in order to raise an Fx to a passing grade is not permissible for this course.
- Students who received a grade of Fx or F on an exam have the right to take at least four additional exams as long as the course is given to achieve at least a grade of E.
- Students who have received the grade Fx or F on an exam twice by the same examiner have the right to request that another examiner be appointed to determine the grade at the next exam. Such a request must be made to the head of the department in written form.

Rules for Home Assignments (12ST and 14ST)

Home assignments are to be solved in groups of 2 or 3 students. The following applies to both assignments:

- Collaboration *within* the group is of course permitted and encouraged. Note, however, that grades are set individually and can vary between group members, it is the individual's performance in the group work that is examined. All group members are responsible for and must be able to answer for all parts of the work reported in the report. All team members must participate in all parts of the work. Collaboration *between* groups is permitted, but each group must submit its unique report.
- Any type of plagiarism is prohibited. This includes text generated by AI tools. The use of AI tools is permitted as an aid in terms of knowledge acquisition and study but not to produce material for any kind of examination. The use of AI tools for the improvement of an originally self-written text is not permitted. Text matching software is used by the institution. Read Guidelines for disciplinary matters at Stockholm University.
- If an assignment handed in by the first submission date fails, students have the opportunity to correct mistakes and hand in a revised assignment by the second deadline. The second deadline constitutes the second examination opportunity. If you do not submit your assignment before the first deadline, and submit your report for the first time by the second deadline, you will not have the opportunity to revise and correct your report.

- Feedback for the home assignments should be available around 5–7 working days after the deadline.

Grading criteria for Exam Part 1 (11ST)

A: (At least 90 points on the written exam.) Excellent. The student can use concepts and results in data analysis, probability, and regression (that are not necessarily directly covered in the course) in an excellent way. The student can solve and interpret problems of an in-depth nature about data analysis, probability, and regression in a well-structured manner.

B: (Between 80 and 89 points on the written exam.) Very good. The student can use concepts and results in data analysis, probability, and regression (covered in the course) in a very good way. The student can solve and interpret problems of an in-depth nature about data analysis, probability, and regression in a well-structured manner.

C: (Between 70 and 79 points on the written exam.) Good. The student can use concepts and results in data analysis, probability, and regression (covered in the course) in a good way. The student can solve and interpret problems of an in-depth nature about data analysis, probability, and regression in some concrete situations.

D: (Between 60 and 69 points on the written exam.) Satisfactory. The student can satisfactorily use concepts and results in data analysis, probability, and regression that are covered in the course. The student can correctly solve and interpret problems of an in-depth nature about data analysis, probability, and regression.

E: (Between 50 and 59 points on the written exam.) Sufficient. The student can adequately use concepts and results in data analysis, probability, and regression that are covered in the course part. The student can mostly correctly solve and interpret problems of an in-depth nature about data analysis, probability, and regression.

Fx: (Between 40 and 49 points on the written exam.) Fail, some more work required. Re-examination is required.

F: (Less than 40 points on the written exam.) Fail, much more work required. Re-examination is required.

Grading criteria for Home Assignment 1 (12ST)

G: (Passed.) The student has set up appropriate statistical models for given situations, has demonstrated sufficient ability to use basic terminology in data analysis, has used the statistical software R in a correct way, has presented the results in a written report in accordance with the instructions and has been present at all compulsory teaching sessions.

U: (Fail.) The student's performance is insufficient with respect to at least one of the criteria for G.

Grading criteria for Exam Part 2 (13ST)

A: (At least 90 points on the written exam.) Excellent. The student can make excellent use of concepts and results within the theory of probability and statistical inference theory (not necessarily directly covered in the course). The student can solve and interpret problems about point and interval estimation of parameters as well as perform hypothesis testing of parameters in a well-structured manner.

B (Between 80 and 89 points on the written exam.) Very good. The student can make a very good use of concepts and results within the theory of probability and statistical inference theory (covered in the course). The student can solve and interpret problems about point and interval estimation of parameters as well as perform hypothesis testing of parameters in a well-structured manner.

C: (Between 70 and 79 points on the written exam.) Good. The student can make a good use of concepts and results within the theory of probability and statistical inference theory (covered in the course). The student can solve and interpret problems about point and interval estimation of parameters as well as perform hypothesis testing of parameters in some concrete situations.

D: (Between 60 and 69 points on the written exam.) Satisfactory. The student can satisfactorily use concepts and results within the theory of probability and statistical inference theory (covered in the course). The student can correctly solve and interpret problems about point and interval estimation of parameters as well as perform hypothesis testing of parameters.

E: (Between 50 and 59 points on the written exam.) Sufficient. The student can adequately use concepts and results within the theory of probability and statistical inference theory (covered in the course). The student can mostly correctly solve and interpret problems about point and interval estimation of parameters as well as perform hypothesis testing of parameters.

Fx: (Between 40 and 49 points on the written exam.) Fail, some more work required.

F: (Less than 40 points on the written exam.) Fail, much more work required.

Grading criteria for Home Assignment 2 (14ST)

G: (Passed.) The student has demonstrated sufficient ability to calculate and interpret point and interval estimates and perform hypothesis testing of parameters, has demonstrated sufficient ability to use basic statistical terminology, has used the statistical software R correctly, has presented the results in a written report in accordance with the instructions and has been present at all compulsory teaching sessions.

U: (Fail.) The student's performance is insufficient with respect to at least one of the criteria for G.

Course literature

- De Veaux, R., Velleman, p. and Bock, D. *Stats: Data and Models.* Pearson, Global Edition. Fifth edition.
- Other course material such as lecture notes, exercises, instructions for submissions and more are posted on the course page in Canvas. There you can also find some preliminary instructions of reading and exercises exercises for each teaching session.

Teaching format

The teaching consists of 26 lectures, 15 exercise sessions and 8 computer exercises. In addition there are 11 recitations (jour) scheduled. See the reading instructions that are available in Canvas for a reading list and for a list of exercises from the literature that will be addressed at the exercise sessions. You can see the full schedule here.

Attendance at the first lecture on Wednesday, November 4, is mandatory. This session will provide important information about the course, its structure, home assignments, and other relevant details. Attendance to any other session in the course is voluntary but strongly recommended.

Teachers and general information

Teacher	Role
Ulf Högnäs	Examiner & Course Coordinator Lectures (Part 2)
Edgar Bueno	Lectures (Part 1)
Akram Mahmoudi	Exercise sessions Recitations (Jour)
Mina Poshtvar	Home assignments Computer labs
Mathias Stenutz	Home assignments Computer labs

All teachers have reception times by appointment.

The Department of Statistics is located at Campus Albano, Albanovägen 12, Building 4, Floor 6. General information about the department can be found on the department's website. You can also visit the course webpage.