STOCKHOLM UNIVERSITY
Department of Statistics
Spring 2025

Johan Koskinen

Course Description for Generalised linear models, 7.5 ECTS Credits, Code: ST2201

Course content

The course is divided into two interrelated parts:

- Exam, 5 ECTS Credits
- Assignment, 2.5 ECTS Credits

both of which are based on the same teaching material but test different learning outcomes. The exam is a sat, pen and paper, exam, that tests your theoretical and technical grasp on the core material, as well as overall comprehension. In particular, emphasis will be put on how you relate the new material to previous knowledge and skills that you have acquired in statistics. In the assignment you will demonstrate that you have the practical skills, sound statistical literacy, and critical thinking to apply the methods in this course to real data.

COURSE LITERATURE

The course material that is the basis for the examinations is

- Faraway, J. J. (2016). Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. CRC press
- Extra material posted on Athena, when relevant
- Relevant course material from previous courses in Semester I and Semester II, or equivalent

During the course, online material will be suggested for further reading and to deepen and solidify your knowledge.

Other course material such as, for example, formula sheets, and hand-in assignments will be available on Athena.

TEACHING AND EXAMINATION
The full course consists of 13 lectures (L1-L13), 6 problem-solving sessions (E1-E6) and 4 computer exercises (C1-C4).

Details of the content at each teaching occasion can be found at the end of this document.

The examination consists of two exams and two compulsory hand-in assignments:

- **Exam**: Regression analysis, 5 ECTS Credits
- **Hand-in assignment**: Compulsory exercise in GLM, 2.5 ECTS Credits.

Allowed tools for the exam: Pocket calculator and own formula sheet that the student themselves collate (1 double-sided A4 page),

The topics of the hand-in assignment will be announced through Athena. The assignments are done as a group project (a group consists of 2-3 students) and are presented as a written report. We encourage discussions among students between groups during the work on the assignment, but each group must finally be able to answer for their own handed-in written report.

Dates for exams

Exam: Monday 28 April, 2025, 8am-12noon
Resit Exam: Wednesday 11 June, 2025, 2pm-6pm

Remember to sign up for the exam <u>at least 10 days</u> before the examination. If you do not register, you cannot take the exam.

Deadline for handing in group assignments:

- Assignment: 3 May, 2025
- Resubmission: 15 June, 2025

GRADING CRITERIA FOR THE EXAMS
At each exam, a maximum total credit of 100 points will be awarded. Grades are given according to the following seven-point rating scale:

Grade	Grading points on the exam
A	90-100
B	80-89
C	70-79
D	60-69
E	50-59

Fx	40-49
F	0-39

To pass the exam a minimum grade of E is required. Failing grades are F and Fx. When obtaining a failing grade F or Fx in the written examination, we will not give extra exercises or extra assignments to obtaining a passing grade.

Hand-in assignments:
The compulsory hand-in assignments are graded as Pass (G) or Fail (U). If a compulsory assignment is graded as Fail (U), the student will have the chance to re-submit an assignment.

FINAL GRADE FOR THE COMPLETE COURSE (15 ECTS CREDITS)
To pass the entire course, a minimum grade of E for the exams and a pass (G) grade for the hand-in assignment are required.

A	**Excellent.** The student should in a proper and well-structured way be able to apply generalised linear models and associated statistical inference that are not necessarily directly addressed in the course material. The student is also able to clearly present and interpret his/her results; explain concepts, methods and theories used in the implementation of generalised linear models.	Total: 90-100 points Written exam: ≥40 points Home assignment: ≥10 points
B	**Very good.** The student will correctly and in a well-structured way be able to apply generalised linear models and associated statistical inference that are directly addressed in the course material. The student is also able to clearly present and interpret his/her findings; explain the concepts, methods and theories used in the implementation of generalised linear models.	Total: 80-89 points Written exam: ≥40 points Home assignment: ≥10 points
C	**Good.** The student will correctly and in a well-structured way be able to apply generalised linear models and associated statistical inference that are directly addressed in the course material. The student should also in a good way be able to present and interpret his/her findings; explain concepts, methods and theories used in the implementation of generalised linear models	Total: 70-79 points Written exam: ≥40 points Home assignment: ≥10 points
D	**Satisfying.** The student will be able to apply generalised linear models with related statistical inference that are directly addressed in the course material. The student will forward in a satisfactory way to present and interpret his/her findings; explain concepts, methods	Total: 60-69 points Written exam: ≥40 points Home assignment: ≥10 points

	and theories used in the implementation of generalised linear models.	
E	**Sufficient:** The student will be able to apply generalised linear models practices directly addressed in the course material. The student will, in a satisfactory way, present and interpret his/her findings; explain the concepts, methods and theory used in the implementation of generalised linear models.	Total: 50-59 points Written exam: ≥40 points Home assignment: ≥10 points
Fx	**Insufficient.** The student's achievements with respect to at least one of the criteria in E have serious shortcomings. .	Total: ≥30 points Written exam: <40 points and/or Home assignment: <10 points
F	**Totally insufficient.** The learning outcomes have not been achieved.	Total: 0-29 points Written exam: <40 points and/or Home assignment: <10 points

TEACHERS AND COMMUNICATION DETAILS

Course coordinator, examiner, and lecturer: Johan Koskinen
Email: johan.koskinen@stat.su.se
Office hours: On agreement.

Teaching assistants: Yulia Ryanova
Email: yulia.ryanova@stat.su.se
Office hours: On agreement.

Fasna Kottakkunnan
Email: fasna.kottakkunnan@stat.su.se
Office hours: On agreement.

Information of the department of Statistics can be found at: www.statistics.su.se.
Course information can be found on the course webpage:
https://www.su.se/english/search-courses-and-programmes/st2201-1.644481?open-collapse-boxes=course-time-table

TEACHING PLAN

Below a preliminary course schedule over the contents, reading preparation and exercises for each teaching occasion is given. Note that *Far* refers to the main course book **Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models** (first, Far.1, or second, Far.2, editions). Lectures that are prefaced by 'Repetition' are meant to be refreshers on content from previous courses and based on Past Material (PM), with some additional Extra Material (EM) that will be posted on Athena. Extra material may be distributed through Athena during the course. Exercises for each Exercise session will be provided ahead of the exercise session.

Lectures		
	Contents	**Readings**
F1	Information about the course Summary of course Repetition: Expected values, Common distributions	Far.2.1 (Far.1. 1) PM & EM
F2	Repetition: Expected values, Common distributions (cont.)	
F3	Repetition: Linear regression	Far.2.1 (Far.1.1) & PM
F4	Repetition: Linear regression (continued)	Far.2.1 (Far.1.1) & PM
F5	Linear probability model and Logistic regression	Far.2.2 (Far.1. 2.1-2.5)
F6	Logistic regression (cont.) and Probit link function	Far.2.3.6, Far.12.4.2 (Far.1. 2.1-2.5)
F7	Poisson regression	Far.2.5.1 (Far.1. 3.1-3.2)
F8	Negative binomial	Far.2.5.4 (Far.1. 3.3)
F9	Multinomial	Far.2.7.1, 7.4 (Far.1. 5)
F10	Revision of models in GLM framework	Far.2.3.1,3.3,3.4 (Far.1. 6)
F11	Revision of models in GLM framework (continued)	Far.2.4.2, 5.5, 8.1 (Far.1.7)
F12	Inference	Far.2. A1 & chapters 8.1, 8.2, 8.3, 8.4 (Far.1. A1 & chapters 1-7)
F13	Going further: overview of extensions	Far.2.9.1, 10 intro, 10.1, 10.2+10.6, 10.8, 10.9,10.10, 13.1 (Far.1.7-12)

Exercises	
	Contents
E1	Linear Regression and distributions
E2	Expected values and common distributions
E3	Logistic regression and probit regression
E4	Count data
E5	Multinomial
E6	Using different link functions

Computer labs	
C1	Far.2: 2.1, 3.1, 3.2, 3.3 (Far.1: 2.1, 2.2, 2.4, 2.7)
C2	Far.2: 5.1, 5.2, 5.3, 5.5 (Far.1: 3.1, 3.2, 3.4, 3.5)
C3	Far.2: 7.1, 7.3, 7.4 (Far.1: 5.1, 5.3, 5.4)
C4	Far.2: 9.7, 9.2, 10.2, 13.1 (Far.1: 6.2, 7.2, 8.2, 10.1)

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, and this includes AI-generated text.
- The use of AI tools for the improvement of a text is not permitted even if the text in question were an original text written by you.
- 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