Course description

Generalized linear models, 7.5 HEC, ST425A

Advanced level
Fall 2024

1 COURSE CONTENTS

The course introduces models belonging to the class of the Generalized Linear Models that extend the linear modelling framework to response variables that are not normally distributed. The focus of the course is on the analysis of discrete data, e.g. counts, proportions, success/failure records, questionnaire items.

Both the statistical inference and applications of the models to real-life data using statistical software will be considered. Prior knowledge in probability and inference theory, calculus, matrix algebra, and computer literacy are assumed. The topics that will be covered in more details include:

• Likelihood theory and inference

• Exponential family of distributions

• Generalized linear models

• Linear models for continuous responses

• Logistic regression models for binary, ordinal and nominal responses

• Poisson regression and Log-linear models

• Models for longitudinal and clustered data

2 LEARNING GOALS

After completing the course, the student is expected to:

• explain the theory behind generalized linear models and statistical methods for data analysis covered by the course;

• choose an appropriate model for data analysis and argue for its adequacy;

• interpret, evaluate and critically review results of the data analysis with regard to relevant scientific aspects;

• discuss limitations and possible error sources in the analysis;

• solve problems and perform standard type calculations for the models and methods covered of the course;

• perform calculations and conduct data analysis using software R for the models covered in the course.

3 TEACHING FORMAT

Teaching consists of 12 lectures (L1 - L12) and 6 computer-sessions (C1-C6) according to schedule. More detailed instructions on these will be provided during the first lecture (L1) and/or first computer session (C1). The teaching consists of lectures, exercises and computer exercises. The course is given on campus, though some digital teaching may occur. Necessary information will be provided in Athena during the course.

4 COURSE LITERATURE AND TEACHING MATERIALS

• Dobson, A. J. and Barnett, A. G. (2018). An Introduction to Generalized Linear Models (Fourth Edition). Boca Raton, FL: Chapman and Hall/CRC. •

Software: R

Some additional material will be available in Athena during the course.

5 COURSE INSTRUCTOR, EXAMINER AND TEACHING ASSISTANT

Lecturer & Examinator: Per Gösta Andersson Room A4604 (Bldg. 4, Albanovägen 12, 6th floor, Albano Campus), E-mail: per.gosta.andersson@stat.su.se Consultation hours: In connection with Lectures or by appointment.

Teaching assistant: Oscar Oelrich, E-mail: oscar.oelrich@stat.su.se

6 EXAMINATION AND GRADING

The course is examined through an individual examination (code 11ST) and a home assignment (code 12SI). The written examination is graded according to a seven-grade grading scale: A = Excellent, B = Very good, C = Good, D = Satisfactory, E = Adequate, Fx = Inadequate, F = Totally Inadequate. Both Fx and F are failing grades and require re-examination. The Home Assignment is graded according to a two-grade grading scale: G = Passed, U = Failed. An individual performance within the group work will be assessed.

Individual examination (max 50 points) consists of two parts: Part I: a written examination (max 30 points) that covers the material of the course, Part II: a written examination (max 20 points) on practical data analysis using software R.

Usually the student should sign up for the written exam within 10 days before the exam, but special cases can change this, so this should be checked by the student. The student can sign up for the exam 30 days in advance. If that is done correctly an anonymous code is generated immediately on the screen. If there is trouble doing this contact the student expedition (expedition@stat.su.se). **If you have not signed up you cannot take the exam!**

The minimum points required to pass the course are 20 points for Part I and 10 points for Part II. Requirements for the Part II:

• The student should properly use the theory of Generalized Linear Models when conducting data analysis.

• The student should be able to correctly use statistical terminology and clearly present and interpret the obtained results.

• The statistical software used should be R. • The results should be reported in the form of a written report.

The final grade of the individual examination is based on the sum of points for Part I and Part II, according to the criteria given in Table 1.

• A grade of E or higher is required in order to pass the course.

• Students who have attained the grade Fx or F on an individual examination are entitled to at least four additional individual examinations to achieve at least grade E as long as the course is offered.

• Students who have received at least grade E on an individual examination may not retake another examination in an attempt to achieve a higher grade.

• Students who have received the grade Fx or F on individual an examination on two occasions by the same examiner have the right to request that a different examiner be appointed to set the grade of the examination on the next possible occasion. The request must be in writing and sent to the head of the department. The examination denotes all compulsory elements of the course.

The following criteria are used for assessing the Home Assignment:

Pass: The student can properly apply the theory of Generalized Linear Models when conducting the task specified in the Home Assignment. The student can identify and use an appropriate statistical model for data analysis and argue for its adequacy. The student can in a satisfactory way present in a written form his/her findings, interpret them using proper graphics, descriptive statistics, significance test etc., explain important statistical concepts relevant to the task, discuss limitations and possible error sources in the analysis. The student can correctly use statistical software R and summarize his/her work in a well-written report using a correct statistical language.

Fail: The student cannot correctly apply the theory of Generalized Linear Models when conducting the task specified in the Home Assignment. The student cannot identify and/or use an appropriate statistical model for data analysis and argue for its adequacy. The student is not able in a satisfactory way to present his/her findings, the student cannot interpret the obtained results using proper graphics, descriptive statistics, significance test etc. The student cannot explain important statistical concepts relevant to the task, and/or discuss limitations and possible error sources in the analysis. The student cannot correctly use statistical software R. The student is not able to summarize his/her work in a well-written report using a correct statistical language.

For a passing grade on the course, a minimum grade of E on an individual examination and a grade of G on Home Assignment is required. The combined grade for the entire course is equated with the grade for the individual examination. Examination assignments that are not submitted on time will not be graded.

Grading criteria for the individual examination Grade Criteria (Requirements)

Excellent A The student can correctly describe and use the theory of generalized linear models as it is considered in the course. Moreover, the student can apply this theory to practical problems that have not been directly discussed in the course. The student can clearly present correct solutions to

problems relevant for the course and use a correct statistical language. The grade corresponds to a total of 47.5 -50 points of the examination score.

Very Good B The student can correctly describe and use the theory of generalized linear models as it is considered in the course. The student can apply this theory to practical problems of a type that has been discussed in the course. The student can clearly present correct solutions to problems relevant for the course and use a correct statistical language. The grade corresponds to a total of 45 - 47 points of the total examination score.

Good C The student can correctly describe and use the theory of generalized linear models as it is considered in the course. Moreover, the student can apply this theory to practical problems of a type that has been discussed in the course. The student can present correct solutions to problems relevant for the course and use a correct statistical language. The grade corresponds to a total of 40 - 44.5 points of the total examination score.

Satisfactory D The student can correctly describe and use the theory of generalized linear models as it is considered in the course. The student can apply this theory to a majority of practical problems that have been discussed in the course. The student can present correct solutions to a majority of problems relevant for the course and use a satisfactory statistical language. The grade corresponds to 35 - 39.5 points of the total examination score.

Sufficient E The student can correctly describe and use the theory of generalized linear models as it is considered in the course. The student can apply this theory to a majority of practical problems that have been discussed in the course. The student can present satisfactory solutions to problems relevant for the course and use a satisfactory statistical language. The grade corresponds to 30 - 34.5 points of the total examination score.

Insufficient Fx The student cannot correctly describe and use the theory of generalized linear models as it is considered in the course. The student cannot present satisfactory solutions to problems relevant for the course or cannot use a satisfactory statistical language. The grade corresponds to 20 - 29.5 points of the total examination score.

Totally Insufficient F The student cannot describe and use the theory of generalized linear models as it is considered in the course. The student cannot present satisfactory solutions to problems relevant for the course or cannot use a statistical language. The grade corresponds to 0 - 19.5 points of the total examination score.

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-handl¨aggningsordning-f¨or-disciplin¨arenden-1.605869