STOCKHOLM UNIVERSITY
Department of Statistics
Spring Semester 2025

Johan Koskinen

Course Description
Multivariate Methods, 7.5 ECTS credits, ST3101

COURSE CONTENTS

This course provides an introduction to several important multivariate methods. Among others, the following key concepts will be explained: principal components, exploratory and confirmatory factor analysis, discriminant analysis, logistic regression, and cluster analysis. Basics of matrix algebra required for the course are introduced. The statistical software R will be used in laboratory sessions to exemplify and complement lecture material.

LEARNING GOALS

To pass the course the student should be able to:

- Give an account of the most common multivariate methods
- Apply statistical software to analyze data modelled by some of the multivariate methods as well as be able to interpret relevant outputs

COURSE LITERATURE AND ADDITIONAL MATERIAL

The course book:

- Zelterman, D. (2015). Applied Multivariate Statistics with R. Springer.

Other recommended books:

- Johnson, R.A., Wichern, D.W. (1988). Applied Multivariate Statistical Analysis
- Afifi, A.A., Virginia, C. (2003). Computer-Aided Multivariate Analysis

Additional material distributed during the course, e.g. lecture notes, exercises, previous exams, etc., will be posted on Athena https://itslearning.com/se/.

GENERAL INFORMATION, TEACHERS

The Department of Statistics is located on the 6th floor in House 4 (Albanovägen 12). General information about the department (office hours, phone numbers, schedules etc.) is posted on the department website, www.statistics.su.se.

Course coordinator and examiner: Johan Koskinen, room A4630
E-mail: johan.koskinen@stat.su.se
Reception hours: Mondays 1-2 PM or by appointment

TEACHING

The course covers the course book sections that are listed in the table below. There are 14 scheduled lectures (L) and 7 computer labs (C). Each lecture will typically address specified chapters and sections in the course literature including some exercises. During the teacher-led computer labs students solve computer-based exercises related to the various topics of the course. A preliminary outline of the course is provided in the following tables. You should always check the current schedule (accessible via link) for correct information and updates about the schedule.

Lecture	Content	Course book section
L1	Introduction and overview	Ch. 1-3
L2	Introduction to matrix algebra: definitions and basic operations	Ch. 4 and extra material (EM) posted on Athena
L3	Introduction to matrix algebra: linear combinations, the matrix inverse, and determinants	Ch. 4 and EM
L4	Introduction to matrix algebra: linear combinations, the matrix inverse, and determinants (continued)	-"-
L5	The data matrix and properties of the multivariate normal distribution	Ch.1-3, and ch. 4 and EM; Ch 5-7 and EM
L6	Principal Components Analysis (PCA)	Ch. 8.1-8.4 and EM
L7	PCA cont.	Ch. 8.1-8.4 and EM
L8	Factor Analysis (FA)	Ch. 8.5-8.6 and EM
L9	FA cont.	Ch. 8.5-8.6 and EM
L10	Introduction to discrimination and classification and Logistic regression (LR)	Ch. 10, 10.2 and EM
L11	LR cont.	Ch. 10.2 and EM
L12	Discriminant analysis (DA)	Ch. 10.3 and EM
L13	Cluster Analysis (CA)	Ch. 11 and EM
L14	Previous exam - Recorded	

Lecture	Content
C1	Introduction, information about home assignment
C2	PCA
C3	PCA/FA
C4	FA
C5	LR
C6	DA
C7	CA

MANDATORY ATTENDANCE

Attendance is voluntary but strongly recommended in order to better achieve the learning goals.

COURSE EVALUATION

In connection with each course offering an evaluation of the course is made. The course evaluation is used as a basis for the departments quality work on the course, and as part of student influence. After the course is completed a questionnaire is sent via e-mail to all registered students. Students' responses to the questionnaire will be compiled and added together with the course coordinators' assessment to a final report, which will be posted on Athena.

EXAMINATION

Students are assessed by examination of the expected learning outcomes. The examination consists of a portfolio of two separate examination parts:

(1) Part I: an individual written (re-)exam with a selection of theoretical problems, and
(2) Part II: an individual home assignment with practical data analysis to be handed in as a written report.

A more detailed description, the grading criteria and weighting of the two components are provided below.

The final grade on the course is based on the total count of points received for both parts taken together. Grading of the course is on a criterion-referenced seven-grade scale.

A	Excellent
B	Very Good
C	Good
D	Satisfactory
E	Sufficient
Fx	Insufficient
F	Totally insufficient

NOTE! To pass the course (grades A – E) the student must pass both part I and part II. Both parts must be completed and passed during the continuation of the course. Partial exam or home assignment credits cannot be transferred to later semesters.

- Students who receive the grade E or higher on an examination may not re-do any part of the examination for a higher grade.
- Fx as well as F are both insufficient grades and require re-examination.
- Students who receive the grade Fx or F are entitled to at least four additional examination opportunities to achieve at least the grade E as long as the course is still given.
- Students who receive the grade Fx or F on an examination twice by the same examiner are entitled to request that a different examiner be appointed to set the grade of the next examination. Such a request must be in writing and sent to the head of the department

Note that the term 'examination' denotes all examination parts of the course.

EXAMINATION DATES AND RULES

Written exam	**Wednesday 3rd March, 14.00-19.00** Results will be announced no later than **24th March.**
Written re-exam	**Thursday 24th April, 14.00-19.00** Results will be announced no later than **15th May.**

- When the results of the written exam are announced the exams will be available at the Student Office and may be collected during office hours (see website).

- The written test is an individual assessment and collaboration of any kind is not allowed. The following aids are allowed for this exam: calculator and a specified list of formulas. Cell phones are not allowed.

NOTE! Registration for examination and re-examination is compulsory. Normally, you must sign up **at least 10 days** before the date of the examination/re-examination through minastudier.su.se. Occasionally, the **deadline is earlier that 10 days** prior to the exam. Therefore, it is your responsibility to check in Ladok what the last date of registration is. Registration is open from 30 days prior to the exam. If you experience any technical issues registering for the exam, contact the Student expedition (expedition@stat.su.se) prior to the registration deadline**. If you have not registered for the exam you are not allowed to sit the exam**.

Home assignment	**Deadline 1: Saturday 8 March at 17.00** Assessed by **16th March** **Deadline 2: Thursday 23rd March at 17.00** Assessed by **18th April**

- The home assignment should be handed in as a written report. Detailed instructions for the assignment will be posted on Athena.
- The home assignment is an individual assessment: collaboration is allowed but the written report is an individual effort. Note that your submission may be run through text matching software to discover plagiarism.
- If the assignment is not passed after the first round of assessment there is a possibility to revise the report, for details see the instructions for the assignment posted on Athena. Revisions should be handed in by deadline 2. Note that it is not possible to revise the report if it is not passed after deadline 2.

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

ASSESSMENT CRITERIA

The final grade is based on the total count of the scores of the two examination parts:

(1) Part I : written exam with maximum 80 points, to pass a minimum of 40 points is required
(2) Part II: home assignment with maximum 20 points, to pass a minimum of 10 points is required

The maximum total score is thus 100 points. In order to pass the course (grades A-E) the student must pass both examination parts; score at least 40 points on the written exam and at least 10 points on the home assignment. Grading of the course is on a criterion-referenced seven-grade scale:

A	**Excellent.** The student should in a proper and well-structured way be able to apply multivariate methods 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 multivariate analysis.	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 the multivariate methods 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 multivariate analysis.	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 the multivariate methods 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 multivariate analysis	Total: 70-79 points Written exam: ≥40 points Home assignment: ≥10 points
D	**Satisfying.** The student will be able to apply multivariate methods 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 and theories used in the implementation of multivariate analysis.	Total: 60-69 points Written exam: ≥40 points Home assignment: ≥10 points
E	**Sufficient:** The student will be able to apply multivariate practices directly addressed in the course material. The student will, in a satisfactory way, present and interpret his/her findings;	Total: 50-59 points Written exam: ≥40 points

	explain the concepts, methods and theory used in the implementation of multivariate analysis.	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