



## Course and Examination Fact Sheet: Spring Semester 2020

### 8,273: Multivariate Statistical Analysis

ECTS credits: 4

#### Overview examination/s

(binding regulations see below)

Decentral - Oral examination (individual) (50%, 15 mins.)

Decentral - Group examination paper (all given the same grades) (50%)

#### Attached courses

Timetable -- Language -- Lecturer

[8,273,1.00 Multivariate Statistical Analysis](#) -- Englisch -- [Fengler Matthias](#)

#### Course information

#### Course prerequisites

Initial programming knowledge, not necessarily in Python, solid knowledge and interest in statistics, econometrics, and data analytics.

#### Learning objectives

- Students will understand the challenges of multivariate data.
- Students will know the basic statistical techniques of variable selection, clustering, dimensionality reduction, and factor models.
- Students will know how to apply and interpret these methods for multivariate data analysis.
- Students will know how to take advantage of Python for multivariate data analysis.

#### Course content

##### Short summary

The class introduces students to the most important methods in multivariate statistics. Thereby, students will learn how to use Python for multivariate data analysis.

##### Detailed Description

Increasingly larger, richer and more versatile datasets are nowadays available for economic and financial research - a phenomenon usually referred to as the "Big Data paradigm". Corporates, banks, and policymakers aim at extracting knowledge from such data sets for various purposes, e.g., for understanding their costumers, improving their product lines, offering new products and services, and for policy decisions and policy implementations. Most of the ideas and methods used for such "fancy and modern" analyses are, in fact, methods coming from multivariate statistics.

The class covers the key methods from multivariate statistics in order to extract insights from multivariate data, in particular variable selection, factor space reduction, factor modelling, and cluster analysis. The objective is to combine conceptual/theoretical material with the practical application using the open source programming language Python.

A typical application could be this: suppose you like to use indicators of economic activity to predict or describe some variable of interest, but you have hundreds of them. How can you find a small set of relevant predictors for your purpose? Is there a way to comprise all variables in a single index such as to distil the essential information from all data? Can you distinguish certain groups among these variables, which may help understand their economic content?



The class connects to and expands on topics being part of "Data Analytics I: Predictive Econometrics" and naturally complements classes that more explicitly focus on the challenges of big data, such as big data analytics and machine learning. The course is recommended for students wishing to develop a profile in data science.

## Course structure

1. Descriptive techniques for multivariate data
2. Multivariate random variables and distributions
3. Theory of Estimation and Hypothesis testing
4. Variable selection: Lasso, LARS, Scad, Lasso extensions
5. Principal Component Analysis
6. Factor Models
7. Cluster analysis
8. Discriminant analysis

## Course literature

W. Härdle, L. Simar (2012): Applied Multivariate Statistical Analysis, Springer-Verlag

## Additional course information

--

## Examination information

### Examination sub part/s

#### 1. Examination sub part (1/2)

##### Examination time and form

Decentral - Oral examination (individual) (50%, 15 mins.)

##### Remark

--

##### Examination-aid rule

Extended Closed Book

The use of aids is limited; any additional aids permitted are exhaustively listed under "Supplementary aids". Basically, the following is applicable:

- At such examinations, all the pocket calculators of the Texas Instruments TI-30 series and mono- or bilingual dictionaries (no subject-specific dictionaries) without hand-written notes are admissible. Any other pocket calculator models and any electronic dictionaries are inadmissible.
- In addition, any type of communication, as well as any electronic devices that can be programmed and are capable of communication such as notebooks, tablets, PDAs, mobile telephones and others, are inadmissible.
- Students are themselves responsible for the procurement of examination aids.

##### Supplementary aids

none

##### Examination languages

Question language: English

Answer language: English

---

#### 2. Examination sub part (2/2)

##### Examination time and form



Decentral - Group examination paper (all given the same grades) (50%)

## Remark

Assignments

## Examination-aid rule

Term papers

- Term papers must be written without anyone else's help and in accordance with the known quotation standards, and they must contain a declaration of authorship.
- The documentation of sources (quotations, bibliography) has to be done throughout and consistently in accordance with the APA or MLA standards. The indications of the sources of information taken over verbatim or in paraphrase (quotations) must be integrated into the text in accordance with the precepts of the applicable quotation standard, while informative and bibliographical notes must be added as footnotes (recommendations and standards can be found, for example, in METZGER, C. (2017), Lern- und Arbeitsstrategien (12th ed., Cornelsen Schweiz).
- For any work written at the HSG, the indication of the page numbers both according to the MLA and the APA standard is never optional.
- Where there are no page numbers in sources, precise references must be provided in a different way: titles of chapters or sections, section numbers, acts, scenes, verses, etc.
- For papers in law, the legal standard is recommended (by way of example, cf. FORSTMOSER, P., OGOREK R. et SCHINDLER B. (2018, Juristisches Arbeiten: Eine Anleitung für Studierende (6. Auflage), Zürich: Schulthess, or the recommendations of the Law School).

## Supplementary aids

none

## Examination languages

Question language: English

Answer language: English

---

## Examination content

1. Descriptive techniques for multivariate data
2. Multivariate random variables and distributions
3. Theory of Estimation and Hypothesis testing
4. Variable selection: Lasso, LARS, Scad, Lasso extensions
5. Principal Component Analysis
6. Factor Models
7. Cluster analysis
8. Discriminant analysis

## Examination relevant literature

W. Härdle, L. Simar (2012): Applied Multivariate Statistical Analysis, Springer-Verlag



## Please note

Please note that this fact sheet alone is binding and has priority over any other information such as StudyNet (Canvas), personal databases or faculty members' websites and information provided in their lectures, etc.

Any possible references and links within the fact sheet to information provided by third parties are merely supplementary and informative in nature and are outside the University of St.Gallen's scope of responsibility and guarantee.

Documents and materials that have been submitted no later than the end of term time (CW21) are relevant to central examinations.

Binding nature of the fact sheet:

- Information about courses and examination time (central/decentral) and examination type starting from the beginning of the bidding on 23 January 2020
- Information about examinations (examination aid regulations, examination content, examination-relevant literature) for decentral examinations after the 4th semester week on 16 March 2020
- Information about examinations (examination aid regulations, examination content, examination-relevant literature) for central examinations as from the starting date for examination registration on 6 April 2020

Please consult the fact sheet again after these deadlines have expired.