GHENT UNIVERSITY

Course Specifications

Valid as from the academic year 2020-2021

Statistical Modelling (C003805)

Due to Covid 19, the education and evaluation methods may vary from the information displayed in the schedules and course details. Any changes will be communicated on Ufora.

Course size (nominal values; actual values may depend on programme)

Credits	Study time	Contact hrs
8.0	240 h	97.5 h

Course offerings and teaching methods in academic year 2020-2021

A (semester 1)	English	Gent	lecture	40.0 h
			project	7.5 h
			seminar: coached exercises	5.0 h
			group work	5.0 h
			microteaching	5.0 h
			seminar: practical PC room classes	15.0 h
			self-reliant study activities	20.0 h

Lecturers in academic year 2020-2021

Goetghebeur, Els	WE02	lecturer-in-charge
Loeys, Tom	PP01	co-lecturer

Offered in the following programmes in 2020-2021	crdts	offering
Master of Science in Statistical Data Analysis	8	A

Teaching languages

English

Keywords

Position of the course

The course assumes students come with basic statistical understanding and skills, that is, they master the material offered in the course `Principles of statistical data analysis'. It focuses on regression models to describe, explain and predict continuous or categorical outcomes from a set of measured covariates. Several other courses in the program build in turn on this course, most directly all other courses involving more advanced regression modeling, such as `Causality and missing data analysis', `Experimental Design', 'Analysis of clustered and longitudinal data', 'Survival Analysis'. This course aims to provide the student with a basic insight in and skills for statistical modeling. This will primarily be developed in the context of generalized linear models, starting from linear regression and with ample attention given to the analysis of tabulated data and logistic regression.

Contents

- The simple and generalized linear regression model for the analysis of continuous and categorical data
- The (iteratively reweighted) least squares algorithm
- The maximum likelihood principle for statistical models
- Confidence intervals for conditional means, prediction intervals for new observations
- Graphical and formal diagnostic methods for the inspection of residuals
- Goodness-of-fit tests, checking the model assumptions, detection of influential observations
- The impact of transforming variables
- Box-Cox regression for non-linear associations
- The link between linear regression and analysis of variance
- The multiple regression model with joint predictors that can be binary, categorical

and/or continuous
- Confounding and effect modification
- Procedures for stepwise building of a regression model
- Penalized regression methods
- Distribution and inference for categorical data
- Analysis of contigency tables
- Logistic regression
- Generalized additive models

Initial competences

Basic knowledge of probability and statistics, such as that covered in `Principles of Statistical Data Analysis'; and basic working knowledge of matrix algebra.

Final competences

1 The student can recognize practical problems which can be solved by means of the (multivariate) (generalized) linear model.
2 The student knows the distinction between association, prediction and causation, and understands the basic role of confounding in this context.
3 The student can interpret the generalized linear regression model correctly, can fit it to a data set and draw justified conclusions in the theoretical as well as the practical sense.
4 The student can apply diagnostic techniques to check the fit of a (multivariate) (generalized) linear model.
5 The student can adopt appropriate remedial measures when the current regression model does not fit.
6 The student knows the basic properties of the estimators and can work with the distinction between estimation and prediction.
7 The student can link regression analysis with analysis of variance.
8 The student recognizes the limitations of the (generalized) linear model, can suggest appropriate extensions and develop some methods via the maximum likelihood approach.
9 The student has theoretical knowledge about the most frequently used methods of categorical data analysis.
10 The student can correctly interpret and critically assess the results of a categorical data analysis.
11 The student can select appropriate statistical methods for categorical data analysis.
12 The student can adequately report in writing results of the statistical analysis of continuous and categorical data.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Guided self-study, group work, lecture, microteaching, project, seminar, self-reliant study activities, lecture: plenary exercises, seminar: coached exercises, seminar: practical PC room classes, lecture: response lecture, online lecture, online lecture: plenary exercises, online lecture: response lecture, online project, online seminar, online seminar: coached exercises, online seminar: practical PC room classes

Learning materials and price

A syllabus is available. Cost: 20 EUR

References

- M. Kutner, C. Nachtsheim, J. Neter, W. Li. 'Applied Linear Regression Models', 5th edition. McGraw-Hill Education, 2004.
- Agresti A. (2002). Categorical Data Analysis. New York: Wiley.
- McCullagh P. & Nelder J.A (1989). Generalized Linear Models. CRC press

Course content-related study coaching

Students are coached by assistants during PC-labs and regular exercice classes which prepare students for problem solving, the final analysis project and the open book exam. Through the electronic learning environment they can exchange questions and answers outside lecture hours among themselves and with lecturers. A take home project will provide the students with practical experience in data analysis. The students will receive coaching and feed back on this projects through organized sessions.

Evaluation methods

end-of-term evaluation and continuous assessment

Examination methods in case of periodic evaluation during the first examination period

Written examination with open questions, open book examination

Examination methods in case of periodic evaluation during the second examination period

Written examination with open questions, open book examination

Examination methods in case of permanent evaluation

Assignment

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible in modified form

Extra information on the examination methods

The end-of-term exam aims to assess if the student understands the basic theory of Statistical regression modelling of continuous and categorical outcome data. The exam consists of a critical understanding of theory and data analysis reports, with exercises that are also concerned with the design and analysis plan for a statistical study and the interpretation of given software output.
The assessment takes the form of a data analytic project. The written report on the take home project is graded.

Calculation of the examination mark

Periodic evaluation (50%) + non-periodic evaluation (50%). Students must pass both parts to pass the course. Partial results for which the student scored at least half of the points, can be transferred to the next examination period within the same academic year.
Partial results will never be rounded.