Statistical Methods

Unit code:	MATH20802
Credit Rating:	10
Unit level:	Level 2
Teaching period(s):	Semester 2
Offered by	School of Mathematics
Available as a free choice unit?:	N

Requisites

Prerequisite

- MATH10141 - Probability 1 (Compulsory)
- MATH20701 - Probability 2 (Compulsory)
- MATH10282 Introduction to Statistics (Compulsory)

Desirable

- Basic knowledge of the R statistical programming language

Aims

- to introduce the general principles of likelihood-based inference and testing for general models (i.e. for both discrete and continuous distributions)
- to offer a first overview of Bayesian statistical inference,
- to provide an introduction to linear regression models,
- and to demonstrate corresponding computational procedures in R, and
- to provide an introduction to data protection and privacy.

Overview

This module provides students with the methodological foundations in model-based statistical learning, in particular likelihood estimation and inference and simple linear regression models. The theoretical and methodological discussions are complemented by practical computer application. In addition key issues in data protection and privacy are discussed.

This module thus links the core level 1 module (Introduction to Statistics) and the optional theoretical and applied level 3 modules (Statistical Modelling, Statistical Inference, Extreme Values and Financial Risk, Time Series, Multivariate Statistics and Machine Learning, Medical Statistics).

Learning outcomes

On successful completion of the course students will be able to:

- apply model-based approaches in statistical data analysis
- derive maximum likelihood estimates and compute corresponding confidence intervals
- perform statistical testing from a likelihood perspective
- address simple inference problems from a Bayesian point of view

- analyse and fit linear regression models
- use R to apply these techniques on actual data
- understand data protection and privacy issues

Assessment methods

- Other - 20%
- Written exam - 80%

Assessment Further Information

- Coursework (1 in-class exam): weighting 20%
- End of semester examination: 2 hours weighting 80%

Syllabus

- Likelihood-based inference (7 lectures): likelihood function, maximum likelihood estimator (MLE), Fisher information, likelihood intervals, invariance principle, relationship to least-squares estimation (OLS) and entropy learning, (generalised) likelihood ratio tests.
- Bayesian learning (7 lectures): Bayes‘ theorem, prior and posterior probabilities, information update, credible intervals, properties of Bayes‘ estimators, shrinkage effect, Bayes factor.
- Linear regression (7 lectures): standard linear regression model, OLS/MLE estimation of regression coefficients and their variances, coefficient of determination, prediction intervals, testing of regression coefficients, variable selection.
- Data protection (1 lecture): discussion of data ethics, data protection and privacy, legal regulations (GDPR, CCPA)

Recommended reading

- Held and Bove. 2014. Applied Statistical Inference: Likelihood and Bayes. Springer.
- Wood. 2015. Core Statistics. Cambridge University Press.
- Gelman et al. 2014. Bayesian Data Analysis (third edition). Chapman and Hall/CRC
- Faraway. 2015. Linear Models with R (second edition). Chapman and Hall/CRC.

Feedback methods

Tutorials and computer labs will provide an opportunity for student to get feedback on their understanding from the instructor. The in-class test also offers an opportunity for students to receive feedback. Students can also get feedback on their understanding directly from the lecturer, for example during the lecturer's office hour.

Study hours

- Lectures - 24 hours (12 x 2 hours)
- Tutorials – 9 hours (9 x 1 hours) - 2 tutorials will be computer-based sessions.
- Independent study hours - 65 hours

Teaching staff

Korbinian Strimmer - Unit coordinator