Course syllabus for

Biostatistics, 6 credits

Biostatistik, 6 hp
This course syllabus is valid from spring 2012.
Please note that the course syllabus is available in the following versions:
Spring2012 , Spring2016

Course code	4BI085
Course name	Biostatistics
Credits	6 credits
Form of Education	Higher Education, study regulation 2007
Main field of study	Biomedicine
Level	AV - Second cycle
Grading scale	Pass with distinction, Pass, Fail
Department	Department of Medical Epidemiology and Biostatistics
Decided by	Programnämnd 7
Decision date	2011-11-25
Course syllabus valid from	Spring 2012

Specific entry requirements

Bachelors degree or professional qualification worth at least 180 credits in biomedicine, biotechnology, cellular and molecular biology or medicine

Objectives

The main aim of the course is to equip the participants with statistical concepts and tools for relating biological outcomes to multiple possible explanatory variables. Suitable tools and methods for two different settings are introduced: a) classic statistical modeling, with a relative small number of explanatory variables and a comparatively large number of subjects, leading to different kinds of multivariate regression and conventional experimental and observational studies; b) supervised and unsupervised statistical learning in situations with many more explanatory variables than subjects, leading to prediction and classification algorithms for high-dimensional data common in bioinformatics and suitable for modern high-throughput data. At the end of the course, students should be able to analyse a realistic data set from either of these settings independently. On completion of the course the student should be able to: Knowledge and understanding - explain the concept of random variation in biological phenomena as it relates to experimental and observational studies in research, - describe appropriate statistical methods to quantify random and systematic effects in complex biological data, - discuss the relevance of multiple testing for modern biological research, - discuss the distinction between explanatory and predictive modelling. Applying knowledge and understanding - choose and fit

multivariate regression models of intermediate complexity using a standard statistical software package, - choose and apply basic machine learning algorithms using a standard scripting language, - communicate the results in a manner suitable for oral presentation, technical reporting and scientific publication, - understand, discuss and evaluate critically corresponding findings of intermediate complexity in the relevant scientific literature. Making judgments - demonstrate the ability to weigh and integrate conflicting empirical evidence in the literature.

Content

The course content is organised in three distinct units, corresponding to three instructional periods: Part 1 is a brief recapitulation of the required basic material. Part 2 covers classical multivariate regression methods, and Part 3 deals with more algorithmic methods relevant for modern high-throughput biology and associated bioinformatics. Part 1: Basic concepts and methods Randomness of biological observations. Experimental and observational data. Types of data: nominal, ordinal, continuous variables. Data summary measures. Graphical representations. Concepts of probability and probability distributions. Parameter estimation: mean, proportion, standard deviation, standard error. Concepts of statistical inference: confidence intervals and hypothesis tests. Elementary parametric hypothesis tests. Univariate linear regression. Part 2: Statistical modeling and multivariate regression methods Multivariate linear regression and general linear model. Continuous and categorical predictors. Interactions. Model fitting and diagnostics. Generalised linear models and logistic regression. Survival analysis and Cox proportional hazard models. Part 3: Modeling and classification of high-dimensional data Multiple testing and types of errors. Control of errors, including false discovery rate. Feature selection. Distance measures and clustering. Classification and prediction algorithms. Validation and cross-validation.

Teaching methods

Student activities will be shared class-room activities (lectures, discussions) and practical activities in smaller groups (tutorials, computer labs, project work). The course ends with a written exam.

Examination

Written examination. One micro-project will be assigned during each part of the course. Having handed in all assigned projects successfully is required for sitting the final written exam. The final grade will be determined by the projects handed in (60%) and the written exam (40%). Compulsory attendance Attendance is compulsory during the introduction to the course, supervised tutorials and group work. Before the student has participated in all compulsory parts or compensated absence in accordance with the course director's instructions, the student's results for respective part will not be registered in LADOK. All compulsory parts need to have been completed in order to qualify for taking the written exam. Limitations of the number of examinations or practical training sessions A student who does not pass the examination on the first occasion is offered a maximum of five additional opportunities to sit the examination. If a student has not passed the examination after a total of four attempts then it is recommended that the student retake the whole course at the next opportunity. Following this the student is permitted to sit the examination on another two occasions. A student who fails the examination on six occasions is not permitted to sit the examination again or to retake the course. Participation in an examination is defined as an occasion on which a student attends an examination, even if the student submits a blank examination paper. If a student has registered to sit an examination, but does not attend the examination, this is not defined as participation in the examination.

Transitional provisions

After each course occasion there will be at least six occasions for the examination within a 2-year period from the end of the course.

Other directives

The course will be evaluated in accordance with the guidelines established by the Board of Education. The course language is English.

Literature and other teaching aids

Bland, Martin
An introduction to medical statistics

3. ed. : Oxford : Oxford University Press, 2000 - xvi, 405 s.
ISBN:0-19-263269-8 (hft.) LIBRIS-ID:4603394
Library search

Holm, Sture
Biostatistisk analys

1. uppl. : Lund : Studentlitteratur, 2008 - 327 s.
ISBN:978-91-44-05378-3 LIBRIS-ID:11173905
URL: Omslagsbild
Library search