



**Karolinska  
Institutet**

Course syllabus for

## **Biostatistics, 6 credits**

Biostatistik, 6 hp

This course has been cancelled, for further information see Transitional provisions in the last version of the syllabus.

Please note that the course syllabus is available in the following versions:

[Spring2017](#) , [Spring2018](#) , [Spring2019](#) , [Spring2020](#) , [Spring2021](#)

Course code	4BI101
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	Institute of Environmental Medicine
Decided by	Programme Committee 7
Decision date	2016-10-17
Revised by	Programme committee for study programmes in biomedicine
Last revision	2017-11-02
Course syllabus valid from	Spring 2018

## **Specific entry requirements**

A Bachelor's degree or a professional degree worth at least 180 credits in biomedicine, biotechnology, cellular and molecular biology or medicine. English language skills equivalent to English B at Swedish upper secondary school.

## **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: 1) 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; 2) 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.

Upon completion of the course the student should be able to:

Regarding 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 distinction between explanatory and predictive modelling.

Regarding competence and skills

- 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.

Regarding judgement and approach

- demonstrate the ability to weigh and integrate conflicting empirical evidence in the literature.

## Content

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.

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.

Quantile regression. Models for longitudinal data.

## Teaching methods

Teaching will be in the form of lectures and practical computer activities.

## Examination

The examination consists of written examination.

Compulsory participation

Attendance is compulsory during the introduction to the course and practical computer activities.

The course director assesses if and, in that case, how absence can be compensated.

Before the student has participated in all compulsory parts or compensated absence in accordance with the course director's instructions, the student's results will not be registered in LADOK.

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. 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 language is English.

Course evaluation will be carried out in accordance with the guidelines established by the Board of Higher Education.

Oral evaluation in the form of course council meetings will be carried out during the course.

## Literature and other teaching aids

### *Recommended literature*

*Altman, Douglas G.*

#### **Practical statistics for medical research**

London : Chapman and Hall, 1991 - xii, 611 s.

ISBN:0-412-38620-8 (hft.) LIBRIS-ID:8286190

[Library search](#)

*Bland, Martin*

#### **An introduction to medical statistics**

3. ed., [Nachdr.] : Oxford : Oxford University Press, 2009 - XVI, 405 S

ISBN:978-0-19-263269-2 LIBRIS-ID:11926588

[Library search](#)

*Dalgaard, Peter*

#### **Introductory statistics with R**

2. ed. : New York : Springer, cop. 2008 - xvi, 363 s.

ISBN:978-0-387-79053-4 (pbk. : alk. paper) LIBRIS-ID:11305121

[Library search](#)

*Pagano, Marcello; Gauvreau, Kimberlee*

#### **Principles of biostatistics**

2. ed. : Pacific Grove : Duxbury, cop. 2000 - xvi, 525 s. , [42] s.

ISBN:0-534-22902-6 ; No price LIBRIS-ID:5036554

[Library search](#)

*Rosner, Bernard*

#### **Fundamentals of biostatistics**

6. ed. : Belmont, Calif. : Thomson/Brooks/Cole, cop. 2006 - xx, 868 p.

ISBN:0534418201 LIBRIS-ID:9942420

[Library search](#)