Karolinska Institutet

SYLLABUS FOR DOCTORAL COURSE

Course number
5303

Title in Swedish
Biostatistik i praktiken (statistisk konsultation)

Title in English
The Practice of Biostatistics (a.k.a., Statistical Consulting)

Language
en

Credits
3.0

Responsible KI department
Department of Medical Epidemiology and Biostatistics

Specific entry requirements
The specific entry requirements are one of the following: 1/. At least 60 higher education credits (HEC) from 1st and/or 2nd cycle courses in the subjects: statistics, probability, mathematics, or data science, of which at least 20 HEC must be in statistics or probability; or 2/. A course in probability theory of at least 7.5 HEC at 2nd cylcle level, a course in statistical inference of at least 7.5 HEC at 1st or 2nd cycle level, and completed the half-time review of doctoral education at Karolinska Institutet or the equivalent at another university.

Grading
Pass/Fail

Purpose of the course
The primary target group for the course is students with a previous degree in mathematics, statistics, or a related area. Students in 1st and 2nd cycle programs in such fields typically receive limited or no training in the application of statistics in real-world collaborations with subject-matter experts. Most courses focus on mathematical theory or details of a specific statistical method: research problems are typically neatly translated into theoretical questions, data are usually simplified or simulated, and results are interpreted and communicated from a statistical perspective. Furthermore, statistical modelling and analysis are presented appear as abstract activities, with little or no regard for the practical, legal or ethical context in which they take place.

The purpose of this course is to fill this gap in the education of biostatisticians by introducing students to challenges they will face when working as a collaborative or consulting biostatistician in biomedical research: these include efficient communication with collaborators in order to understand and formulate research questions, thorough problem formulation prior to statistical analysis, organization and management of data flows, interpretation and communication of results to non-statisticians, but also legal frameworks and ethical dilemmas with which they will be confronted. The course will provide opportunities for both high-level discussion and reflection as well as practical skill training, exposing the students to a set of conceptual and technical tools for meeting these challenges.

Intended learning outcomes
After successfully completing this course, students will be able to:

- articulate an understanding of the collaborative nature of medical science and the role(s) of the biostatistical scientist in medical research,

- demonstrate via role-play the ability to communicate effectively with non-statisticians about statistical aspects of medical research,

- argue for the importance of reproducible research and demonstrate familiarity with common tools and procedures for documentation of a research project,

- explain the importance and benefits of structured procedures for collecting, managing, and archiving research data,

- summarise the relevant Swedish laws and articulate the responsibilities of researchers in using personal data in biomedical research,

- describe some of the common ethical dilemmas faced by practicing statisticians and contrast strategies to address them.

Contents of the course

The course will cover the following topics:

- overview of the research process and the role(s) of the biostatistician in this process, with special emphasis on problem formulation and planning prior to data analysis;

- role(s) of the biostatistician in a research collaboration, e.g., data manager, data analyst, collaborative scientist, and co-author, and the responsibilities associated with each of these roles;

- communication with clients and collaborators, with a particular focus on how to plan, structure, and execute effective meetings;

- overview of structured procedures for collecting, managing, analyzing, presenting, and archiving research data, including a brief overview of version control systems for source code and the FAIR principles for open science and reproducible research;

- overview of Swedish laws governing the use of personal data in biomedical research and the responsibilities of researchers, including an overview of the principles for processing personal data according to GDPR and examples of how to write data processing and data transfer agreements;

- ethical and moral dilemmas, e.g., when clients advocate inappropriate research practices such as p-hacking or harking, as well as strategies for dealing with them, including when and how to leave an unsuccessful collaboration, when and how to decline authorship or acknowledgment, and how to respond to offers of co-authorship in lieu of payment;

Teaching and learning activities

The course will run over one term (20 weeks), with one in-person teaching session each two weeks. Students will be provided reading material and questions to prepare for each of the in-person sessions. The course will include lectures, group discussions, oral presentation by students, and role play sessions where, for example, one student will play the role of the biostatistician and another student or a teacher will play the role of the client.

Compulsory elements

All group discussions, role plays, and seminars are compulsory. If a student misses one of these, they will be given an opportunity to compensate by presenting the topic for the teacher at a separate occasion.

Examination

To pass the course the student must show that the learning outcomes are achieved. This is judged by an assessment of the student's contribution in group discussions, performance when making oral presentations, and performance in the role-play sessions.

Literature and other teaching material

There is no mandatory course literature, however, the following texts may provide relevant background reading:

- Cabrera & McDougall (2001). Statistical Consulting, Springer.

- Chatfield (1995). Problem solving: A statistician's guide. Chapman & Hall.

- Christopher Gandrud (2015) Reproducible Research with R and R Studio. Chapman & Hall/CRC The R Series.

Notes

☐ (*) The course meets the requirements for a general science course.
☐ (#) Suitable for doctoral students without medical education at university level.
☐ A Laboratory Animal Science course.

Syllabus established by the Course and Programme Committee on behalf of the Board of Doctoral Education.

Date of decision	Date of revision

Course responsible

Paul Dickman
Paul.Dickman@ki.se
0852486186
Department of Medical Epidemiology and Biostatistics

Contact person

Gunilla Nilsson Roos (gunini)
gunilla.nilsson.roos@ki.se
08-524 822 93
Department of Medical Epidemiology and Biostatistics