Summary
This course covers topics in applied biostatistics, with an emphasis on practical aspects of data analysis using R statistical software. Topics include types of studies and their design and analysis, high dimensional data analysis (genetic/genomic) and other topics as time and interest permit.

Content
Types of studies
Design and analysis of studies
R statistical software
Reproducible research techniques and tools
Report writing
Exploratory data analysis
Linear modeling (regression, anova)
Generalized linear modeling (logistic, Poission)
Survival analysis
Discrete data analysis
Meta-analysis
High dimensional data analysis (genetics/genomics applications)
Additional topics as time and interest permit

Keywords
Data analysis, reproducible research, statistical methods, R, biostatistical data analysis, statistical data analysis

Learning Prerequisites

Required courses
This course will be very difficult for students with no previous course or experience with statistics. Previous experience with R is neither assumed nor required.

Recommended courses
Undergraduate statistics course

Important concepts to start the course
It is useful to review statistical hypothesis testing.
Learning Outcomes
By the end of the course, the student must be able to:
• Synthesize analysis into a written report
• Report plan of analysis and results obtained
• Justify analysis plan
• Plan analysis for a given dataset
• Interpret analysis results
• Analyze various types of biostatistical data

Transversal skills
• Write a scientific or technical report.
• Assess one’s own level of skill acquisition, and plan their on-going learning goals.
• Take feedback (critique) and respond in an appropriate manner.
• Use a work methodology appropriate to the task.

Teaching methods
Lectures and practical exercises using R. Typically, each week covers an analysis method in the lecture and then the corresponding exercise session consists of an R practical showing how to implement the methods using R. In each practical, students use R to carry out analyses of the relevant data type for that week.

Expected student activities
Students are expected to participate in their learning by attending lectures and practical exercise sessions, posing questions, proposing topics of interest, peer reviewing of preliminary reports, and interacting with teaching staff regarding their understanding of course material. In addition, there will be a number of short activities in class aimed at improving English for report writing.

Assessment methods
Evaluation is based on written reports of projects analyzing biostatistical data.

Supervision
Office hours Yes
Assistants Yes
Forum Yes

Resources
Virtual desktop infrastructure (VDI)
No

Bibliography
To be provided during the course.
Pre-recorded lectures (videos) will also be provided.

Moodle Link
• https://go.epfl.ch/MATH-493