Statistical Methods for Bioinformatics [I0U31a]

Part I: General information

Next

Part I:

Objectives:

The students will get familiar with the techniques covered in the Teaching Plan (see below) and will be able to apply them to solve real life problems in an independent way

Teaching Plan:

- Lecture 1: Linear regression and correlation
- Lecture 2: Generalized linear models: Logistic Regression. Model building and model selection using Akaike Information Criterion (AIC)
- Lecture 3: Multilevel Models: Longitudinal data
- Lecture 4: Multilevel Models: Cluster data
- Lecture 5: Missing data
- Online Q&A session.

Evaluation:

The evaluation will be divided in a practical project and an exam. The total score for part I (20 points) will be divided in the following way

- *Project*: 4 points
- *Exam*: 16 points

Project:

- Students will do a project working in tutorial groups. The list with the groups is available on Toledo.
- After the first lecture students should check that they are assigned to one and only one group and they should inform Prof Alonso if there is any problem.
- Once the groups have been checked the project will be assigned.
- The assignment is on Toledo.
- Each tutorial group have to write a report with a detailed discussion of the analysis. A report should contain no more than 3 sheets in total, including the title page, i.e., if you print your report it should not exceed 3 sheets two sided. Please use an A4 page format, a times new roman 12 font and a 1.2 spacing between lines.
- Each tutorial group send an electronic copy of the report of the project, as well as, the R code used in the analysis to Prof. Ariel Alonso Abad maximum 5 days after the last lecture (01/04/2025 before 23:00) via email. Please send the report and the R code as

(01/04/2025 before 23:00) via email. Please send the report and the R code as independent files. The subject of the email should be "Bioinformatics project report group XX" where XX denotes the number of the group.

- The title page of the report **should contain the number of the group and a list with the names and student numbers of all the members of the group**. Please note that these reports are part of your evaluation, thus follow theses instructions very carefully.
- If you have to do a second chance exam then the grade of the project will be carried forwards to your second chance exam.

Exam:

The exam will be written and will have a methodological and practical part. The methodological part will consist of multiple choice questions and the practical part will consist of a small practical assignment (analysis of small data set) and will take part in the computer room.

Text books:

Applied Linear Statistical Models, 5th Edition, Kutner et al. (2005)

Applied Longitudinal Data Analysis: Modeling Change and Event Occurence, 1st Edition, Singer and Willett. (2003)

Software:

The R package will be the main software package. Codes will be discussed in the lectures

Lecturer:

Prof.dr. Ariel Alonso Abad

L-BioStat: Leuven Biostatistics and statistical Bioinformatics Centre. KULeuven

Kapucijnenvoer 35 - blok D (G01.56)

3000 LeuvenEmail: ariel.alonsoabad@kuleuven.be (mailto:ariel.alonsoabad@kuleuven.be)

tel. +32 16 32 58 44

of +32 16 3 73389

fax +32 16 3 37015