

Data Analysis

Instructors: I.NTZOUFRAS

Course Code: 61104

Core Course, 1st semester, 7.5 ECTS units

Course level: Graduate (MSc)

Language: English

Course Description

Primary aim of this course is the understanding and the application of statistical method in real life problems of various scientific fields such as Management, Marketing, Psychology, Medicine, Sports and Social Sciences. Focus is given on the review of parametric and non-parametric hypothesis tests for one and two samples (t-tests και Wilcoxon tests), analysis of variance and regression models. Emphasis is given in the implementation of all methods using R and in problem solving. Interesting real-life datasets and problems are analyzed during this course with aim to provoke their attention and motivate them. Finally, the students are introduced to the basic principles of scientific report writing and story telling either in the form of a written report or in form of oral presentation.

Prerequisites

Students should have good knowledge of estimation and statistical inference. It is also useful to have basic knowledge of the statistical language R and to be familiar with the statistical theory of regression.

Target Learning Outcomes

Upon completion of the course, students will be able to:

- 1) To manipulate and analyze data in R
- 2) To perform basic hypothesis tests
- 3) To build and interpret regression models

To write statistical reports in a professional way.

Recommended Bibliography

- Diez, D., Barr, C., & Cetinkaya-Rundel, M. (2012). *OpenIntro statistics* (Second. Edition). Free Open Book; available at <http://www.openintro.org/stat/textbook.php>
- Fox J. & Weisberg H.S. (2011). *An R Companion to Applied Regression*. 2nd edition. SAGE Publications Inc.
- Faraway, J. (2002). *Practical regression and ANOVA using R*; available at <http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf>

- Ntzoufras I. & Karlis D. (2015). Introduction to programming and statistical analysis with R. Athens: Association of Hellenic Academic Libraries. <http://hdl.handle.net/11419/2601>, ISBN: 978-960-603-449-7.
- Fouskakis D. (2013). Data Analysis Using R. Tsotras Publications. Athens. (Book code in Eudoxos: 33134029).
- Field A, Miles J and Field Z. (2012). Discovering Statistics Using R. Sage Publications. Translated into Greek edition (2021): Discovering Statistics Using R. Propompos Publications.

Teaching and Learning Activities

- Introductory motivational talk about the value and the fun part of Statistics.
- Introductory videos (for R, for interpretation of Statistics by David Spiegelhalter, for the necessity of statistics in our daily life).
- Teaching in a classroom and computer labs.
- Laboratory exercises conducted during an extended lab session.
- Online game Kahoot (all together in the room - online version and asynchronously - offline).
- Guess the correlation game.
- Introductory course in R

Assessment and Grading Methods

- 25% project/assignment accompanied with long scientific report
- 25% project/assignment on high dimensional problem accompanied with short scientific report and oral presentation (focus is given on the latter)
- 50% Written examination (mostly multiple choice with 3-4 open questions)
- Three (3) optional lab assignments (small size)
- One (1) optional R exercise