Statistical Methods for Data Analytics (m62111p)

Instructors: ST. PSARAKIS

Core Course, 1st semester, 5 ECTS units

Course level: Graduate (MSc)

Language: Greek (or English if foreign students attend)

Course Description

This course has two parts. In the first one the basic theory of confidence intervals and hypothesis testing is presented while in second part the basic theory of statistical models is examined.

In the first part the basic theory of the construction and interpretation of confidence intervals for means, proportions and variances for one and two populations both for normal and non-normal data is presented. The basic theory statistical hypothesis testing for population parameters such as means, proportions, variances as well comparison of parameters of two populations the level of significance, the p-value, the power of the test and the determination of sample size is examined.

In the second part the course introduces and presents the fundamental theory of statistical models, methods, and techniques, which are necessary in the research and empirical data analysis. The theory of regression models, single and multiple linear regression, is presented. The variable/model selection problem, the use of dummy variables, and the problem of multicollinearity are examined. Emphasis is given on the application of the theory, estimation of the model parameters, examination of the assumptions of residuals using diagnostic tests, and the interpretation of results. The theory and empirical application of the analysis of variance are introduced and presented in detail. The underline theory, methods and models are implemented to empirical data and problems using the statistical package R.

Prerequisites

Applied Probability – Estimation.

Target Learning Outcomes

The aim of this course is to provide students with the learning of using appropriate statistical methods, models and techniques required for data analysis. After successfully completing the course, students will be able to:

- Know and apply the appropriate methodology, to construct confidence intervals that will contain the unknown parameters with the desirable probability
- Perform hypothesis testing in various problems
- Learn the fundamentals in statistical inference allowing them to understand which type of analysis is necessary and how it can be correctly implemented
- Estimate the parameters of statistical models
- Conduct hypothesis testing and construct confidence intervals for model parameters
- Estimate regression models, construct predictions and interpret the results of the analysis appropriately
- Apply, using the R package, statistical models to empirical problems and applications

Recommended Bibliography

- Draper N, Smith H (1998) Applied Regression Analysis 3rd Edition Wiley
- Montgomery D (2012) Introduction to Linear regression Analysis, 5th Edition Wiley.
- Montgomery D. and Runger GC (2018) Applied Statistics and Probability for Engineers 7th Edition Wiley
- Weisberg, S. (2005). Applied Linear Regression, 3rd edition, Wiley
- Fox, J., and Weisberg, S. (2011). An R Companion to Applied Regression, 2nd edition, SAGE Publications Inc.

Teaching and Learning Activities

One three-hour lecture per week and study exercises as homework (some to be submitted).

Assessment and Grading Methods

90% Written examination. 10% project/assignment based on simulated data applying the methodologies and techniques described during the course accompanied with short scientific report.