Statistical Data Modeling (62105)

Instructors: V.VASDEKIS

Core Course, 2nd semester, 5 ECTS units
Course level: Graduate (MSc)
Language: Greek

Course Description

The principles of statistical modeling, experimental design, and the use of regression techniques for data following an exponential family distribution are presented. The R language is used for statistical modeling. Regression techniques for analyzing binary or binomial data, Poisson data, and data exhibiting overdispersion are introduced. Models for continuous data, categorical data with multiple categories and ordinal variables are also discussed. Emphasis is placed on predicting new observations and the ability to separate subpopulations as supervised learning techniques.

Prerequisites
Students should have basic knowledge of mathematical calculus and probability theory.

Target Learning Outcomes
Upon successful completion of the course, students are expected to understand if the nature of their data allows application of a generalized linear model (knowledge and understanding). They should also be able to define the appropriate generalized linear model to the data at hand (application). They should be able to fit this model, interpret the results of analysis and provide predictions of future observations (skill). Finally, they should be able to explain to scientists of other disciplines the results of their analysis (synthesis).

Recommended Bibliography
- Agresti (2013). Categorical data analysis, Wiley Atkinson (1985). Plots, transformations and regression, Oxford university Press
- Carroll and Ruppert (1988). Transformation and weighting in regression, Chapman and Hall
- Chatterjee and Price (1977). Regression analysis by example, Wiley.
- Christensen R. (1998). Analysis of variance, design and regression. Chapman and Hall.
- Collett, D. (1991) Modelling Binary data, Chapman and Hall
- Cook and Weisberg, S. (1982). Residuals and Influence in regression, Chapman and Hall
- Dobson, A., Barnett, A.G (2008). An introduction to generalized linear models, Chapman and Hall.
- Draper and Smith (1981). Applied regression analysis, Wiley.
- McCullagh, P and Nelder, J.A. (1989) Generalized Linear Models, Chapman and Hall.
- Montgomery, D.C. (1989) Design and Analysis of Experiments, Wiley
- Montgomery, D.C., Peck, E.A. and Vining, G.G. (2001). Introduction to linear regression analysis. Wiley.
- Ryan (1997). Modern regression methods, Wiley.
- Weisberg, S. (1985) Applied Linear Regression, Wiley
- Venables W.N. and Ripley B.D (1999) Modern Applied Statistics with S-Plus, Springer

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

Assessment and Grading Methods

The final grade is the weighted average of the final examination grade (70%) and the grade of the study exercises to be submitted (30%).