Dipartimento di Economia

ROMA TRE UNIVERSITÀ DEGLI STUDI

Statistical methods in economics (60h – 9CFU)

Prof. Caterina Conigliani
caterina.conigliani@uniroma3.it

Course learning objectives and skill acquisition

The main objective of the course is to provide the fundamental tools for the application of statistical methods to the analysis of economic data. The theoretical part will be supported by an applied part devoted to the analysis of real data sets by means of the software R. A student that has completed the course should be practiced in the application of advanced statistical methods, should be able to interpret the results of a statistical analysis, and should be aware of limitations and possible sources of errors in the analysis.

Assessment

Attending students will develop and discuss a short dissertation (January/February only). For non-attending students the course assessment will be based on a written exam held in the computer lab, that will involve the analysis of different data sets using the methods and models studied during the course.

Course general schedule

Part 1: Introduction to data analysis and exploratory techniques

- Cluster analysis
- Principal component analysis

Part 2: Normal linear regression and its generalizations

- Polynomial regression
- Multiple regression
- Logistic regression
- Beta regression
- Poisson and negative binomial regression
- Spatial regression models

Part 3: Panel data analysis

- Modeling the level of a dependent variable
- Modeling the change of a dependent variable
- Fixed effects and random effects models for categorical variables and continuous variables
- Spatial regression models for panel data

Teaching material

Teaching material (slides, data sets, labs) will be available to students in a dedicated OneDrive folder.

Textbooks

Chatterjee, S. and Hadi, A.S. (2012), Regression Analysis by Example, 5th Edition, Wiley. Chapters: 1, 2, 3 (excluding 3.9), 4 (excluding 4.3, 4.9.2, 4.9.3, 4.10, 4.12, 4.13, 4.14), 5 (excluding 5.6 and 5.7), 6 (excluding 6.6 and 6.7), 9, 11, 12 (excluding 12.8.3 and 12.8.4), 13 (excluding 13.5, 13.6, 13.7).

Fox,J. and Weisberg, S. (2010), An R companion to applied regression, 2nd Edition, SAGE publications Inc.

Andreb, H-J, Golsch, K., Schmidt, A.W. (2013), Applied panel data analysis for economic and social surveys, Springer. Chapters: 1, 2, 3, 4.

Additional readings

Elhorst, J.P. (2010). Applied Spatial Econometrics: Raising the Bar, *Spatial Economic Analysis*, 5:1, 9-28.

Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R. *Journal of Statistical Software*, *34*(2), 1–24.

Azzalini, A. & Scarpa, B. (2012). Data Analysis and Data Mining: an Introduction. *Oxford University Press*.

Detailed teaching agenda

Lecture #1: Introduction, practical information, data collection of participants.
Lecture #2: An introduction to exploratory data analysis and statistical inference.
Lecture #3: Lab 1: an introduction to the software R.
Lecture #4: Lab 2: importing data into R.
Lecture #5: Cluster analysis: agglomerative hierarchical clustering.
Lecture #6: Cluster analysis: parametric mixture models.
Lecture #7: Lab 3: cluster analysis with R on the OECD Better Life Index data.
Lecture #8: Principal components analysis: theory and practice.
Lecture #9: Principal components analysis: interpretation and use of the components.

Lecture #10: Lab 4: principal component analysis with R on the OECD Better Life Index data.
Lecture #11: Simple regression analysis: OLS estimates, Gauss Markov theorem, the analysis of the residuals.
Lecture #12: Simple regression analysis: prediction, omitted-variable bias.
Lecture #13: Lab 5: simple regression analysis with R.
Lecture #14: Multiple regression analysis: the analysis of collinear data, the use of qualitative predictors.
Lecture #15: Multiple regression analysis: detecting outliers and influential observations.
Lecture #16: Lab 6: multiple regression analysis with R.
Lecture #17: Non linear regression functions: models with variable transformations (polynomial regression, log-level, level-log, log-log models) and models with interactions.
Lecture #18: Lab 7: non linear regression analysis with R.
Lecture #19: Logistic regression models for dichotomous responses.
Lecture #20: Lab 8: non linear regression and logistic regression with R (telekom data).
Lecture #21: Beta regression models for rates and proportions.
Lecture #22: Poisson and negative binomial regression models for count data.
Lecture #23: The likelihood function: likelihood ratio tests and AIC.
Lecture #24: Spatial regression models: SEM, SAR, SLX, SAC, SDM, SDEM. The Moran-I test on the residuals. The spatial Hausman test.
Lecture #25: Panel data analysis: modelling the level of a continuous response. FE and RE models. Hausman test.
Lecture #26: Lab 9: Panel data analysis with R: the analysis of CO2 emissions.
Lecture #27: Panel data analysis: modelling the change of a continuous response. FD models.
Lecture #28: Panel data analysis: modelling the level of a dichotomous response and of a discrete response.
Lecture #29: Lab 10: Panel data analysis with R: the analysis of the number of patent applications.
Lecture #30: Spatial regression models for panel data.