Statistical Methods for Economics

Coordinator Alessia Naccarato
Economic Department, Roma Tre University

The lessons are structured in 4 modules: Statistical Methods for Econometrics (A. Naccarato), Applied Econometrics with R (D. Zurlo), Applied Economics with STATA (G. D'Agostino), Geographic information systems and spatial analysis (F. Benassi).

Statistical Methods for Econometrics

Alessia Naccarato

The participation in the course requires some basic knowledge that can be acquired either through individual study or by participating in to the first 12-13 lessons (24-26 hours) of the Statistical Methods for Econometrics - course for Master Degrees, from 24th September to 20th October.

Ph.D lessons will take place from November 2018 to January 2019 for a total of 22 hours.

The exam will take place in March 2019 and consists of an oral interview on the subjects of the course.

Basic Knowledge I – References to statistical inference

Random Variables, Distribution Functions, Density Functions, Expected Values, Moments, Conditional Distributions and Independence, Covariance and Correlation, Basic Concepts of Random Samples, Sampling from the Normal Distribution, Methods of finding Estimator, Methods of Evaluating Estimator, Hypothesis Testing.

Basic Knowledge II – Review on linear regression

Classical Linear Regression: Descriptive Linear Regression, Classical Linear Regression Model, Statistical Inference in Classical Linear Regression.

Basic Knowledge III - Extensions of linear regression

Functional forms, Analysis of variance, Nonspherical disturbances: heteroskedasticity and autocorrelation, Multicollinearity, Variables selection, Stability of the regression function, Use of auxiliary information.

Course Programme

Simultaneous Equation Models - *System of linear regression*: The Model, Assumptions on error components and parameters' estimation. *Seemingly Unrelated (SUR)*: The Model, Estimation problems and Empirical Generalized Least Squares (EGLS).
Simultaneous Equation Model: Excursus on conditions for OLS consistency, Structural and Reduced Form equations, Identification and structural parameters, Rank and order conditions for structural parameters identification. *Structural parameters estimation*: Consequences of order conditions, Exact identification: Indirect Least Squares Estimators (ILS), Over identification: Limited and Full Information methods, Two Stage Least Squares estimator (2SLS), Consistency of 2SLS, Three Stage Least Squares (3SLS).
Time Series Analysis – First-Order Difference Equations, pth-Order Difference Equations, Lag Operators, Stationary ARMA Processes, Forecasting, Maximum Likelihood Estimation, Spectral Analysis, Vector Autoregression.

Introduction to Panel Data Models – Fixed effects models: Least squares dummy variable fixed approach, Random effects models: estimation of variance-components models The appropriateness of fixed effects and random effects estimation.
Introduction to Principal Component Analysis – Principal components, Choosing the number of principal components, Justification of the principal component MLE, Multidimensional scaling.

Books
Statistical Inference, G. Casella, R. L. Berger, 2nd Edition, Duxbury Advanced Series
Matrix Differential Calculus with applications in statistics and econometrics, J. R. Magnus, H. Neudecker, Wiley Series in Probability and Statistics
Econometric Analysis, W. H. Greene, 5th Edition, Prentice Hall
Time Series Analysis, J. D. Hamilton, Ed. Princeton
Analysis of Panel Data, C. Hsiao, Ed. Cambridge
New Introduction to Multiple Time Series Analysis, H. Lutkephol, Springer

Applied Econometrics with R

Davide Zurlo

Ph.D lessons will take place from December 2018 to January 2019 for a total of 12 hours.

The exam will take place in March and consists of an application to actual data on the subjects of the course.

Course Programme

Introduction to the statistical software R – Basic commands, Cycles if else and for, Data Structures, Importing datasets, Graphs
Econometrics with R – multiple regression model,OLS and ML estimators, model selection and use of dichotomous variables, test on autocorrelation, heteroschedasticity, structural breaks, multicollinearity, presence of outliers, endogenity. FGLS and IV estimators, dynamic linear models. SUR linear equation systems and structural equation systems, EGLS estimators, 2SLS, 3SLS. Introduction to panel data models, fixed-effect estimator random-effect estimator.

Books
Introductory Statistics with R, Peter Dalgaard, Springer, 2008.
Applied Econometrics with R, Christian Kleiber and Achim Zeileis, Springer, 2008

Applied Economics with Stata

Giorgio D'Agostino

Ph.D lessons will take place from November 2018 to December 2018 for a total of 18 hours.

The exam will take place in March and consists of an application to actual data on the subjects of the course.

Course Programme

Introduction to the statistical software STATA – Reading data into STATA, Import data, Data examination, Creating and manipulating variables, Basic graphs, Saving the data, Advanced data management

Statistical analysis with STATA – The simple linear regression model, Interval estimation and hypothesis testing, Prediction, goodness-of-fit and modelling issues, Extended example – Macroeconomic data. The multiple regression model, Return objects, Global and local macros, Save and use the estimated results, Advanced graphic capabilities, Extended example – Microeconomic data. Using indicator variables, Extended example – Microeconomic data. Panel data models.

Geographic information systems and spatial analysis

Federico Benassi

Ph.D lessons will take place from November 2018 to February 2019 for a total of 18 hours.

The exam will take place in March and consists of an application to actual data on the subjects of the course.

The course provide an introduction to the use of open source software for the representation and analysis of spatial data, i. e. all those data that have a geographical component. In particular, demographic and socio-economic data will be used at different geographical scales of analysis.

Introduction – work with geo-referenced data, why? what does it mean theoretically and operationally? the different types of geo-referenced data, geo-referred data sources, the foundations of the spatial thinking, type and short classification on computer applications and methodologies for elementary and advanced processing of this type of data, an overview on the statistical measures and concepts (contiguity, distances, neighborhood, local quotients, global and local spatial autocorrelation indices etc.)

Description and representation. Software: Quantum Gis (Qgis) – brief description of Qgis, importing geographical data in Qgis, import table data in Qgis, the "join" function and monitoring of results, building a simple thematic cartography in Qgis, how to manage layers, scales, printing area (geographic scale, north arrow, legend), how to export the file and manage it as an image for playback to other files, the management of the table fields in Qgis (the filter, creates, and deletes functions), construct an advanced thematic cartography in Qgis (management of multiple layers), export a SHP file to different formats for management on other applications.

Spatial analysis methods and techniques: global and local spatial autocorrelation. Software: GeoDa – brief description of GeoDa, the spatial contiguity matrix, conceptual and logical differences among global and local spatial indices, global spatial autocorrelation

analysis using the Moran index, local spatial autocorrelation analysis using local indices (Lisa, G and G*), management of the local results, import into Qgis for the production of advanced classification maps.

Geographically Weighted Regression Model – How to perform a local form of linear regression to model spatially varying relationships with GWR4.0

References bibliography

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis*, 27, 93-115.

Champion T. e Hugo G. (2004). New forms of urbanization. Beyond the urban-rural dichotomy, Ashgate, Adelrshot, England.

Dueker K.J. (1979). Land Resource Information Systems: A Review of Fifteen Years of Experience, *Geo-Processing*, 1, 105-128.

Duncan, O.D., & Duncan B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20, 210-217

Gehlke, C. E.; Biehl, Katherine (1934). "Certain effects of grouping upon the size of the correlation coefficient in census tract material". *Journal of the American Statistical Association*. 29 (185A): 169–170.

Isard W. (1960). Methods of regional analysis: an introduction to regional science. The MIT Press, Cambridge.

Janelle, D. G. and M. F. Goodchild (2011). Concepts, Principles, Tools, and Challenges in Spatially Integrated Social Science. In Nyerges, T.L., H. Couclelis, and R. McMaster (Eds.) The Sage Handbook of GIS & Society. Sage Publications. Pp. 27-45.

Kemp, K. (2008). Encyclopedia of Geographic Information Science, SAGE, p.146-147

Masser I. (1998). Governments and Geographic Information, Taylor & Francis, London

Matthews S.A., Parker D.M. (2013). Progress in Spatial Demography, *Demographic Research*, 28, 10, 271-312

Miller, H.J. (2004). Tobler's First Law and Spatial Analysis, *Annals of the Association of American Geographers*, 94(2), 2004, 284–289

Moran, P.A.P. (1948). The interpretation of Statistical Maps. *Journal of the Royal Statistical Society*, 243-251.

Morrill, R.L. (1991). On the measure of geographic segregation. «Geography Research Forum», 11, p.25-36

Openshaw, S. (1983). The modifiable areal unit problem. Norwick: Geo Books.

Romei P., Petrucci A.(2003). L'analisi del territorio. I Sistemi Informativi Geografici. Carocci Editore.

Tobler W. (1970) A computer movie simulating urban growth in the Detroit regio. *Economic Geography*, 46(2): 234-240.

Wong, D. W.S. (1993) Spatial indices of Segregation. *Urban Studies*, 30(3), 559-572

http://docs.qgis.org/2.18/it/docs/user_manual/ (Manuale utente Qgis)

http://geodacenter.github.io/documentation.html (GeoDa documentation)