UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Generalized Linear Model

2627-1-F8205B003-F8205B003-1

Learning objectives

** Knowledge and understanding **
This course will provide knowledge and understanding of:

linear models
GLS models
multivariate models
multilevel models

** Ability to apply knowledge and understanding **
At the end of the course, students will be able to:

Diagnose and resolve violations of the OLS model assumptions
Apply templates for data with hierarchical structure
Apply multivariate models (with more dependent variables)
Use the main SAS and R procedures dedicated to linear, GLS, multivariate and multilevel models

Contents

The aim of the course is to present linear model extending the classical ordinary least squares model. The main topics are

- generalized linear models
- multivariate models
- multilevel models

Detailed program

The course aims at introducing at the specification, estimation and verification of the interpretative advanced linear models compared to the classical linear model. It also presents:

· Generalized linear models that do not meet the assumptions of the classical linear model: heteroschedastic models, autoregressive models, non-linear models, models with outliers, GLS

· Multivariate linear models: from least squares models to seemingly unrelated regressions

· Multilevel models: hierarchical data and fixed effects anova, mixed models (random slope, random intercept)

Each area will be the specific object of a course module. The course activity comprises theoretical lecture and lab activity with SAS and R. The material of the course (both the theoretical lessons both practical lessons) and additional information will be posted on the web page in the e-learning platform unimib: http://elearning.unimib.it/.

Prerequisites

It is requested a good knowledge of:

Univariate descriptive statistics: position indeces; variability indices: symmetry and kurtosis indices.

Bivariate descriptive statistics: connection, average dependence, linear correlation,linear bivariate,

Multiple, multivariate, polynomial, non-linear regressions.

Probability theory: population and sample; probability in the classic version; combinatorial calculation elements; sampling types; distributions of univariate random
variables; random variables Normal, t of Student, F d Snedecor; random sampling distributions

Inference:estimation theory, property of the punctual estimators; interval estimation;hypothesis tests: general theory, Neyman Pearson hypothesis tests,
hypothesis tests on mean (Normal t of Student) and variance.
Basics of matrix algebra
Therefore student that do not have these previous knowledge in statistics are requested to attend biostastistics courses and pass exams of
of: probability calculation, introduction to stastistical inference, introduction to statistical models, statistical models for data categorical . The students also have to know the R or SAS statistical packages.

(in the three-year degree or in the degree course of biostatistics) exams of: univariate and bivariate statististics, probability calculation, introduction to stastistical inference, introduction to statistical models, statistical models for data categorical . The students also have to know the R and SAS statistical packages.

Teaching methods

The course includes theoretical lecture and labs. In theoretical lecture, the methodological frameworks related to the course are presented and then applied during the practical labs. In the lab, SAS and R are going to be used with the aim to code and interpret model outputs. Lessons and exercises will be recorded on the e-learning platform.

Assessment methods

The test takes place in the laboratory. The student must answer two theoretical questions from a set of predetermined questions that he/she will already know at the beginning of the course. It is necessary to argue the answer in understandable and comprehensive terms by reporting the required demonstrations. The reference point for answers is the slides and the handout: of course, knowledge gained from the recommended books can be reported. Formulas and graphs should be reported: if it is difficult, they can be written on paper with pen and then scanned. The required length of the answers will depend on the question: answers not exceeding four typed sheets in 12-gauge spacing 1.5 (12000 characters including spaces) are suggested.
The second part of the test will consist of a practical exercise on real or simulated data provided by the lecturer through the use of statistical packages. The statistical tools to be used will be those learned in the course. In the paper all graphs and outputs should be appropriately commented on, both from a theoretical point of view and with respect to the application under examination. The development is done through the R environment. The student may use the exercise codes during the exam. These codes will be provided on the day of the test.

Textbooks and Reading Materials

-All the reading material is included in the course digital textbook uploaded on the elearning platform. The textbook covers both theoretical topics and practical examples.

Suggested readings

- Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Cengage learning.
- Freund, R. J., Wilson, W. J., and Sa, P. (2006), Regression Analysis: Statistical Modeling of a Response Variable, 2nd edition, Academic Press
- Baltagi B. H. (2008), Econometrics, fourth Edition, Springer Berlin

-Rencher , A. C. ,Methods of Multivariate Analysis, Wiley

-Tom Snijders, T., Bosker, R., Multilevel Analysis: An Introduction To Basic And Advanced Multilevel Modeling, SAGE Publictaions Inc.

- Littell, R. C., Freund, R. J., and Spector, P. C. (2002), SAS for Linear Models, 4th Edition, Cary, NC: SAS Institute Inc.

Manual SAS/STAT 15.1 - Faraway, J. J. (2004). Linear models with R. Chapman and Hall/CRC.

Semester

Second semester, third period

Teaching language

Italian

Sustainable Development Goals

QUALITY EDUCATION