UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Introduction To Statistical Models

2526-1-F8205B005

Learning objectives

The course deals with the fundamental methodological and applied aspects of statistical modelling, according to both the descriptive and inferential perspectives. The main topics are the simple and multiple linear regression model for population and sample data and their analysis with the SAS software. The course is addressed expressly to three-year-degree graduates without previous education in statistics or quantitative methods.

Knowledge and understanding. This teaching will provide knowledge and understanding concerning:

- the principal methods for the study of the dependency relations of a quantitative variable with one or more quantitative variables, possibly together with qualitative variables, and the construction of the pertaining linear modelling in both descriptive and inferential contexts
- the application of the primary methods for bivariate analysis through numerical exercises to be solved with the pocket calculator (i.e., without the aid of any statistical software)
- the logic and working of the SAS software and its use to carry out bivariate analyses and set up linear regression models along with the relative graphical representations
- the reading and interpretation of the analysis outputs produced by the SAS software.

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

- apply the basic methods of bivariate analysis, establish the correct role for the variables (i.e., whether dependent or independent), and build the most suitable regression models according to the purposes of the analysis and the nature of the available data
- select a linear regression model that is both parsimonious and of fair goodness of fit through descriptive or inferential statistical criteria
- interpret analysis results critically and identify any margin for improvement in the construction of the basic models applied according to either an interpretative-descriptive or a predictive perspective
- import external data files of various sources and formats into SAS and use the basic SAS syntax autonomously.

The course allows the student to acquire the main theoretical and applicative bases relating to the bivariate analysis and linear regression models necessary in any working context where data files are used and for advancing university studies.

Contents

Mean and linear dependence relationships, mathematical versus statistical interpolation, regression function. Simple and multiple linear regressions for population data and sample data. Regression model selection using descriptive and inferential criteria. Analyses of empirical cases with the SAS software.

Detailed program

- Introduction to the principal concepts of dependence for quantitative population data: Mean dependence and linear dependence. Mathematical interpolation versus statistical interpolation. Regression function. Least-squares fitting criterion
- Linear regression: Least-squares line, goodness-of-fit, extension to more than one explicative variable, multiple and partial correlations, qualitative regressors
- Simple and multiple linear regression models for sample data: Normal multivariate distribution and its properties, model specification, assumptions, least-squares and maximum likelihood estimation methods, confidence intervals, testing statistical hypotheses
- Multiple regression model selection: Descriptive and inferential criteria
- Analyses of empirical cases with the SAS software

Prerequisites

Knowledge of the topics covered in undergraduate Statistics and Statistical Inference courses is recommended.

Teaching methods

Theoretical lectures in the classroom and practical exercises in the statistical-informatics laboratory with the SAS software.
All lectures will be delivered in face-to-face mode. Asynchronous video recordings of both theory lectures and laboratory exercises will be made available.

Assessment methods

The exam consists of a written test (total duration: 2 hours; language of the exam text: Italian) with three questions (divided into several points) that deal with both the theoretical and applicative aspects of the statistical models covered in the course. The theoretical questions concern the methodological aspects of the topics covered in the course and aim at verifying the theoretical knowledge acquired on the logic and essential aspects underlying

population data analyses or sample data analyses, bivariate analyses, formal specification of regression models, determination or estimation of the parameters, assessment of model goodness-of-fit, and drawing of statistical inference in a modelling perspective. They also aim to verify the ability to use the symbolic-formal statistical language autonomously, provide definitions appropriately and prove the main theoretical results analytically. The practical questions concern the transposition of the theory into the practice and aim to verify the ability to comprehend and apply the theory and read and interpret the analysis results obtained with the SAS software. Furthermore, the exam in written form allows verifying the ability of expression through proper use of the statistical technical language.

The oral exam is optional (on request by the professor or student) and covers theoretical and practical topics. Access to the oral test is subject to passing the written test with a mark of at least 18/30. It should be noted that the oral test may involve either the increase, the maintenance, or the decrease in the evaluation achieved in the written test.

Given the abundance of teaching material uploaded on the e-learning platform of the course, no distinction is made between exams for attending students and exams for non-attending students. Finally, there is no ongoing test.

Textbooks and Reading Materials

- Teaching material uploaded on the course e-learning website (restricted access)
- Delwiche, L.D., Slaughter, S.J. (2012), The Little SAS Book: A Primer, SAS Institute
- Freund, R. J., Wilson, W. J., and Sa, P. (2006), Regression Analysis: Statistical Modeling of a Response Variable, 2nd edition, Academic Press
- Johnston, J. (1993), Econometrica, 3rd edition, Franco Angeli, Milano
- Littell, R. C., Freund, R. J., and Spector, P. C. (2002), SAS for Linear Models, 4th Edition, Cary, NC: SAS Institute Inc.
- Piccolo, D. (2010), Statistica, Il Mulino, Bologna
- Spencer N. (2004), SAS Programming - The One-Day Course, 1st Edition, Chapman and Hall/CRC, New York
- Zelterman, D. (2010), Applied Linear Models with SAS, Cambridge University Press, New York
- Zenga, M. (2014), Lezioni di Statistica Descrittiva, Giappichelli, Torino

Semester

First semester, first period

Teaching language

Italian

Sustainable Development Goals

QUALITY EDUCATION