Statistics for social sciences II: multivariate techniques

Academic Year: (2023 / 2024) **Review date: 28-04-2023**

Department assigned to the subject: Statistics Department

Coordinating teacher: GUERRERO LOZANO, VANESA

Type: Compulsory ECTS Credits : 6.0

Year : 2 Semester : 1

Branch of knowledge: Social Sciences and Law

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

Statistics for Social Sciences I or a similar introductory statistics course.

OBJECTIVES

Specific competences:

1. Understanding the basic concepts of statistical multivariate analysis and its applications in the social sciences.
2. Capacity to apply simple linear regression and interpret the results.
3. Capacity to apply multiple linear regression and interpret the results.
4. Capacity to apply binomial logistic regression and interpret the results.
5. Capacity to apply principal components analysis and interpret the results.
5. Capacity to apply cluster analysis and interpret the results.
7. Capacity to use effectively statistical software.

Transversal competences:

1. Capacity for analysis and synthesis.
2. Capacity for mathematical and statistical modeling.
3. Problem solving.
4. Critical reasoning.
5. Oral and written communication.

DESCRIPTION OF CONTENTS: PROGRAMME

Topic 1. Linear regression.
1.1. Linear regression. Introduction; simple and multiple regression; motivation; graphical data analysis; model formulation; dummy variables; parameter interpretation; examples; applications.
1.2. Fitting the model to the data; the least squares criterion; using the fitted model.
1.3. Model assumptions; inference on model parameters I: confidence intervals; inference on the response.
1.4. Inference on model parameters II: hypothesis testing; statistical significance of estimated parameters.
1.5. Assessing model fit; ANOVA.
1.6. Selection of predictor variables; multicollinearity; model diagnostics; model validation.

Topic 2. Binomial logistic regression.
2.1. Motivation; model assumptions and formulation; parameter interpretation; examples; applications.
2.2. Fitting the model to the data; using the fitted model; inference on model parameters; statistical significance of estimated parameters.
2.3. Assessing model fit; selection of predictor variables; multicollinearity.

Topic 3. Principal component analysis.
3.1. Motivation; formulation; variance explained; examples; applications.
3.2. Deciding how many components to keep; component scores; interpretation of components; graphical representations.

Topic 4. Cluster analysis.
4.1. Motivation; k-means clustering.
4.2. Hierarchical methods; similarity measures; dendrograms.

4.3. Applications and examples.

LEARNING ACTIVITIES AND METHODOLOGY

Theory (3 ECTS). Theory classes with supporting material available in the course's web page. Practical classes (3 ECTS). Problem-solving classes. Practical classes in computer rooms. Weekly individual tutoring sessions.

The teaching methodology will be eminently practical, being based on the study of diverse data sets through multivariante analysis techniques, both in the theory and practical classes, as motivation and illustration of the theory.

ASSESSMENT SYSTEM

Continuous evaluation: 70%. It will be based on two midterm exams and on exercises to be done in the computer labs.

Final exam: 30%.

Extraordinary call:

If the student followed the continuous assessment process:
- The exam in the extraordinary call will have the same percentage value as in the ordinary call, and the final grade of the subject will take into account the grade of the continuous assessment and the grade obtained in the exam. However, the student will have the right to be graded in the extraordinary call taking into account only the grade obtained in the exam when it is more favourable.

For those students who do not follow the continuous assessment process:
- The grade in the extraordinary call is exclusively obtained through an exam consisting of both theoretical questions and practical problems, with a 100% of value in the final mark.

% end-of-term-examination:	30
% of continuous assessment (assigments, laboratory, practicals…):	70

BASIC BIBLIOGRAPHY

- A. Agresti Statistical Methods for the Social Sciences, Pearson Education Limited, 2017
- D.J. Bartholomew, F. Steele, I. Moustaki, J. Galbraith Analysis of Multivariate Social Science Data, 2nd ed., Chapman & Hall/CRC, 2008
- J.F. Hair, W.C. Black, B.J. Babin, R.E. Anderson Multivariate Data Analysis: A Global Perspective, 7th ed. , Pearson Education, 2010

ADDITIONAL BIBLIOGRAPHY

- Chatterjee, S. Regression analysis by example, Wiley, 2000
- N.R. Draper Applied Regression Analysis, 3rd ed., Wiley , 1998
- Peña, D. Regresión y diseño de experimentos, Alianza, 2002
- Peña, D. Análisis de datos multivariantes, McGraw-Hill, 2002
- Pérez López, C. Técnicas de análisis multivariante de datos : aplicaciones con SPSS, Pearson Prentice Hall, 2004
- YOUNGER, M. S. A First Course in Linear Regression, Duxbury Press, 1985

BASIC ELECTRONIC RESOURCES

- J.F. Hair et al. . Multivariate Data Analysis: A Global Perspective, 7th ed.: http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9781292035116