Multivariate Analysis

Course code:	1608
Number of ECTS credits:	6
Semester:	2nd (February-June)
Prerequisites:	60 ECTS of basic subjects and 60 ECTS of compulsory subjects.
Recommended components:	Inferencia Estadística (1598)
Language of instruction:	Spanish (students are allowed to ask questions and write homeworks and exams in English)

Course description

This is an introductory course on the methods of multivariate data analysis. Firstly, we study the theory of linear models used in regression for modelling the relationship between random variables and predictor variables. Secondly, we study how to summarize the information contained in a data set with many random variables by using a Principal Component Analysis (PCA). Finally, we study how to classify data into prior populations by using the measures in different random variables and a Discriminant Analysis. All these techniques are used to analyze data sets by using the free statistical software R.

Learning outcomes and competences

After completion of this course you will:

1. know the basic multivariate analysis tools.

2. know the basic properties of Linear Models.

3. know how to perform a Linear Model to estimate one or more random variables.

4. know how to compute the principal components and their basic properties.

5. know how to perform a principal components analysis (PCA) of a data set.

6. know how to compute the linear and quadratic discriminant functions and their basic properties.

7. know how to perform a discriminant analysis (DA) over a data set to classify a data into one of the given populations.

Course contents

1. Multiple Regression Analysis.

   Introduction to lineal regression models. Estimation and tests of parameters. Coefficient of determination. Confidence intervals and prediction intervals. Model validation and diagnostics of residuals. Extensions of multiple regression and related applications using real data sets.

2. Principal Component Analysis.

   Definitions and theoretical properties of principal components. Sampling computation and properties. Principal Component Analysis of data sets. Methods to determine the number of significant components.

3. Discriminant Analysis.

   Linear Discrimination methods. Fisher's lineal discriminant function. Quadratic Discriminant methods. Maximum likelihood methods. Sampling properties. Discriminant Analysis of data sets. Probabilities of misclassification.

References

Main texts

1. Draper, N.R., Smith, H. (1998). Applied Regression Analysis, 3rd. John Wiley.

2. Mardia, K.V., Kent, J.T., Bibby, J.M. (1998). Multivariate analysis. Academic Press.

3. Rencher, A.C. (2002). Methods of multivariate analysis. Wiley.

Supplementary references

1. Peña, D. (1992). Estadística Modelos y Métodos. Modelos Lineales y Series Temporales. Alianza Universidad Textos.

2. Peña, D. (2002). Análisis de datos multivariantes. McGraw Hill.

Faculty of Mathematics - University of Murcia
Campus Universitario de Espinardo - Murcia - Spain
T. +34 868 88 4181 - F. +34 868 88 4182 - mailto:decamate@um.es - www.um.es/web/matematicas/