# uc3m Universidad Carlos III de Madrid

### Statistics for data analysis

Academic Year: (2022 / 2023) Review date: 17/05/2022 12:40:08

Department assigned to the subject: Statistics Department Coordinating teacher: AUSIN OLIVERA, MARIA CONCEPCION

Type: Compulsory ECTS Credits: 3.0

Year: 1 Semester: 1

#### **OBJECTIVES**

#### Basic skills

To achieve and understand the knowledge that provide foundations or changes of being original in the development and/or application of ideas, often in a context of research

To learn habilities that allow oneself to continue studying in a self-sufficient way.

#### General skills

To apply the theoretical foundations of data collection techniques, storage, treatment and presentation of the information, specially for large data sets, as a basis for the development and adaptation of these techniques to specific problems.

### Specific skills

To use the advanced knowledge in Linear Algebra for its application in the analysis of large data sets.

Apply basic concepts of programming and data bases in the learning process of technologies and advanced methods for the treatment of large data sets.

To employ the basic results of inference and regression as foundations for the advanced methods of prediction and classification

## Learning achievements:

- Using the tools of Linear Algebra (matrices, factorization) and Calculus (multidimensional integration) in the design and analysis of methods for treatment of data
- Habilities for the application of the basic statistical techniques (different data displays, probabilities, distributions, sampling, estimation) to data analysis.
- Knowledge and use of basic programming concepts, structured programming, design of algorithms and relational data bases for the development of applications for the treatment of data.

## **DESCRIPTION OF CONTENTS: PROGRAMME**

- 1 Descriptive statistics
- 1.1. Introduction to Statistics
- 1.2. Description of one variable
- 1.3. Relation between two variables
- 1.4. Variable transformations

### 2 Probability theory

- 2.1. Introduction to Probability
- 2.2. Discrete random variables
- 2.3. Continuous random variables
- 2.4. Transformations of random variables
- 3. Statistical inference
- 3.1. Estimation methods
- 3.2. Inference under normality assumptions
- 3.3. Inference for large samples
- 3.4. Chi-square tests

# 3.5. Correlation and dependence

### LEARNING ACTIVITIES AND METHODOLOGY

Learning activities:

Theoretical lessons

Exercises

Practical lessons in class using the computer

Individual project

### Teaching methodologies:

Presentations of the professor in class with computing and visual media, where the professor develops the mail concepts of the subject and provides bibliography supplementing the knowledge of students.

Critical reading of texts recommended by the professor: Press articles, reports, manuals and/or academic papers, either for their posterior discussion in class, or for widening and consolidating the subject matter.

Solving practical case studies, problems, etc., established by the professor either individually or in groups

### ASSESSMENT SYSTEM

% end-of-term-examination/test:

% of continuous assessment (assigments, laboratory, practicals...):

- Quizzes (10%) Students should complete a few self-assesment tests during the classes
- Intermediate test (20%).
- Final project (70%) based on a descriptive and statistical analyusis of a real data set collected by the student from a real survey.

30

### **BASIC BIBLIOGRAPHY**

- Dalgaard, Peter Introductory statistics with R, Springer, 2008
- Michael W. Trosset An Introduction to Statistical Inference and Its Applications with R, Chapman and Hall/CRC, 2009
- Ugarte, María Dolores Probability and statistics with R, CRC Press, 2008