uc3m Universidad Carlos III de Madrid

Statistical Data Analysis

Academic Year: (2022 / 2023) Review date: 21-06-2022

Department assigned to the subject: Statistics Department Coordinating teacher: MARIN DIAZARAQUE, JUAN MIGUEL

Type: Basic Core ECTS Credits: 6.0

Year: 2 Semester: 1

Branch of knowledge: Social Sciences and Law

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

None

OBJECTIVES

- 1. Analyze univariate and bivariate data
- 2. Solve probability problems
- 3. Use random variables
- 4. Show and understand basic concepts in Estimation techniques
- 5. Be able to solve problems in Estimation
- 6. Be able to solve problems using the statistical software R.
- 1. Capacity for analysis and synthesis.
- 2. Knowledge of statistical software.
- 3. Resolution of problems.
- 4. Team working.
- 5. Critical reasoning.
- 6. Oral and written communication skills.

DESCRIPTION OF CONTENTS: PROGRAMME

- 1. Introduction.
- 1.1. Concepts and use of Statistics.
- 1.2. Statistical terms: populations, subpopulations, individuals and samples.
- 1.3. Types of variables.
- Analysis of univariate data with R.
- 2.1. Representations and graphics of a qualitative variable: bar plots.
- 2.2. Representations and graphics of a quantitative variable: histograms, densities and box-plots.
- 2.3. Graphics for related observations in time and space: line graphics and map visualization.
- 2.4. Numerical summaries.
- 3. Analysis of bivariate data with R.
- 3.1. Association among quantitative variables: scatter plots and correlograms.
- 3.2. Association among qualitative variables: mosaic plots.
- 3.3 Association among qualitative and quantitative variables: box-plots.
- 3.4 Associations among statistical units and variables: heat maps.
- 3.4 Numerical summaries of associations: covariance and correlation.
- 4. Probability and probabilistic models.
- 4.1. Random experiments, sample space, elemental and composite events.
- 4.2. Properties of Probability. Conditional Probability and its properties.
- 4.3. Random variables and their characteristics.
- 4.4. Discrete probability models: Bernoulli variables and related distributions.
- 4.5. Continuous probability models: the normal distribution and related distributions.
- 4.6. Introduction to the bivariate normal distribution.
- Introduction to Statistical Inference.
- 5.1. Parameter point estimation.
- 5.2. Goodness-of-fit to a probability distribution. Graphical methods.
- 5.3. The sample mean distribution.

- 5.4. Confidence interval for the mean.
- 5.5 Hypothesis testing on a mean: use and interpretation of a p-value.

LEARNING ACTIVITIES AND METHODOLOGY

14 Theoretical support materials available on the Web, and 14 sessions based on problem-solving sessions and practical computing tasks.

No group tutorials except during the last week before the final exam.

ASSESSMENT SYSTEM

60% of the final grade will be achieved by a final examination for assessing the acquired knowledge. A minimum of 4 points (out of 10) is required in the final exam. The remaining 40% is obtained by two midterm exams (15%+20%) and the compulsory tasks assigned in the computational labs (5%). Theoretical questions as well as queries on computational laboratories can be asked in the exams.

% end-of-term-examination: 60 % of continuous assessment (assigments, laboratory, practicals...): 40

BASIC BIBLIOGRAPHY

- Newbold, P. Estadística para los Negocios y la Economía., Prentice-Hall, 1997
- Newbold, P., Carlson, W., & Thorne, B. Statistics for business and economics. , Pearson., 2012

BASIC ELECTRONIC RESOURCES

- Charte Ojeda, F. . Análisis exploratorio y visualización de datos con R. : http://www.fcharte.com/libros/ExploraVisualizaConR-Fcharte.pdf