

**COURSE DATA****Data Subject**

Code	33860
Name	Introduction to the Statistical Analysis and Processing of Data
Cycle	Grade
ECTS Credits	6.0
Academic year	2022 - 2023

Study (s)

Degree	Center	Acad. year	Period
1007 - Degree in Information and Documentation	Faculty of Geography and History	1	Second term

Subject-matter

Degree	Subject-matter	Character
1007 - Degree in Information and Documentation	13 - Statistics	Basic Training

Coordination

Name	Department
FUSTER ORTI, MARIA ANGELES	130 - Statistics and Operational Research

SUMMARY

Statistical data analysis is a basic tool for its practical application on different areas of education and research in Information and Documentation, such as the assessment and management of the resources, processes and services of information units; the development of studies on the community of users, with particular attention to gender studies; or the quantitative analysis of the scientific and documentary production in general. Moreover, statistics is a basic analytical tool for problem-solving and decision-making in any organisation or system. The aim is to provide the tools and basic concepts of statistics that are needed to formulate hypotheses, recognise simple probabilistic models, analyse data sets and make decisions based on statistical techniques.



PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

Other requirements

No hay

OUTCOMES

1007 - Degree in Information and Documentation

- Capacity to write analytical reports and summaries with regard to management and organisation of information.
- Demonstrate organisational and planning skills.
- Have computer skills related to the field of study.
- Have skills for information management.
- Have problem-solving skills.
- Have decision-making capacity.
- Be able to apply critical reasoning to the analysis and assessment of alternatives.
- Be able to undertake improvements and propose innovations.

LEARNING OUTCOMES

Be familiar with the variability that is intrinsic to the data obtained in any type of study and with the problems that this may cause for drawing conclusions from research.

Model simple random experiments by identifying the appropriate probability model.

Handle techniques related with statistical inference, such as estimation and testing of hypotheses, which will be used in the statistical analysis of experiments.

Analyse the applicability conditions and limitations of the most common statistical methods. Know how to use statistical software to work with databases. Interpret graphs and tables obtained with statistical software.

DESCRIPTION OF CONTENTS



1. Introduction to Statistics.

Concept and classification. Basic concepts. Types of experiments. Population and sample. Sampling. Sampling error. Representativeness of the sample. Sampling methods.

2. Description of a sample.

Descriptive statistics. Organisation of data. Qualitative and quantitative variables. Measures of central tendency. Measures of position. Measures of dispersion. Measures of shape. Box plots. Basic ideas for data presentation.

3. Relationship between two variables.

Two-dimensional frequency distribution. Concept and types of correlation. Linear correlation coefficient. Linear regression. Regression line. Quality of fit. Multiple regression.

4. Description of a population: Probability distributions.

Introduction. Interpretation of density. Examples of continuous distributions. The normal distribution. The standard normal distribution. Properties. Calculation of areas. Sampling variability.

5. Inferential analysis in a population. 5a) Estimation. 5b) Hypothesis testing.

5a) Statistical inference. Point and interval estimation. Standard error of the mean. Confidence interval for the mean of a population. Student's t-distribution. Interpretation of the confidence interval. Sample size selection.

5b) Hypothesis testing. Introduction. Bilateral or unilateral contrast. Choice of alternative hypothesis. The test statistic. P-value. Level of significance. Decision and conclusion of a test. Relationship of the bilateral hypothesis test to the confidence interval. Errors and power of the test. Contrast on a mean. Normality tests.

6. Inference with two populations.

Introduction. Paired and independent samples. Comparison of two populations. Confidence interval and hypothesis testing of the difference of means.

7. Analysis of categorical data.

Categorical variables. Types of statistical studies. Confidence interval for a proportion. Chi-square test of goodness of fit. Hypothesis testing on a proportion. Contingency tables. Independence and homogeneity tests.

**WORKLOAD**

ACTIVITY	Hours	% To be attended
Theory classes	30,00	100
Computer classroom practice	30,00	100
Development of individual work	15,00	0
Study and independent work	27,00	0
Preparation of evaluation activities	3,00	0
Preparing lectures	15,00	0
Preparation of practical classes and problem	30,00	0
TOTAL	150,00	

TEACHING METHODOLOGY

All the documents that will be used, both in the theoretical and practical sessions, will be available in the Virtual Classroom environment.

PRESENTIAL CLASSES:

In the theory classes real problems will arise whose resolution requires the methodology corresponding to each topic, previously explained.

The practical sessions, in the computer classroom and synchronized with the theory, will allow the student to apply these procedures to solve problems. Each student will have a dossier in which the content of each practice will be described, and will include the problems that will be solved in it.

PREPARATION OF PRACTICAL WORK:

The student must attend each practical session having previously read this dossier and having reviewed the theoretical contents related to it

EVALUATION**Written tests:**

Taking into account the fundamentally practical approach of the subject it is impossible to conceive, at the time of evaluation, a separation between theory and practice, so the evaluation will consist of a theoretical-practical examination whose resolution will require the interpretation of different results presented in the standard format of the statistical software used. It will be essential to obtain a grade on the exam equal to or greater than 5 out of 10 to average with the rest of the grades.

**Individual and/or group work:**

They will consist in the resolution, in group, of the problems proposed in some practical classes and/or the realization of some individual works analyzing, from a statistical point of view, some data set.

The composition of the final grade will be taken, in summary, to the following table:

- EXAMINATION: 50%
- INDIVIDUAL AND/OR GROUP WORK: 50%
- TOTAL: 100%

This evaluation is based on the premise that teaching at the Universitat de València is, by definition, a face-to-face teaching. In this sense, students should bear in mind that attendance, both theoretical and practical, is essential for proper monitoring of the contents of the subject. The student must also bear in mind the possibility of a part-time enrollment when it is not possible to attend all the subjects that make up a complete course (60 credits). However, the possibility will be established, in cases that are adequately justified and for those students who request it, the possibility of being evaluated without having to attend all or part of the classes. In these cases the student should proceed as follows:

- At the beginning of the course, the teacher/s responsible for the subject must be informed about the incident for which it is impossible for them to attend the class, which must be adequately justified in a documentary way.
- The responsible teacher, in view of this information will decide the possibility of evaluation without total or partial assistance to the classes of the subject.

Students who are in this situation must submit for evaluation all work required by the lecturer (not necessarily the same to those required for the course) and may also be called to defend them orally to the lecturer, and conduct a knowledge test. The weight of the final grade work will be 50% and the test the remaining 50% knowledge. Students who do not attend Theoretical activities and / or practices, and individual and collective practical work, will read a series of supplementary texts.

REFERENCES

Basic

- Egghe, L. y Rousseau, R. (1990). Introduction to Informetrics: quantitative methods in library, documentation and information science. Elsevier
- Glenberg, A. M. y Andrzejewski, M. E. (2008). Learning from data. An introduction to statistical reasoning. Lawrence Erlbaum Associates, 3ª Edición
- Marín, J. (2008). Estadística Aplicada a las Ciencias de la Documentación. Diego Marín Librero-Editor, 3ª Edición
- Peña, D. y Romo, J. (1997) Introducción a la Estadística para las Ciencias Sociales. McGraw-Hill.



Additional

- Milton, J. S. (2001). Estadística para Biología y Ciencias de la Salud. Madrid: Ed. Interamericana - McGraw-Hill. 3ª Edición
- Samuels, M.L., Witmer, J.A. y Schaffner (2012) Fundamentos de Estadística para las ciencias de la vida. Pearson Educación SA. 4ª Edición