

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

<b>Code</b>	34177
<b>Name</b>	Statistical modelling
<b>Cycle</b>	Grade
<b>ECTS Credits</b>	6.0
<b>Academic year</b>	2023 - 2024

**Study (s)**

<b>Degree</b>	<b>Center</b>	<b>Acad. year</b>	<b>Period</b>
1107 - Degree in Mathematics	Faculty of Mathematics	4	Second term

**Subject-matter**

<b>Degree</b>	<b>Subject-matter</b>	<b>Character</b>
1107 - Degree in Mathematics	17 - Models of statistics and operations research	Optional

**Coordination**

<b>Name</b>	<b>Department</b>
ARMERO CERVERA, MARIA CARMEN	130 - Statistics and Operational Research

**SUMMARY**

This course aims to give an overview of how to perform a complete statistical analysis, specifically how to perform previously the most essential part correctly modeling the situation to be studied and the selection of the best model that fits the data analyzed.

To do so, the basics of modeling and the most important techniques available to perform the appropriate statistical analysis are introduced. The course describes with more detail those concepts of statistical inference previously introduced on the first year course on "Basic Statistics". It will also make use of the basic concepts and techniques of estimation (point and confidence intervals) and contrast hypothesis presented in the course of "Mathematical Statistics" (third year of the degree).



## PREVIOUS KNOWLEDGE

### Relationship to other subjects of the same degree

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

### Other requirements

Have studied the subjects of Basic Statistics (1st year of the degree), Probability (2nd year) and Statistics Mathematics (3rd year).

## OUTCOMES

### 1107 - Degree in Mathematics

- Learn autonomously.
- Adapting to new situations.
- Apply the knowledge in the professional world.
- Argue logically in decision-making.
- Expressing mathematically in a rigorous and clear manner.
- Reason logically and identify errors in the procedures.
- Capacity of abstraction and modeling.
- Participate in the implementation of software and learn mathematical software.
- Knowing the time and the historical context in which occurred the great contributions of women and men in the development of mathematics.
- Visualize and interpret the solutions obtained.

## LEARNING OUTCOMES

- Formulate and solve problems of various types using statistical models.
- Learn to assess the conditions of application of statistical models.
- Use statistical analysis software.
- Correctly interpret the results provided by statistical packages.

## DESCRIPTION OF CONTENTS

### 1. Linear models.

1.1 Simple linear regression model

1.2 Linear models (LM). Least squares estimation. The distribution of the estimators of the regression coefficients. Hypothesis tests. Influence matrix, residuals and fitted values. Gauss Markov theorem.

1.3 The geometry of linear models.

1.4 LM and R



1.5 Model fitting and model checking. Model selection. Confidence and prediction intervals. Collinearity, confusión and causality. Factors and interaction

## 2. Generalized linear models.

- 2.1 Exponential family of distributions
- 2.2 Generalized lineal model (GLM). Model fitting. Asymptotic distribution of the estimators of the regression coefficients. Model comparison. Canonical link functions. Residuals. Quasi-likelihood
- 2.3 The geometry of linear modelling
- 2.4 Logistic regression and Poisson regression.
- 2.5 Tweedie distribution and negative binomial distribution. Cox model for survival data.
- 2.6 GLMs and R.

## 3. Linear mixed models

- 3.1 Linear mixed models (LMM) for balanced data.
- 3.2 Models with one and two factors.
- 3.3 Maximum likelihood estimation REML. Degrees of freedom. EM Algorithm.
- 3.4 LMM and R.

## WORKLOAD

ACTIVITY	Hours	% To be attended
Theory classes	37,50	100
Computer classroom practice	15,00	100
Other activities	7,50	100
Development of group work	5,00	0
Study and independent work	30,00	0
Preparation of evaluation activities	10,00	0
Preparing lectures	10,00	0
Preparation of practical classes and problem	10,00	0
Resolution of online questionnaires	20,00	0
<b>TOTAL</b>	<b>145,00</b>	

## TEACHING METHODOLOGY

**Theoretical activities.** Lectures on the subject by the teacher with the participation of the students in the solution of specific questions.



**Practicals and seminars.** Learning through the resolution of exercises, problems and tasks, as well as the preparation of reports. These activities will be carried out individually or collectively in small groups in the computer classroom (practical sessions) or in the theory class (seminars).

## EVALUATION

The evaluation of the subject will be based on three elements

- Practical work. Individual assignments collected in the practical classes. Percentage of the total mark: 15%.
- Seminars. A group task on a subject related to the subject. Percentage of the total mark: 15%.
- Examination. Percentage of the total mark: 70%.

In order to pass the course it will be necessary, but not sufficient, that the student obtains a minimum mark of 4 in the exam, scored out of 10, and a minimum mark in the joint evaluation of the practicals and seminars, jointly scored out of 3.

Those students who do not pass the minimum grade in the practical and seminars will have to sit an exam on both subjects in the second general exam.

Students who have a minimum mark in practicals and seminars, but not in the exam, will be able to keep their mark in the second session.

## REFERENCES

### Basic

- Referència b1: S. N. Wood. (2017). Generalized Additive Models. An Introduction with R (Second Edition). CRC Press. Taylor & Francis Group.

### Additional

- Faraway, J. J. (2016). Extending the Linear Model with R . Taylor & Francis, 2016.
- G. James, D. Witten, T. Hastie, R. Tibshirani (2021). An Introduction to Statistical Learning with Applications in R (Second Edition). Springer.
- T. Hastie, R. Tibshirani and J. Friedman (2021). The Elements of Statistical Learning. Data Mining, Inference, and Prediction (Second Edition). Springer.