Course guide
33104 - TAEDDPE - Techniques of Statistic Analysis of Data and Design and Planning of Experiments

Last modified: 25/05/2024

Unit in charge: Manresa School of Engineering
Teaching unit: 749 - MAT - Department of Mathematics.

Degree: MASTER'S DEGREE IN NATURAL RESOURCE ENGINEERING (Syllabus 2015). (Compulsory subject).

Academic year: 2024 **ECTS Credits:** 5.0 **Languages:** Spanish

LECTURER

Coordinating lecturer: Rossell Garriga, Josep Maria

Others:

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
1. Analyse field and laboratory data and design experiments following computer methods.
2. Identify analytical techniques for the characterisation of inorganic natural resources and waste in different states, use these techniques and interpret the results.
3. Use scientific and technical information to respond efficiently to any demand for the preparation of an analytical method for characterising a material of natural or anthropogenic origin.

TEACHING METHODOLOGY

The course is based on directed learning activities, which include studying the main statistical methods and solving applied problems using a computer. In the face-to-face format, practical lectures serve to introduce new concepts, guide students on the use of computational resources and solve doubts. In the non-face-to-face mode, an activity programme that allows the course's objectives to be met independently is proposed. In this second case, in order to facilitate the follow-up of the course, the recordings of the classes that are given to the students in face-to-face mode will be provided. Using advanced computational resources is a fundamental part of applying statistical methods. In this course, students use the MINITAB computing environment, which provides a tool with outstanding features for professional practice, research and learning in advanced statistical methods.

LEARNING OBJECTIVES OF THE SUBJECT

The principal aim of the course is to familiarise students with the theory and applications of the main statistical methods. Fundamental concepts in statistics and basic statistical tools are introduced and/or revised, allowing us to take samples, process data, analyse the results, make statistical inferences, formulate and fit models, design experiments and work with data mining. To illustrate the multidisciplinary nature of the concepts and methods studied, environmental and industrial data are used.

STUDY LOAD

Type	Hours	Percentage
Hours large group	30,0	66.67
Hours medium group	15,0	33.33

Total learning time: 45 h

CONTENTS

Unit 1: Basic concepts of statistics. Introduction to MINITAB.

Description:
Types of data. Descriptive and inferential statistics. Random sampling. Types of sampling. Basic concepts in probability. Basic functioning of Minitab. Calculation of statistical descriptors with Minitab. Graphic representation with MINITAB.

Specific objectives:
Revision of the main concepts in statistics. Basic use of MINITAB.

Related activities:
Students must do and hand in the E1 set of exercises.

Full-or-part-time: 25h
Theory classes: 5h
Laboratory classes: 5h
Self study : 15h

Unit 2. Univariate statistics.

Description:
Random variables. Discrete distributions: uniform, binomial, geometric, Poisson. Continuous distributions: exponential, normal, t-Student, chi-square, F-Fisher. Calculation of special distributions with MINITAB. Simulation of random variables with MINITAB. Confidence intervals. Calculation of confidence intervals with MINITAB. Hypothesis testing. Hypothesis testing with MINITAB.

Specific objectives:
Revision of the main probability distributions and their application with Minitab. Revision of the main tools in inferential statistics for univariate problems and their implementation with MINITAB.

Related activities:
Students must do and hand in the E2 set of exercises.

Full-or-part-time: 20h
Theory classes: 5h
Laboratory classes: 2h
Self study : 13h

Unit 3. Fitting of models with a single predictor.

Description:
Models with a single predictor. Linear regression. Pearson coefficient and coefficient of determination. Fitting of polynomial models. Fitting of non-linear models. Fitting of models with MINITAB.

Specific objectives:
Revision of the main concepts in the fitting of models with a single predictor and the implementation of the models with MINITAB.

Related activities:
Students must do and hand in the E3 set of exercises.

Full-or-part-time: 20h
Theory classes: 5h
Laboratory classes: 2h
Self study : 13h

Unit 4: Multivariate statistics.

Description:
Introduction to multivariate statistics. Multiple linear regression. Multiple linear regression with MINITAB.

Specific objectives:
Presentation of the basic concepts in multivariate statistics. Discussion of the fitting of models with a set of predictors and the implementation of the models with MINITAB.

Related activities:
Students must do and hand in the E4 set of exercises.

Full-or-part-time: 20h
Theory classes: 5h
Laboratory classes: 2h
Self study : 13h

Unit 5: Principal component analysis.

Description:
Linear dimensionality reduction technique with applications in exploratory data analysis, visualization and data preprocessing. The data is linearly transformed onto a new coordinate system such that the directions (principal components) capturing the largest variation in the data can be easily identified.

Specific objectives:
Presentation of the basic concepts and techniques of principal component analysis. Implementation with MINITAB.

Related activities:
Students must do and hand in the E5 set of exercises.

Full-or-part-time: 20h
Theory classes: 5h
Laboratory classes: 2h
Self study : 13h

Unit 6: Experimental design and analysis.

Description:
Observational studies and experiments. Single-factor experiments. Multifactorial designs. Experimental design with MINITAB.

Specific objectives:
Presentation of the main elements of experimental design and their implementation with MINITAB.

Related activities:
Students must do and hand in the E6 set of exercises.

Full-or-part-time: 20h
Theory classes: 5h
Laboratory classes: 2h
Self study : 13h

GRADING SYSTEM

Assessment (Face-to-Face). The final mark takes into account the following:
1. The individual assignments handed in throughout the course. Six assignments are planned, that is, an assignment every two weeks. Each assignment submitted as a pdf file is corrected and returned as soon as possible so that students can solve the questions they have not answered correctly. Students have the option of improving their assignments and submitting them for a second time. Each submission (a maximum of two for each assignment) is given an individual numerical mark, so the mark for an assignment is the average of the marks for the submissions. The final mark for the assignments is the arithmetic average of the marks for the six assignments handed in and has a weight of 70% of the final mark for the module.
2. Students' attendance, which is 30% of the final mark. An attendance sheet must be signed by students at the end of each face-to-face class.
3. Students' participation and degree of involvement during the course. This assessment, which is obviously more subjective, is used to round off the final mark.

Assessment (Non-Face-to Face). Exercises to be done and submitted are uploaded regularly (approximately every two weeks) to the ATENEA digital platform. Exercises must be done individually using the MINITAB program. Summaries of the topics covered will be offered, as well as links to web pages of interest for those who wish to delve deeper into a topic. The final mark takes into account the following:

1. The submission and assessment of individual assignments carried out by students during the course. Six assignments are planned, that is, an assignment every two weeks. Each assignment submitted as a PDF file in ATENEA is corrected and returned as soon as possible so that students can solve the questions they have not answered correctly. Students have the option of improving their assignments and submitting them for a second time. Each submission (a maximum of two for each assignment) is given an individual numerical mark, so the mark for an assignment is the average of the marks for the submissions. The final mark is the arithmetic average of the marks for the six assignments handed in and has a weight of 70% of the final mark for the module.
2. At the end of the course, students take an individual exam on site on the six assignments, in which they must demonstrate that they have done the assignments and assimilated the course content. Thus, students must present and defend one of the six assignments, which will be chosen at random just before the defence. The mark for this exam accounts for 30% of the final mark for the module.

EXAMINATION RULES.

Students must do the E1-E6 sets of exercises individually and hand them in before the deadlines in the ATENEA digital platform so that they can be corrected and given a mark. If a set of exercises is not handed in, students are awarded a mark of 0. There will be a maximum of two submissions for each of the E1-E6 sets of exercises. The first submission will be given a mark and comments will be made so that the sections that are wrong can be corrected and improved. Students are given the option of handing in a corrected version of the assignment in a second, voluntary submission. This second submission is given a mark. The final mark for the assignment is the arithmetic average of the two submissions, if applicable, or the mark for the first submission if the student does not make a second submission.

BIBLIOGRAPHY

Basic:
- Scheaffer, Richard L.; Mendenhall, William; Ott, Lyman. Elementos de muestreo. 6ª ed. Madrid: Thomson, 2007. ISBN 8497324935.
- Walpole, Ronald E.; Myers, Raymond H.; Myers, Sharon L.; Ye, Keying. Probabilidad y estadística para ingeniería y ciencias [on line]. 9ª ed. México: Pearson, 2012 [Consultation: 28/07/2022]. Available on: https://www-ingebook-com.recursos.biblioteca.upc.edu/ib/NPcd/IB_BooksVis?cod_primaria=1000187&codigo_libro=6766. ISBN 9786073214179.
- Montgomery, Douglas C. Design and Analysis of Experiments. 8th ed., International Student Version. New York: John Wiley & Sons, 2013. ISBN 9781118097939.
- Introducción a Minitab Statistical Software [on line]. Minitab, 2022 [Consultation: 27/07/2022]. Available on: https://www.minitab.com/content/dam/www/en/uploadedfiles/documents/getting-started/MinitabGettingStarted_ES.pdf.
- Grima Cintas, Pere; Tort-Martorell Llabrés, Xavier; Marco Almagro, Lluís. Estadística práctica con Minitab. Madrid: Pearson Educación, cop. 2004. ISBN 8420543551.
- Witten, I. H; Hall, Mark A; Frank, Eibe. Data mining: practical machine learning tools and techniques [on line]. 3rd ed. Burlington: Morgan Kaufman, cop. 2011 [Consultation: 28/07/2022]. Available on: https://www-sciencedirect-com.recursos.biblioteca.upc.edu/book/9780123748560/data-mining-practical-machine-learning-tools-and-techniques. ISBN 9780123748560.

Complementary:

- Jolliffe, I. T. Principal component analysis [on line]. Second Edition. New York: Springer, 2002 [Consultation: 04/06/2024]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/b98835. ISBN 0387962697.

RESOURCES

Computer material:

- Aules d'informàtica. Resource

Other resources:

Videos of the recordings of the face-to-face classes.