Descriptive Statistics (250227)

General Information

School	ETSECCPB
Departments	Departament d'Enginyeria Civil i Ambiental (DECA)
Credits	7.5 ECTS
Programs	GRAU EN ENGINYERIA D'OBRES PÚBLIQUES (pla 2010)
Course	2022/23

Main teaching language at each group

- Group 10Q1 Spanish (Q1)

Faculty

Responsible Faculty: María Isabel Ortego Martínez
Faculty: Agustin Medina Sierra, María Isabel Ortego Martínez

Objectives of Education

Students will acquire the skills to represent and process data, including basic knowledge of databases and computer software with applications in engineering. They will also learn statistical concepts.

Upon completion of the course, students will have acquired the ability to: 1. Carry out a data analysis of a construction engineering problem using a computer tool that employs the studied techniques. 2. Carry out a multiple linear regression analysis using computer software. 3. Carry out data simulations and transformations of random variables, as well as studies of distributions.

Data analysis; Regression models, parameter estimation; Probability and uncertainty; Basics of point estimation and interval estimation; Hypothesis testing

Competencies

Especific

Ability to solve the types of mathematical problems that may arise in engineering. Ability to apply knowledge of: linear algebra; geometry; differential geometry; differential and integral calculus; differential equations and partial derivatives; numerical methods; numerical algorithms; statistics and optimisation.

Transversal

EFFICIENT ORAL AND WRITTEN COMMUNICATION - Level 2. Using strategies for preparing and giving oral presentations. Writing texts and documents whose content is coherent, well structured and free of spelling and grammatical errors.
TEAMWORK - Level 2. Contributing to the consolidation of a team by planning targets and working efficiently to favor communication, task assignment and cohesion.
EFFECTIVE USE OF INFORMATI0N RESOURCES - Level 3. Planning and using the information necessary for an academic assignment (a final thesis, for example) based on a critical appraisal of the information resources used.
SELF-DIRECTED LEARNING - Level 3. Applying the knowledge gained in completing a task according to its relevance and importance. Deciding how to carry out a task, the amount of time to be devoted to it and the most suitable information sources.
THIRD LANGUAGE. Learning a third language, preferably English, to a degree of oral and written fluency that fits in with the future needs of the graduates of each course.

Total hours of student work

		Hours	Percentage
Supervised Learning	Large group	45.0 h	54.55 %
	Medium group	15.0 h	18.18 %
	Laboratory classes	15.0 h	18.18 %
	Guided Activities	7.5 h	9.09 %
Self Study		105.0 h	

Contents

Exploratory data analysis

Simulation of data. Qualitative variables, discrete, continuous. Graphical representation. Location measures. Calculation of statistics. Transformacón variables.

Simulation data. Qualitative variables, discrete, continuous. Graphical representation. Measures of location and dispersion. Transformation of data.
Simulation data. Qualitative variables, discrete, continuous. Graphical representation. Measures of location and dispersion. Transformation of data.
Multivariate analysis and data processing. Tables of double entry. Least-square adjustment. Correlation. Trends.
Multivariate analysis and data processing. Tables of double entry. Least-square adjustment. Correlation. Trends.
Multivariate analysis and data processing. Tables of double entry. Least-square adjustment. Correlation. Trends.
Study and graphical representation. Analysis of the trend. Components. Smoothing and seasonal index computations.
Study and graphical representation. Analysis of the trend. Components. Smoothing and seasonal index calculations.
Study and graphical representation. Analysis of the trend. Components. Smoothing and seasonal index calculations.

Elementary Probability. Probability models.

Counting techniques. Definition of probability. Conditional probability. Bayes Formula.
Counting techniques. Discrete Probability. Conditional probability. Bayes Formula.
Counting techniques. Discrete Probability. Conditional probability. Bayes Formula.
Definition. Distributions of Probability. Discrete and continuous variables . Probability density function. Moments.
Discrete distributions: Bernoulli, binomial, geometric, Poisson, ...
Representation of probability functions.

Definition. Distributions of Probability. Discrete and continuous variables . Probability density function. Moments.
Discrete distributions: Bernoulli, binomial, geometric, Poisson, ...
Representation of probability functions.

Definition. Probability distributions. Discretes and continuous variables . Probability density function. Moments.
Discrete Distributions: Bernoulli, binomial, geometric, Poisson, ... Representation of probability functions.

Continuous models: uniform, normal, exponential, ... Central limit theorem. Processing methods.
Representation of densities. Point processes in time. Return periods.

Continuous models: uniform, normal, exponential, ... Central limit theorem. Processing methods.

Representation of densities. Point processes in time. Return periods.
Continuous models: uniform distribution, normal, exponential, ... Central limit theorem. Transforming variables. Representation of densities. Puntual processes in time. Return periods.

Statistical Inference

Samples. Estimation by the method of moments. Maximum likelihood estimation. Properties of estimators. Examples of estimation.

Samples. Estimation by the method of moments. Maximum likelihood estimation. Properties of estimators.

Samples. Estimation by method of moments. Concept of likelihood. Most likelihood estimate. Properties of estimators.
Concept of testing hypotheses. Applications in normal sampling. Decision rules. Errors type I and II. Power. Test on average and variance of normal populations. T distributions, Chi2, F. Estimation of the p-value.

Concept of contrasting hypotheses
Applications in normal sampling. Decision rules. Errors type I and II. Power. Contrasts of mean and variance of normal populations. T distributions, Chi2, F.
Estimate p-value.
Concept of contrasting hypotheses
Applications in normal sampling. Decision rules. Errors type I and II. Power. Contrasts of mean and variance of normal populations. T distributions, Chi2, F.
Estimation of the p-value.

Linear regression model

Regression model and its extensions. Adjustment by least squares. Usual assumptions of the model.
Regression model and its extensions. Adjustment by least squares. Common model assumptions
Regression model and its extensions. Adjustment by least squares. Common model assumptions
First example. F tests on the regression. T Test on coefficients. Normality of residuals.
First example. F tests on the regression. T Test on coefficients. Normality of residuals.
First example. F tests on the regression. T Test on coefficients. Normality of residuals.
Multiple regression. ANOVA. Factors.

Multiple regression. ANOVA. Factors
Multiple regression. ANOVA. Factors.
Review and synthesis of the contents of theory, problems and laboratory of issues 3 and 4. Evaluation of them by conducting written tests.
Review and synthesis of the contents of theory, problems and laboratory issues 1, 2, 3 and 4. Final evaluation of them by conducting written tests for students that fail or to improve the final qualification.

Activities

Course project

Dedication

7h 30m

Teaching Methodology

The subject consists of 5 hours per week of classroom lessons face to face in the classroom or online.

These lecture hours include theory, problems and laboratory sessions and they are not strictly distinguished. Full size and medium size teaching groups are combined, following the teaching program.

Statistics has an eminently applied and computational component. Therefore, practical classes, that are taught in the same classroom, should be followed using a laptop. Specific software is used (R + RStudio, among others).

Students should use the support materials that will be available in the virtual campus ATENEA: updated information about subject organization, contents, scheduling of activities, learning assessment and bibliography.

The main language of instruction will be Catalan/Spanish/English, depending on the group. The other languages will also be used. Teaching materials may be written in any of the three languages.

Although most of the sessions will be given in the language indicated, sessions supported by other occasional guest experts may be held in other languages.

Grading Rules

(*) The evaluation calendar and grading rules will be approved before the start of the course.

The mark of the course is obtained as a weighted average of the marks of the continuous assessment activities.

This continuous assessment consists of activities of different types, both individual and group, of an additional and formative nature, carried out during the course (inside and outside the classroom). The course project is one of the most important activities within continuous assessment.

The evaluation tests consist of a part with questions about concepts associated with the learning objectives of the course with regard to knowledge, and a part with a set of application exercises.

Criteria for re-evaluation qualification and eligibility: Students who that not reach a minimun of five and have a score strictly greater than zero in the assessment activities will have the opportunity of carrying out a re-evaluation exam during the period specified in the academic calendar. The maximum mark for the re-evaluation exam will be five over ten (5.0). The non-attendance of a student to the re-evaluation test, in the date specified will not grant access to further re-evaluation tests.

Test Rules

If any of the continuous assessment activities are not carried out in the scheduled period, it will be considered as a zero score.

Office Hours

In order to get an appointment, an email should be sent to the professor.

Bibliography

Basic

- Kottegoda, N.T.; Rosso, R. Applied statistics for civil and environmental engineers. 2nd ed. Oxford: Blackwell, 2008. ISBN 9781405179171.
- Devore, J.L. Probabilidad y estadística para ingeniería y ciencias. 9a ed. México: Cengage Learning, 2016. ISBN 9786075228280.
- Ross, S. M. Introduction to Probability and Statistics for Engineers and Scientists. 4th ed. Amsterdan: Elsevier, 2009. ISBN 9780123704832.
- DeGroot, M.H.; Schervish, M.J. Probability and statistics. 4th ed. Boston: Pearson, 2012. ISBN 9780321709707.
- Ang, A. H-S; Tang, W.H. Probability concepts in engineering. Emphasis on Applications to Civil and Enviromental Engineering. 2nd ed. New York: Wiley, 2007. ISBN 9780471720645.

Complementary

- Kabacoff, R.I. R in action: data analysis and graphics with R. 2nd ed. Shelter Island, [New York]: Manning, 2015. ISBN 9781617291388.

- Pawlowsky-Glahn, V. [et al.] (eds.). Modeling and analysis of compositional data: theory and applications. Hoboken, N.J.: Wiley, 2015. ISBN 9781119003144.
- Castillo, E. [et al.]. Extreme value and related models with applications in engineering and science. Hoboken, New Jersey: John Wiley & Sons, 2005. ISBN 047167172X.
- Canavos, G.C. Probabilidad y estadística : aplicaciones y métodos. México [etc.]: McGraw Hill, 1988. ISBN 9684518560.
- Mood, A. M; Graybill, F.A; Boes, D.C. Introduction to the theory of statistics. 3rd. McGraw-Hill, 1974. ISBN 0070428646.