

Degree programme	Type	Course
Applied Statistics	FB	1

Contact lecturer

Name : Francisco Javier Delgado Vences

Email : francisnojavier.delgado@uab.cat

Teaching staff

Queralt Miro Catalina

Group languages

You can consult this information at the [end](#) of the document.

Prerequisites

A good knowledge of the contents of the subjects studied during the first semester is considered very important, especially those of Introduction to Probability, Calculus 1 and Exploratory data analysis.

Objectives

This subject is the first of the Degree dedicated to Statistical Inference, which is the part of the Statistics that allows to obtain, in a controlled way, information about a population based on the data of a "representative" sample. The subject has a central character within the studies, as different concepts and techniques that will be used in many of the subjects that will be studied from now on are introduced here. Specifically, an introduction to the Statistics will begin, and then the estimation of parameters, both punctual and by confidence intervals, will be treated, as well as classical parametric hypotheses testing, both for one and two normal and dichotomous populations, ending with the chi-square tests.

Learning outcomes

- CM08 (Determine the sample size and the sampling strategies required to conduct a specific study in the field of applications.) Determine the sample size and the sampling strategies required to conduct a specific study in the field of applications.
- KM09 (Discover the fundamental properties of estimators: invariance, sufficiency, efficiency, bias, mean square error and asymptotic properties, in the classical and Bayesian domains.) Discover the fundamental properties of estimators: invariance, sufficiency, efficiency, bias, mean square error and asymptotic properties, in the classical and Bayesian domains.
- KM11 (Identify exact and asymptotic sampling distributions of different statistics.) Identify exact and asymptotic sampling distributions of different statistics.
- SM09 (Analyse data through different inference techniques using statistical software.) Analyse data

- through different inference techniques using statistical software.
- SM10 (Use different estimation methods depending on the context of application.) Use different estimation methods depending on the context of application.

Contents

1. Statistical inference: introduction and basic concepts

1.1. Introduction, objectives and program of the assignment

1.2. Population and sample.

1.3. Statistics.

1.4. Distribution of the proportion, mean and sample variance: normal distribution and Central Limit Theorem.

2. Punctual forecasting

2.1. The problem of punctual estimation. Parameter and estimator

2.2. Properties of Estimators

2.2. Estimator of a proportion

2.3. Estimator of a mean and population variance.

2.4. How to find a good estimator: the method of moments and the method of maximum likelihood.

3. Estimated by confidence intervals

3.1. Concept of Confidence Interval

3.2. Confidence interval for a proportion

3.3. Confidence intervals for the average (with known population variance or unknown population variance). Normal case and general case

3.4. Interval of confidence for variance. The Normal Case

3.5. Interval for the difference in averages (paired data or independent samples with known, unknown and equal population variances, or unknown and different). Normal case and general case

4. Proof of hypothesis for a population. Basic Concepts

4.1. The problem of a hypothesis test. Type of hypothesis. Type I and II errors

4.2. Level of significance and critical region. The P-value. The Power Function

4.3. Contrast to a proportion.

4.4. Contrast to the average population. The Z-test and the Student t-test

4.5. Determining sample size to ensure a given level of confidence and accuracy

4.6. Contrast to Variation

4.7. Relationship between the region of acceptance of a hypothesis test and the confidence interval

5. Contrasts of hypothesis and confidence intervals to compare two populations

5.1. Comparison of the averages of two populations from paired data

5.2. Comparison of the averages of two independent populations

5.3. Comparison of the proportions of two independent populations

5.4. Comparison of the variances of two independent normal populations. The F-Test

5.5. Goodness-of-fit tests: Chi-square, Kolmogorov-Smirnov.

5.6 Tests of independence

IMPORTANT: In teaching, the gender perspective involves reviewing androcentric biases and questioning the assumptions and hidden gender stereotypes. This revision involves including the contents of the subject the knowledge produced by scientific women, often forgotten, seeking the recognition of their contributions, as well as that of their works in the bibliographical references. Efforts will also be made to introduce the most practical part of the subject, the analysis and comparison of statistical data by sex, commenting on the classroom causes and the social and cultural mechanisms that can sustain the observed inequalities.

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Workshop resolution	20	0.8	
Problems resolution	25	1	
Exams	15	0.6	
Theory classes	30	1.2	
Problems classes	18	0.72	
Practical classes	12	0.48	

The subject is structured from theory classes, problems and practices.

In theory classes we will introduce the concepts and techniques described in the course program. Considering that the content is essentially the standard of a first course of statistical inference, one can follow the course making use of the recommended basic bibliography. The material corresponding to each topic explained in the classroom will also be posted on the Virtual Campus.

The classes of problems are intended to work and understand statistical concepts. In the Virtual Campus the lists of problems will be posted and, when they have already been solved in class, also the solutions.

The objective of the practices is the use of statistical software R, to obtain and clarify the results of the procedures that have been introduced in theory classes and problems. In the Virtual Campus the statement of each practice will be posted in advance.

IMPORTANT: To work more comfortably with R, it is recommended to use the RStudio interface: it is free, "Open source" and works with Windows, Mac and Linux. <https://www.rstudio.com/>

OBSERVATION: The gender perspective in teaching goes beyond the contents of the subjects, since it also implies a revision of the teaching methodologies and of the interactions between the students and the teaching staff, both in the classroom and outside. In this sense, participatory teaching methodologies, where an egalitarian, less hierarchical environment is generated in the classroom, avoiding stereotyped examples in gender and sexist vocabulary, with the aim of developing critical reasoning and respect for the diversity and plurality of ideas, people and situations, tend to be more favorable to the integration and full participation of the students in the classroom, and therefore their effective implementation in this subject will be sought.

Annotation: within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continuous assessment activities

Title	Weight	Hours	ECTS	Learning outcomes
Practical exam (P)	0,30	12	0.48	SM09, SM10
midterm exam	0,35	8	0.32	CM08, KM09, KM11, SM10
Final exam / Reassessment (E)	0,35	10	0.4	CM08, KM09, KM11, SM10

The course grade will be obtained from a first midterm exam E1 (35%), continuous assessment of the course practicals P (30%), and a second midterm exam E2 (35%). With the grades E1, P, and E2, the course grade, N, is calculated as follows:

$$N = 0.35 \times E1 + 0.35 \times E2 + 0.30 \times P$$

Important: The assessment can be:

- Either Continuous Assessment, following the established calendar throughout the semester.
- Or Single Assessment, where on the same day as the Continuous Assessment's second midterm, the student must take all three tests at the same time (E1, E2, and P). Important: to be eligible for the Single Assessment, students must use the appropriate channel and submit the request within the deadlines established by the Faculty of Sciences.

Given the quantitative nature of the course, Continuous Assessment is recommended by the course instructors and the degree program.

Retake and/or improvement of the exam grade: The student passes the course if N is greater than or equal to 5 and, at the same time, E1, E2, and P are greater than 4. Otherwise, or if the student wishes to improve their grade, there is a possibility to improve the exam grade portion for E1 or E2 through a retake exam. Thus, the final course grade is obtained from this retake grade:

$$NF = 0.35 \times \max(E1, E1_retake) + 0.35 \times \max(E2, E2_retake) + 0.30 \times P$$

- Observation 1: The continuous assessment grade P cannot be retaken.
- Observation 2: The student is considered to have attended the exam call if they sit for either of the two exams that result in grades E1 or E2. Otherwise, the status will be "Not Attended" (No Presentado), even if they have a continuous assessment grade P.
- Observation 3: Taking the retake exam means obtaining a maximum grade of 9 for the course.
- Observation 4: To obtain honors (Matrícula de Honor), it is recommended to achieve an excellent grade in all three parts. The honors awards will be decided before the resit exam.

Bibliography

- Berger, R.L., Casella, G.: Statistical Inference. Duxury Advanced Series. 2002.
- DeGroot, M. H.: Schervish, M.J. Probability and Statistics. Pearson Academic. 2010.
- Canavos, G. C.: *Probabilidad y estadística* (p. 651). México: McGraw Hill. 1987.
- Daalgard, P.: Introductory Statistics with R. Springer. 2008.
- Daniel, W.W.: Biostatistics. Wiley. 1974.
- Silvey, S.D.: Statistical Inference. Chapman&Hall. 1975.
- Novales, A.: Econometria. McGraw-Hill 2000
- Peña, D.: Estadística. Fundamentos de estadística. Alianza Universidad. 2001.
- R Tutorial. An introduction to Statistics. <https://cran.r-project.org/manuals.html>. juny 2019.
- Silvey, S.D.: Statistical Inference. Chapman&Hall. 1975.

Software

The software to be used to work with the data will be Excel and the statistical program R.

Course groups and languages

The information provided is provisional until November 30. After this date, you will be able to consult the language of each group through this [link](#). To access the information, you will need to enter the course CODE

Type of teaching	Group	Language	Semester	Shift
(TE) Theory	1	Catalan	second semester	afternoon
(PAUL) Classroom practices	1	Catalan	second semester	afternoon
(PLAB) Practical laboratories	1	Catalan	second semester	afternoon
(PAUL) Classroom practices	2	Catalan	second semester	afternoon
(PLAB) Practical laboratories	2	Catalan	second semester	afternoon