UAB
Universitat Autònoma de Barcelona

Probabilistic and Statistical Descriptions

Code: 104348
ECTS Credits: 6

2024/2025

Degree	Type	Year
2503758 Data Engineering	FB	2

Contact

Name: Joan Porti Pique

Email: joan.porti@uab.cat

Teachers

Joan Porti Pique

Teaching groups languages

You can view this information at the of this document.

Prerequisites

There are no prerequisites. It is recommended to have followed the courses in Algebra and Calculus.

Objectives and Contextualisation

The goal of the course is to introduce the basic tools of probability and statistics to analyze data from natural, social or economic phenomena or experiments, focusing on its correct use and the interpretation of the results. The theory and problem sessions are going to be complemented with practice classes, with the aim that students learn to use computer tools for statistical analysis.

Competences

- Design efficient algorithmic solutions to computational problems, implement them in the form of robust software developments which are structured and easy to maintain, and verify their validity.
- Develop critical thinking and reasoning and know how to communicate it effectively in both your own language and in English.
- Search, select and manage information and knowledge responsibly.
- Use techniques of probability and statistics to analyse and model complex phenomena and solve optimisation problems.

- Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Learning Outcomes

1. Acquire and consolidate the basic concepts of probability theory, mathematical expectation and conditional probability.
2. Choose and apply the most suitable data structures and strategies to solve an informatics problem efficiently.
3. Develop critical thinking and reasoning and know how to communicate it effectively in both your own language and in English.
4. Develop programmes that are well documented, using a good programming style and be able to debug, test and correct them.
5. Identify the most common descriptors for a data set and assess their applicability to a known data set.
6. Identify the statistical distributions and their application to engineering problems.
7. Search, select and manage information and knowledge responsibly.
8. Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Content

Topic 1. Descriptive statistics.

Descriptive study in one variable: frequency distributions, graphic representations, numerical summaries (position, dispersion and shape measurements).
Descriptive study in two variables: correlation and regression line, tables of contingency.

Topic 2. Probability.

Notion of probability ant main properties. Conditional probability. Formula of total probabilities. Bayes Formula.
Independence of events.
Expectation and variance of a random variable.
Discrete random variables. Bernoulli, Binomial and Poisson distributions.
Continuous random variables. Normal and Exponential distributions. Approximation of the Binomial by the Normal distribution.
Independence of random variables.
Central limit theorem.

Topic 3. Statistical inference.

Sample and population. Most frequent statistics.
Confidence intervals for the mean and for the variance of a Normal population and for the proportion.
Hypothesis test concept. Test for the mean and for the variance of a Normal population.
Test for the proportion. Test of comparison of means and variances for two Normal populations. Comparison test of proportions.
Chi-square test for the goodness of fit and independence.

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes

Type: Directed			
Practices in the computer room	12	0.48	1, 2, 3, 4, 5, 6, 7, 8
Problem sessions	12	0.48	1, 2, 3, 4, 5, 6, 7, 8
Theoretical classes / lectures	26	1.04	1, 2, 3, 4, 5, 6, 7, 8
Type: Supervised			
Tutoring and consultations	10	0.4	1, 2, 3, 4, 5, 6, 7, 8
Type: Autonomous			
Independent study and preparation	60	2.4	1, 2, 3, 4, 5, 6, 7, 8

The center of the learning process is the work of the student. The student learns working, being the mission of the teaching staff help him/her in this task by providing information or showing him/her the sources where one can get it and directing his/her steps in a way that the learning process can be carried out effectively. In line with these ideas, and in accordance with the aims of the subject, there are theory classes (lectures), problem sessions, and practices sessions. In these sessions and with individual work of the student, the specific skills are achieved.

New material will be mainly introduced in the lectures, but explanations must be complemented with the autonomous study and personal work of the student, with the help of the references and the material in the CV.

The problems sessions will be devoted to the oriented resolution of some proposed problems. Attention will be paid to correctness and rigor, as well as to vocabulary, mathematical expression and clarity in writing. There will be a partial test of theory and problems.

In the practice sessions, software for statistical analysis will be introduced. Descriptive and inferential methodologies are introduced. These tools will be used to solve problems and to do some assignments that students will have to hand in.

The Campus Virtual UAB is a key tool for following the subject: access to material, consultation of deadlines and monitoring of the pace of the course.

The lectures, in which mathematical models are discussed, with the problem sessions, in which several solutions to problems will be proposed, and with the individual work of the student allow to reach the transversal skills. (T01.01,T01.02 and T01.03).

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

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Parcial exams	70%	5	0.2	1, 2, 3, 4, 5, 6, 7, 8
Practices assignments	30%	20	0.8	1, 2, 3, 4, 5, 6, 7, 8

Recovery exam	70%	5	0.2	1, 2, 3, 4, 5, 6, 7, 8

The continued evaluation of theory and problems consists of two partial exams, the first one with a weight of 30% and the second one with a weight of 40%. These two partials will be the recoverable part of the subject.

The evaluation of the practices with computer will have a weight of 30% in the final evaluation of the subject. To evaluate them, at the end of each session, the students will submit the practice assignment and complete a test . This part is not recoverable.

To participate in the recovery examination, the students must have been previously evaluated in a series of activities whose weight equals to a minimum of 2/3 of the total grade of the subject.

The unique evaluation consists of a single summary exam in which the contents of the entire theory and problems program of the subject will be assessed. The grade obtained in this final exam will account for 70% of the final grade of the subject. The date of this exam will coincide with that fixed in the calendar for the last continued evaluation exam and the same recovery system will be applied as for the continued evaluation.

The evaluation of practice activities and the delivery of assignments will follow the same procedure as the continued evaluation. The grade obtained will have a weight of 30% in the final evaluation of the subject.

A weighted average of a minimum of 4 out of 10 is required in the partial exams, in the final exam or in their recovery. A minimum grade of 4 out of 10 is also required in the average grade of the practices assignments. If the minimum of each module is reached, the final grade is the weighted mean. Otherwise, the final grade is the minimum between the weighted mean and 4.5 (out of 10).

Those who have not taken tests that add up to 50% of the course will be considered Non-Assessable.

In order to pass the course with honors, the final grade must be equal or higher to 9 points. This will be given to students that, according to the criterion of the professors, have reached in a brilliant manner all the goals of the subject.

Assessment and delivery of works dates will be published on the virtual campus, and the programming may change because of adaptation to possible incidents. Any modification will be informed in the Campus Virtual, which is the usual exchange of information platform between teachers and students.

Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the academic regulations in force, assessment activities will receive a zero (0) score whenever a student commits academic irregularities that may alter such assessment. The assessment activities qualified in this way and by this procedure will not be recoverable. This implies a failing grade in the subject with a final grade not exceeding 4.5 out of 10. These irregularities include, among others: the total or partial copy of a practice, report, or any other evaluation activity; to let copy; to have communication devices (such as mobile phones, smartwatches, etc.) accessible during face-to-face assessment tests.

There is no differentiated treatment for repeat students.

Bibliography

Bardina, X. Farré, M. *Estadística descriptiva.* Manuals UAB, 2009.

Besalú, M. Rovira C. *Probabilitats i estadística.* Publicacions i Edicions de la Universitat de Barcelona, 2013.

Delgado, R. *Probabilidad y Estadística para ciencias e ingenierías.* Delta, Publicaciones Universitarias. 2008.

Devore, J. L. *Probabilidad y estadística para ingeniería y ciencias.* Thomson, 2005.

Montgomery, D. C. Runger, G. C. *Probabilidad y estadística aplicadas a la ingeniería.* Limusa Wiley, 2002.

Walpole, R. Myers, R. H. Myers, S. L. *Probabilidad y estadística para ingenieros.* Prentice Hall, 1999.

Software

In the computer practice sessions, the student will learn to use the free software R with the integrated development environment RStudio, in order to apply the statistical tools for the descriptive analysis of data sets and statistical inference.

Language list

Name	Group	Language	Semester	Turn
(PAUL) Classroom practices	81	Catalan	first semester	morning-mixed
(PAUL) Classroom practices	82	Catalan	first semester	morning-mixed
(PLAB) Practical laboratories	1	Catalan	first semester	morning-mixed
(PLAB) Practical laboratories	2	Catalan	first semester	morning-mixed
(TE) Theory	81	Catalan	first semester	morning-mixed