Universitat Autònoma de Barcelona

2019/2020

Probabilistic and Statistical Descriptions

Code: 104348
ECTS Credits: 6

Degree	Type	Year	Semester
2503758 Data Engineering	FB	2	1

Contact

Name: Xavier Bardina Simorra
Email: Xavier.Bardina@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

Teachers

Giulia Binotto
Xavier Bardina Simorra

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 phenomena or experiments, focusing in its correct use and the interpretation of the results. The theory and problem sessions are going to be complemented with practice classes in the computer room, with the aim ofusing computer tools.

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 statistics. Descriptive study in one variable: categorical (sector diagram) and quantitative (mean, deviation, bar diagram and histogram). Descriptive study in two variables: categorical (contingency tables) and quantitative (regression line, correlation coefficient). Software tools for statistical analysis.

Topic 2. Probability.

Notion of probability. Conditioned probability and independence of events. Statistical distributions. Examples of application to engineering. Random variables. Expected value and variance of a random variable. Examples: binomial and normal. Approximation of the binomial to normal. Independence of random variables. Basic notions of stochastic processes, Poisson and exponential distributions.

Topic 3. Statistical inference.

Sample and population. Most frequent statistics. Confidence intervals: for the average and for the variance of anormal population and for the proportion. Hypothesis test. Mutual information of random variables. Principal component analysis.

Methodology

There are theory classes (lectures), problem sessions, and practices sessions. In these sessions and with individual work the especific 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. There will be a partial test of theory and problems. The problems sessions will be devoted to the oriented resolution of some proposed problems. Attention will be payed to corection and rigorousness, as well as to vocabulary, mathematical expression and clarity in writing. In the practice sessions we shall introduce software with applications to statistics (spread sheets and statistics packages). Descriptive and inferential methodologies are introduced. These tools will be used to solve problems and will be used to work (individually) with real data.The Campus Virtual UAB is a key tool to follow the class: access to material, check information and following the course.T

ransversal skills. 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).

Activities

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

Assessment

The assessment consists of two modules:

Module I: consist of

a) A partial test (20%)

b) A final test (40%)

c) A problems session (15%)

Module II: participation and assessment in the laboratory practices. (25%)

Finally, during the exams period there will be a test to recover the mark of the whole Module I. Tests (a) and (b) contain a written exercise to guarantee reaching the transversal skills. In order to be able to participate inthe compensation test, it is required that the student has followed evaluation activities for a minimum of two thirds of the total mark. Module II is not compensated.The minimal marks required are 4 out of 10 for Module I (or for its compensation) and for Module II. If the minimum of each module is reached, the final mark is the weighted mean. Otherwise the final mark is the minimum between the weighted mean and 4.5 (out of 10). Assessment dates will be published on the virtual campus and on the presentation slides, and the programming may change because of adaptation to possible incidents. Any modification will always be informed in the Campus Virtual, which is the usual exchange of information platform between teachers and students. For the assessment activity, a place, date and time of review will be indicated in which the student will be able to review the activity with the teacher. In this context, claims can be made about the activity grade, which will be graded by the teachers responsible for the subject. If the student does not take part of this review,this activity will not be reviewed later. 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. Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the academic regulations in force, assessment activities will receive a 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. If you need to pass any of these assessment activities to pass the subject,

this subject will be failed directly, without opportunity to recover it in the same course 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, smart watches, etc.) accessible during theoreticalpractical assessment tests (individual exams).

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Exercices and problems	15%	10	0.4	1, 7, 4, 3, 5, 6, 2, 8
Practices in the computer room	25%	15	0.6	1, 7, 4, 3, 5, 6, 2, 8
Written test	60%	5	0.2	1, 7, 4, 3, 5, 6, 2, 8

Bibliography

Arnold O. Allen, Probability, Statistics, and Queueing Theory with Computer Science Applications, Academic Press, Inc. 1990

Jay L. Devore. Probabilidad y estadística para ingeniería y ciencias. Thomson. 2005

Rosa Millones, Emma Barreno, Félix Vásquez y Carlos Castillo, Estadística aplicada a la ingeniería y los negocios. fondo Editorial, Universidad de Lima. 2015.

Douglas C. Montgomery y George C. Runger, Probabilidad y estadística aplicadas a la ingeniería. Limusa Wiley. 2002

Ronald E. Walpole, Raymond H. Myers y Sharon L. Myers. Probabilidad y estadística para ingenieros.Prentice Hall. 1999