UAB
Universitat Autònoma de Barcelona

Probability and Statistics

Code: 106483
ECTS Credits: 6
Type: FB   Year: 1   Semester: 2

ECTS Credits: 6

Degree

1488 - Artificial Intelligence

The contents of this guide are provisional and may be subject to minor changes. The final version of the guide will be available at the beginning of the semester.

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities

Use of languages

Contact

Name: Ramon Antoine Riolobos

Email: Ramon.Antoine@uab.cat

Principal working language: English

Some groups entirely in English: Yes

Some groups entirely in Catalan: No

Some groups entirely in Spanish: No

Prerequisites

There are no prerequisites besides the main contents of Mathematics from high school.

Objectives and Contextualisation

The goal of the course is to introduce the basic tools of probability and statistics used to analyse data coming from either natural, experimental, social or economic phenomena. A special focus will be given to the correct use of these tools and the interpretation of the results by providing the student with the required theoretical background. Moreover, a part of the course will be dedicated to introduce and familiarize the student with the use of the most common computer tools for statistical analysis.

Competences

- To know, understand, use and apply in a proper way the mathematical foundations that are necessary to develop systems for reasoning, learning and manipulation of large volumes of data.
- To develop critical thinking to analyse in a grounded and reasoned way alternatives and proposals, both own and others.
- To analyse and solve problems effectively, generating innovative and creative proposals to achieve the objectives.
- To work autonomously, with responsibility and initiative, planning and managing the time and available resources, adapting to unforeseen situations

Learning outcomes

- To know, understand and apply the basic concepts of probability theory.
- To identify the main statistical distributions and their application to different problems.
- To understand the concept of regression and its application.
- To be able to perform hypothesis tests and correctly interpret the result.

Contents

- **Topic 1.** *Descriptive statistics.*
  - Descriptive study in one variable: frequency distributions, graphic representation and numerical summaries.
  - Descriptive study in two variables: correlation and regression. Contingency tables.
- **Topic 2.** *Probability.*
  - Notion of probability and main properties. Conditional probability. 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.
  - Independence of random variables.
  - Central limit theorem.
- **Topic 3.** *Statistical inference.*
  - Sample and population. Most frequent statistics.
  - Confidence intervals for a single mean or proportion.
  - Hypothesis tests. Main theoretical concepts; p-values.
  - Test for the mean and for the variance of a Normal population,. Test for comparison of means, proportions...
  - Chi-square goodness of fit test for independence.

Methodology

Besides the mandatory student's personal work, the course will have three distinguished types of activities: The core and main theoretical lessons, problem sessions, and practice in the computer lab. With the correct combination of this activities the specific skills will be achieved.

New material will be mainly introduced in lectures were the professor will explain the main theoretical results backing the tools that will be used throughout. Although these explanations can be complemented with support exercises, those will be mainly worked during the problem sessions where the teacher will provide an oriented resolution of some proposed problems. The student is encouraged to attend the sessions having previously tried the exercises by himself.

The main theoretical part together with typical exercises will have a partial evaluation halfway along the semester. This should provide the student with a measure of his or her progression.

There will be practice sessions with specialized computer software for statistical analysis. These sessions will have the double effect of introducing the student with the typical procedures of data analysis, as well as providing lots of examples illustrating the subject.

The practical sessions will have an independent evaluation with assignments to hand in, possibly at the end of some of the sessions.

In all the evaluations, special attention will be given to the correctness and validity of the assertions and arguments used. These includes vocabulary, mathematical correctness and clarity in writing.

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.

Activities

Title	Hours	ECTS	Learning outcomes
Type: Directed			
Theoretical lessons	12	0.48	
Problem sessions	12	0.48	
Practice at the Computer Lab	26	1.04	
Type: Supervised			
Tutorial and consultation	10	0.4	
Type: Autonomous			
Study and preparation	60	2.4	

Assessment

The evaluation of theory and problems will consists of two partial exams. The first one, with a weight of 30% and the second one with a weight of 40%. For these evaluations there will be a second-chance exam at the end of the semester. The remaining 30% of the evaluation weigh will come from the computer practices. This will be obtained from different assignments delivered throughout the course, for which there will not be a second chance evaluation.

In order to attend the recovery examination, students must have been previously evaluated in a series of activities whose weight equals at least two thirds of the total.

A weighted average of a minimum of 4 out of 10 is required in the partial exams or in their recovery. A minimum grade of 4 out of 10 is also required in the average grade of the practice 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 means 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 honours, the final grade must be equal to or higher than 9 (out of 10). This will be given to students that, according to the criterion of the professor, have reached in a brilliant manner all the goals of the subject.

The dates for the assessments and delivery of assignments will be published in a dedicated webpage for the course.

Assessment activities

Title	Weighting	Hours	ECTS	Learning outcomes
First partial exam	30%	2	0.1	
Second partial exam	40%	3	0.1	
Hand in assignments	30%	20	0.8	
Recovery exam	70%	5	0.2	

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

One of the main tools for statistical analysis and development and which has gained a growing popularity in academia is the R language. The R project is is a free software environment providing a large set of libraries and tools aimed at statistical computing and graphical representation of data.

The student will learn the basics of the R language though the use of the integrated development environment RStudio

R: https://www.r-project.org/
Rstudio: https://www.rstudio.com/

No special version of the software is required for the aims of this course.