COURSE DATA

Data Subject	
Code	36409
Name	Probability and simulation
Cycle	Grade
ECTS Credits	6.0
Academic year	2023 - 2024

Study (s)

Degree	Center	Acad. year	Period
1406 - Degree in Data Science	School of Engineering	1	Second term

Subject-matter

Degree	Subject-matter	Character
1406 - Degree in Data Science	2 - Statistics	Basic Training

Coordination

Name	Department
NAVARRO QUILES, ANA	130 - Statistics and Operational Research

SUMMARY

The theory of probability is the area of mathematics which allows us measuring the uncertainty around us. It is the language which allows Statistics to develop as a science.

The main goal of this subject is to learn the language of probability providing students with the following capacities:

- To understand and use the probability calculus.
- To understand the idea of random variable and random vector.
- To recognise the main probability distributions for continuous and discrete variables as well as its interpretation as a theoretical model for a given population.
- To understand the concept of joint, conditional and marginal distribution.

- To calculate the main moments associated with a random variable or vector.
- To simulate from a probability distribution (Monte Carlo methods).

PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

Other requirements

Knowledge of the concepts of function, limit, derivative and integral studied in the subject Mathematical Analysis (cod. 36407)

OUTCOMES

1406 - Degree in Data Science

- (CG01) Knowledge of basic subjects and technologies that enable students to learn new methods and technologies, and to provide them with versatility to adapt to new situations.
- (CG04) Ability to work in a multidisciplinary group in a multilingual environment and to communicate, orally and in writing, knowledge, procedures, results and ideas related to data science.
- (CT01) To be able to access (bibliographical) information tools and appropriately use them in the development of their daily tasks.
- (CE09) To methodologically know and apply the concepts and techniques of probability and statistics necessary for the extraction of useful knowledge from data analysis.
- (CE15) Ability to model and analyse the uncertainty in data-based studies, as well as to know how to interpret and contextualise the results obtained.
- (CB2) Students must be able to apply their knowledge to their work or vocation in a professional manner and have acquired the competences required for the preparation and defence of arguments and for problem solving in their field of study.
- (CB5) Students must have developed the learning skills needed to undertake further study with a high degree of autonomy.

LEARNING OUTCOMES

To understand the concepts of uncertainty, randomness and probability. (CG01, CB5, CE15, CE09)

To know how to solve probability calculus problems. (CB2, CB5, CE09, CE15)

To understand the concept of random variable. (CG01, CB2, CB5, CE09, CE15)

To know the most important probability distributions and the calculus of their associated moments, quantiles and probabilities. (CG01, CB2, CB5, CT01, CE09, CE15)

To understand the concept of random vector and its associated probability distributions: joint, marginal and conditional. (CG01, CG04, CB2, CB5, CT01, CE09, CE15)

Simulation of probability distributions related to random variables and approximated calculus of its probabilistic features. (CG01, CG04, CB5, CT01)

DESCRIPTION OF CONTENTS

1. Introduction to probability

History of probability
Experiments and events
Definition and axioms of probability
Total probability theorem
Conditional probability
Bayes Theorem

2. Random variables and probability function

Definition of random variable
Types of random variables
Probability density function
Cumulative distribution function
Moments of a random variable

3. Main probability distributions

Bernoulli and binomial density functions
Hypergeometric density function
Poisson density function
Negative binomial density function
Normal distribution
Law of large numbers
Central limit theorem
Density functions from the normal distribution (chi-Squared, t-student)
Exponential, gamma and beta density functions

4. Random vectors

Definition of random vector
Joint, conditional and marginal density function.
Covariance and correlation

5. Simulation and Monte Carlo methods

Monte Carlo methods
Introduction to Markov chains
Introduction to Markov Chain Monte Carlo methods.

WORKLOAD

ACTIVITY	Hours	% To be attended
Theory classes	32,00	100
Laboratory practices	20,00	100
Classroom practices	8,00	100
Attendance at events and external activities	2,00	0
Development of group work	10,00	0
Development of individual work	5,00	0
Study and independent work	20,00	0
Readings supplementary material	5,00	0
Preparation of evaluation activities	10,00	0
Preparing lectures	20,00	0
Preparation of practical classes and problem	10,00	0
Resolution of case studies	5,00	0
Resolution of online questionnaires	3,00	0
TOTAL	**150,00**	

TEACHING METHODOLOGY

MD1 – Theoretical activities: Exposition of the concepts with the enrollment of the student by solving specific questions. Completion of individual evaluation questionnaires. (GC01, CB5, CE09)

MD2 – Practical activities. Learning by solving problems, exercises and case studies through which competences are acquired on the different aspects of the subject. (CB2, CE15)

MD4 - Lab work and / or computer classroom. Learning through the performance of activities developed individually or in small groups and carried out in laboratories and / or computer rooms. (CG04, CT01, CE09, CE15)

EVALUATION

SE1 – Written exam with theoretical and practical questions that will represent 40% of the mark. (GC01, CB5, CE09)

SE2 – Preparation of works and/or a report in PDF, alone and/or in groups, of a report of the issues addressed in the practical sessions. This will represent a 40% of the final mark. The evaluation of this part may be recovered by means of a practical examination that will be carried out on the same date as the theoretical examination of the second call. (CB2, CE15)

SE3 – Continuous evaluation based on the degree of implication of the student in the learning process. This degree of implication can be marked in two different ways between which the student should choose:

i) Oral expositions of the theoretical concepts studied (alone and/or in groups) and problem solving in the classroom (this implies attending the class).

ii) Exercises that students should solve on their own and deliver in the virtual classroom in a given date.

No matter the evaluation process chosen, those activities will not be recoverable and will represent a 20% of the final mark. (CG04, CT01, CE09, CE15).

It will be necessary to get more than 4 in each part to be able to average and more than 5 as final grade to pass.

In any case, the evaluation system would be governed by the *Reglamento de Evaluación y Calificación de la Universitat de València para Grados y Másteres*:
https://www.uv.es/graus/normatives/2017_108_Reglament_avaluacio_qualificacio.pdf

REFERENCES

Basic

- Joseph K. Blitzstein and Jessica Hwang. Introduction to Probabilitiy. CRC Press (2015) http://proquest.safaribooksonline.com/9781466575578?uicode=valencia
- Robert P. Dobrow. Probability with applications in R. John Wiley & Sons (2014) http://syndetics.com/index.aspx?isbn=9781118241257/summary.html&client=valenciah&type=rn12

Additional

- Deborah Rumsey. Probability for Dummies. Wiley. (2006)
- Lola Ugarte, Ana Militino y Alan Arnholt. Probability and statistics with R Second ed Chapman & Hall (2016)
- Mark J. DeGroot, Morris H., Schervish
  Probability and Statistics 4th edition Pearson Education Limited (2014