Course guide
270601 - SMDE - Statistical Modelling and Design of Experiments

Last modified: 02/02/2024

Unit in charge: Barcelona School of Informatics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.

Degree: MASTER'S DEGREE IN INNOVATION AND RESEARCH IN INFORMATICS (Syllabus 2012). (Compulsory subject).

Academic year: 2023 **ECTS Credits:** 6.0 **Languages:** English

LECTURER

Coordinating lecturer: PAU FONSECA CASAS

Others:

Primer quadrimestre:
NIHAN ACAR DENIZLI - 10
PAU FONSECA CASAS - 10

Segon quadrimestre:
NIHAN ACAR DENIZLI - 10
PAU FONSECA CASAS - 10

PRIOR SKILLS

Students must have sufficient knowledge of algebra and mathematical analysis to assimilate the concepts related to algebra of sets, numerical series, functions of real variables of one or more dimensions, derivation and integration.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.

Generical:
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.

Transversal:
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

TEACHING METHODOLOGY

The course is practical and aims that students will be able, once the course is completed and from the work done in the sessions, to solve real problems similar to those developed in class.

LEARNING OBJECTIVES OF THE SUBJECT

1.Applying the mathematical formalism to solve problems involving uncertainty.
2.Applying the queuing models for computer systems performance evaluation and/or configurations analysis.
3.Ability to design, conduct experiments and analyze results.

STUDY LOAD

Type	Hours	Percentage
Hours medium group	13,5	9.00
Hours small group	27,0	18.00
Hours large group	13,5	9.00
Self study	96,0	64.00

Total learning time: 150 h

CONTENTS

Introduction to probability

Description:
Students should feel comfortable with the use of set notation and basic statistical terminology. Likewise, the student should be able to write the sample space of simple experiments, including sampling with replacement (like throwing coins or throwing dice), sampling without replacement, from Bernoulli trials and with rules of detention. Likewise, the student should be able to calculate the probabilities in simple cases of the above type of experiment.

Introduction to statistical estimation

Description:
Estimation, in the framework of statistical inference, is the set of techniques with the aim of give an approximate value for a parameter of a population from data provided by a sample. From the different methods that exist (point estimate, estimate intervals, or Bayesian estimation) we focus on the point estimate.

Analysis of data

Description:
The main objective of the section is to know the procedures associated with the analysis of variance (ANOVA terminology in English) and when is useful to be applied.This activity also introduces MANOVA, as a technique useful when there are two or more dependent variables. We also work with the techniques of linear regression and PCA, completing the repertoire of tools for data analysis.

Introduction to experimental design

Description:
Statistical experimental design, a.k.a. design of experiments (DoE) is the methodology of how to conduct and plan experiments in order to extract the maximum amount of information in the fewest number of runs (saving resources). In this section we describe different techniques to achieve that.

Introduction to queuing theory and simulation

Description:
This section will introduce the student to use the techniques of operations research for systems analysis for making quantitative decision in the presence of uncertainty through their representation in terms of queuing models and simulation.

ACTIVITIES

Introduction to probability

Description:
At the end of this activity the Student must be comfortable with using basic set notation and terminology. Also the Student must be capable of write down the sample space for simple experiments, including sampling with replacement (such as tossing coins or rolling dice), sampling without replacement, and Bernoulli trials with stopping rules. Also the Student must be capable of calculate probabilities in straightforward instances of the above types of experiment.

Specific objectives:
1

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 9h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 5h

Introduction to statistical estimation

Description:
Estimation, in the framework of statistical inference, is the set of techniques with the aim of give an approximate value for a parameter of a population from data provided by a sample. From the different methods that exist (point estimate, estimate intervals, or Bayesian estimation) we focus on the point estimate.

Specific objectives:
1

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 16h
Theory classes: 2h
Practical classes: 2h
Laboratory classes: 4h
Self study: 8h

ANalysis Of VAriance

Description:
The main objective of the activity is to know the procedures associated with the analysis of variance (ANOVA terminology in English) and when is useful to be applied.This activity also introduces MANOVA, as a technique useful when there are two or more dependent variables.

Specific objectives:
1

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 9h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 5h

Linear regression

Description:
Linear regression is a mathematical method that models the relationship between a dependent variable Y, independent variables Xi and a random term. This section will examine this method and explain its applicability from different examples.

Specific objectives:
1

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 10h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 6h

Principal component analysis

Description:
The principal component analysis (PCA, PCA in English), in statistics, is a technique that reduces the dimensionality of a dataset. This allows us to represent them graphically in two or three dimensional graphs of various variables grouped the data into factors, or components, consisting of the grouping variables. In this section we will work this technique from a practical point of view.

Specific objectives:
1

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 10h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 6h

Factorial design

Description:
Many experiments are conducted to study the effects of two or more factors. in this case the factorial designs are more efficient, presented in this section.

Specific objectives:
3

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 27h
Theory classes: 3h
Practical classes: 3h
Laboratory classes: 9h
Self study: 12h

Randomized blocks, Latin squares and related designs

Description:
In many research problems is necessary to design experiments that can systematically control the variability caused by different sources. This section will consider some experimental designs for solve these situations.

Specific objectives:
3

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 10h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 6h

Incomplete block design

Description:
Description incomplete blocks design, useful when you can not develop all combinations of treatment within each block.

Specific objectives:
3

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 10h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 6h

General structure of queuing models

Description:
Introduction to the theory of queue models. Notation Kendall. Discreet simulation using Event Schedulling.

Specific objectives:
2

Related competencies :
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 9h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 5h

Queuing models based on birth and death processes

Description:
Introduction to basic concepts and elements of the analysis of Markov processes. Markov queues.

Specific objectives:
2

Related competencies :
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 9h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 5h

Generalized queuing patterns with non-exponential distributions and serial exponential queues.

Description:
Networks of queues: open and closed networks. Introduction to general service distributions and multiple types of work.

Specific objectives:
2

Related competencies :
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 9h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 5h

Validation Verification and Accreditation

Description:
Techniques to Verify, Validate and do the Accreditation of the models.

Specific objectives:
2

Related competencies :
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 9h
Theory classes: 1h
Practical classes: 1h
Laboratory classes: 2h
Self study: 5h

First report

Specific objectives:
1

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 5h
Self study: 5h

Second report

Specific objectives:
2

Related competencies :
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 5h
Self study: 5h

Third report

Specific objectives:
3

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 5h
Self study: 5h

Final exam

Specific objectives:
1, 2, 3

Related competencies :
CG3. Capacity for mathematical modeling, calculation and experimental designing in technology and companies engineering centers, particularly in research and innovation in all areas of Computer Science.
CG1. Capability to apply the scientific method to study and analyse of phenomena and systems in any area of Computer Science, and in the conception, design and implementation of innovative and original solutions.
CEE2.3. Capability to understand models, problems and mathematical tools to analyze, design and evaluate computer networks and distributed systems.
CEE4.1. Capability to analyze, evaluate and design computers and to propose new techniques for improvement in its architecture.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.
CTR6. REASONING: Capacity for critical, logical and mathematical reasoning. Capability to solve problems in their area of study. Capacity for abstraction: the capability to create and use models that reflect real situations. Capability to design and implement simple experiments, and analyze and interpret their results. Capacity for analysis, synthesis and evaluation.

Full-or-part-time: 12h
Guided activities: 2h
Self study: 10h

GRADING SYSTEM

The course will have different exercises that the students must solve during the course (80% of the final grade).
At the end there will be an exam that will weigh 20% of the final grade.

BIBLIOGRAPHY

Basic:
- Robinson, S. Simulation: the practice of model development and use. 2nd ed. Houndmills, Basingstoke, Hampshire, UK: Palgrave Macmillan, 2014. ISBN 9781137328038.
- Box, G.E.P.; Hunter, J.S.; Hunter, W.G. Statistics for experimenters : design, innovation, and discovery. 2nd ed. John Wiley and Sons, 2005. ISBN 0471718130.
- Montgomery, D.C. Design and analysis of experiments. 8th ed. John Wiley & Sons, 2013. ISBN 9781118097939.
- Bose, S.K. An Introduction to queueing systems. Kluwer Academic/Plenum Publishers, 2002. ISBN 0306467348.
- González, J.A. [et al.]. Estadística per a enginyers informàtics. Edicions UPC, 2008. ISBN 9788483019535.
- Baron, M. Probability and statistics for computer scientists. 3rd ed. Boca Raton, FL [etc.]: CRC Press, 2019. ISBN 9781138044487.

Complementary:
- Jain, R. The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling. John Wiley & Sons, 1991. ISBN 0471503363.
- Trivedi, K.S. Probability and statistics with reliability, queuing and computer science applications. 2nd ed. John Wiley & Sons, 2001. ISBN 0471333417.
- Hillier, F.S.; Lieberman, G.J. Introduction to operations research. 10th ed. Mcgraw-Hill, 2015. ISBN 9780073523453.
- Winston, W.L. Operations research: applications and algorithms. 4th ed. Brooks/Cole - Thomson Learning, 2004. ISBN 0534423620.
- O'Connor, P.D.T.; Kleyner, A. Practical reliability engineering. 5th ed. John Wiley & Sons, 2012. ISBN 9781119961260.
- Ross, S.M. Probability models for computer science. Harcourt/ Academic Press, 2002. ISBN 9780125980517.

RESOURCES

Hyperlink:
- http://cran.r-project.org/ - http://wiki.fib.upc.es/sim/index.php/Main_Page/en - http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/p