Course guide
220270 - 220270 - Applied Statistics in Industrial Engineering

Last modified: 19/04/2023

Unit in charge: Terrassa School of Industrial, Aerospace and Audiovisual Engineering
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.

Degree: MASTER'S DEGREE IN INDUSTRIAL ENGINEERING (Syllabus 2013). (Optional subject).

Academic year: 2023 **ECTS Credits:** 5.0 **Languages:** Catalan

LECTURER

Coordinating lecturer: Maria Albareda Sambola

Others:

PRIOR SKILLS

To follow this course it is necessary to have some basic knowledge on statistics (covariance, correlation coefficient, linear regression) and on linear algebra (matrix calculus and eigenvalues and eigenvectors)

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
1. Ability to exercise direction in organizations and departments.
2. Ability to design, develop and apply analytical methods (quantitative methods, statistical models and decision tools) for making strategic, tactical and operational decisions in organizations.
3. Ability to analyze, diagnose, design solutions and manage complex systems that integrate various resources of an organization keeping in mind the business environment.
4. Ability to apply theories and inherent principles of the organization in order to analyze complex and uncertainty situations, and make decisions using engineering tools.

TEACHING METHODOLOGY

The course is organized around three main activities:
Lectures
Practical sessions (exercises and laboratory sessions)
Autonomous work
In the lectures, teachers will introduce the theoretical basis of the subject, concepts, methods and results, given convenient examples to assist the understanding.
During the practical sessions, teachers will guide the students in the application of the theoretical concepts for problem solving, putting special emphasis on critical thinking skills. Exercises will be proposed to solve in and outside the classroom, in order to put the students in contact with the use of the tools developed in the lectures.
Students, on their own, should work on the supplied materials and the results of the practical sessions, in order to assimilate and fix the concepts. Professors will supply a syllabus of the course and guidelines to follow activities.

Observation: Although the materials are written in catalan, classroom sessions might be delivered in Spanish if it is needed.

LEARNING OBJECTIVES OF THE SUBJECT

This course provides the student with a series of advanced statistical tools which have been developed in the area of statistics in response to industrial and managerial problems.
It is a quantitative course, where statistical tools that can be used in decision taking on the basis of collected data. In particular, it covers forecasting computations from data series and some multivariate analysis methods.

STUDY LOAD

Type	Hours	Percentage
Self study	80,0	64.00
Hours large group	30,0	24.00
Hours small group	15,0	12.00

Total learning time: 125 h

CONTENTS

-M1 Module. Introduction to multivariate analysis

Description:
Multivariate statistics is quite a wide area within statistics, specialized in treating multidimensional observations. That is, it can be applied to populations where two or more characteristics are simultaneously measured. Because of its nature, its study requires certain matrix algebra skills and some basic statistics concepts, such as variance, covariance and correlations. This first introduction reviews these concepts.

Specific objectives:
Revise concepts that will be needed for following the course

Related activities:
1,2,3

Full-or-part-time: 8h 30m
Theory classes: 3h 30m
Self study : 5h

M2 Module. Principal Components Analysis

Description:
Principal components address the problems that arise when several variables which have been measured on a large set of individuals need to be analyzed, and the size of the database complicates an easy and efficient interpretation, whilst correlations among variables reduce the efficacy of other methods.
THe graphical representation of the variables and individuals on the space defined by the principal components allows to visualize the relationships among the variables, the similarities among individuals and the mutual assotiations.
Indeed, principal components can be seen as the projection of the individuals on a lower dimension space that minimizes the lost information.

Specific objectives:
Provide the student with the ability to reducer the volume of tables of quantitative data collected on a sample, without loosing much information

Related activities:
1, 2, 3 and 5

Full-or-part-time: 35h
Theory classes: 7h
Laboratory classes: 6h
Self study : 22h

M3 Module: Correspondence Analysis.

Description:
Correspondence analysis is focused on the study of two-entry tables where individuals are classified according to two different criteria. They yield a geometric representation on a low dimensional space, which is simple and precise, and allows to visualize association between cathegories of the same classification criterion, and among both of them.

Specific objectives:
At the end of the course the student whould be able to identify possible relationships between the cathegories used to classify individuals in a two-entry table.

Related activities:
1, 2, 3

Full-or-part-time: 16h 30m
Theory classes: 5h 30m
Laboratory classes: 1h
Self study : 10h

- S1 Module: Introduction to Time Series

Description:
A Time Series is a set of observations of the same magnitude ordered along the time. The objective of this module and the following ones is to analyze a series to extract its behavioral pattern, validate its goodness of fit, and forecast, as long as possible, its future evolution. A necessary tool to achieve this objective is linear regression, which will be revised in this introduction.

Specific objectives:
To present the concept of Time Series, and review some concepts that will be required in the forthcomming modules

Related activities:
1, 2, 4

Full-or-part-time: 5h
Theory classes: 2h
Self study : 3h

-S2 Module: Classical decomposition of Time Series

Description:
This module develops the so-called Classical Decomposition method, which decomposes the series in trend, seasonality, ciclic component, and residuals. To model the series and be able to do forecasts it is necessary to stabilize it by freeing it from the seasonal component, by means of moving averages. Once the conjuntion (aditive or multiplicative) of the components, a final model is obtained, which will allow to forecast future values.

Specific objectives:
Provide the student with the ability to identify additive and multiplicative series, and to model their trend and seasonality components, in order to build adequate models for each time series.

Related activities:
1, 2, 4

Full-or-part-time: 11h
Theory classes: 3h
Laboratory classes: 1h
Self study : 7h

S3 Module: Time series modelling using cathecorical variables

Description:
Modeling Time Series by means of cathecorical variables is a generalization of the classical decomposition method, that allows to model simultaneously the trend and the seasonality. Moreover, it dones not require deciding a priory wether the model should be additive or multiplicative. As opposite, a very general model, which includes both effects is proposed as a starting point, and it is the method itself who determines which of them are relevant in the series evolution. Thanks to this fact, the use of categorical variables in the process of time series modeling allows to overcome one of the principal limitations of classical decomposition, since it allows to build adequate models for series with a mixt behavior.

Specific objectives:
Provide the student the ability to model series with trend and seasonality, independently of their interaction.

Related activities:
1, 2, 4 and 6

Full-or-part-time: 15h
Theory classes: 2h
Laboratory classes: 3h
Self study : 10h

-S4 Module: Autocorrelation

Description:
This module presents the correlogram (graphical representation of the autocorrelation function). It can be used to confirm the seasonality of the series and its periodicity, as well as to determine the maximum number of allowable forecasts.

Specific objectives:
Give the student the necessary skills for correlogram elaboration and interpretation.

Related activities:
1, 2, 4 and 6

Full-or-part-time: 21h
Theory classes: 4h
Laboratory classes: 3h
Self study : 14h

-S5 Module: Other forecasting techniques

Description:
This module develops some methods based on exponential smoothing which can be applied, in particular, to model series that do not present a stable trend along the data collection period, or for which not much information is available.
It also contains brief overview of Box Jenkis methodology.

Specific objectives:
Overview other techniques that become useful to work with series that do not present a strong structure given by a trend and a seasonality

Related activities:
1, 2, 4.

Full-or-part-time: 13h
Theory classes: 3h
Laboratory classes: 1h
Self study : 9h

ACTIVITIES

Activity 1: Large group sessions/Theory and exercises

Description:
Attendance to the theory and practice sessions, and previous and subsequent training.

Specific objectives:
Transfer the necessary knowledge for a correct interpretation of the contents developed in the large group sessions.
Solving of exercises, and discussion on any doubts that might arise in relation to the course contents.
Generic competences training.

Material:
Documentation made available through Atenea
Course bibliography

Full-or-part-time: 67h
Theory classes: 27h
Self study: 40h

Activity 2: Practice sessions

Description:
Attendance to the practical sessions, and previous and subsequent preparation.
During these sessions the problem will guide the students through the solution of practical problems, encouraging the discussion among all attendants.

Material:
Notes and exercise sheets provided through Atenea

Full-or-part-time: 19h
Laboratory classes: 7h
Self study: 12h

Activity 3: Partial exam

Description:
Individual and writen exam concerning the contents of modules M1, M2 and M3

Specific objectives:
The exam should prove that the student has acquired and fixed the concepts, principles and basic foundings related with modules M1, M2 and M3.

Material:
Statement of the partial exam.
Students can use all their notes, and the material made avalilable trhough Atenea. Bibliography is also allowed.

Delivery:
The deliverable will be the exam sheed with marked answers.
It represents 40% of the final course mark

Full-or-part-time: 5h
Theory classes: 1h
Self study: 4h

Activity 4: Final exam

Description:
Individual and written test concerning the contents of modules S1, S2, S2, S4 and S5

Specific objectives:
The test must prove that the student has adquired and fixed the concepts, principles and basic fundaments related with modules S1, S2, S3, S4 and S5.

Material:
Statement of the final exam.
Students can use all their notes, and the material made avalilable trhough Atenea. Bibliography is also allowed.

Delivery:
The deliverable will be the exam sheed with marked answers.
It represents 40% of the final course mark

Full-or-part-time: 10h
Theory classes: 2h
Self study: 8h

Activity 5: Multivariate analysis project

Description:
Individual project on the practical application of the contents of module M2.
In this project the student will have to perform a principal components analysi of a data set provided by the professor, using a worksheet.

Specific objectives:
This project must prove that the student is able to apply correctly the tools developed in module M2, and use the obtained results to raise reasonable conclusions.

Material:
Course notes, project sketch and assigned dataset.
Worksheet prepared as a support for the development of the project.

Delivery:
The deliverable will be a report on the application of principal components analysis on the assigned data set, that the student must hand it within the prespecifided time window.
This task represents 10% of the final course mark

Full-or-part-time: 12h
Laboratory classes: 4h
Self study: 8h

(ENG) ACTIVITAT 6: PRÀCTICA DE SÈRIES TEMPORALS

Description:
Individual project.
Practical application of the contents of modules S3 and S4 to a particular dataset.

Specific objectives:
The report must prove that the student is able to study the data of his/her assigned dataset, model their behavior, give an interpretation of the obtained model, and compute the allowable forecasts.

Material:
Course notes, project sketch and individual dataset

Delivery:
The deliverable will be the report of the application of the necessary tools to the assigned dataset, and must be handed it withn the specified time window.
This task represents 10% of the final course mark

Full-or-part-time: 12h
Laboratory classes: 4h
Self study: 8h

GRADING SYSTEM

The final mark of the course depends on the following evaluation activities:
- Activity 3 (partial exam), weight: 40%
- Activity 4 (final exam), weight: 40%
- Activity 5, weight: 10%
- Activity 6, weight: 10%

Any student who cannot attend to the midterm exam (activity 3) or that wants to improve the obtained grade, will have the opportunity to improve that grade by taking an additional written exam that will take place the same day as the final exam (activity 4). The grade obtained in this test will range between 0 and 10, and will replace that of the midterm exam in case it is higher.

For those students who meet the requirements and submit to the reevaluation examination, the grade of the reevaluation exam will replace the grades of all the on-site written evaluation acts (tests, midterm and final exams) and the grades obtained during the course for lab practices, works, projects and presentations will be kept.
If the final grade after reevaluation is lower than 5.0, it will replace the initial one only if it is higher. If the final grade after reevaluation is greater or equal to 5.0, the final grade of the subject will be pass 5.0.

EXAMINATION RULES.

All evaluation activities (3, 4, 5 and 6) will be written, and performed individually.

BIBLIOGRAPHY

Basic:
- Polo Miranda, C. Estadística multivariable [on line]. 4ª ed. Barcelona: Edicions UPC, 2005 [Consultation: 08/01/2016]. Available on: http://hdl.handle.net/2099.3/36241. ISBN 8483018152.
- Albareda, M.; Algaba, I.; Pepió, M. Series temporales y previsiones. Barcelona: Omnia Science, 2013. ISBN 9788494062469.

Complementary:
- Everitt, Brian. An R and S-PLUS companion to multivariate analysis [on line]. London: Springer, cop. 2005 [Consultation: 10/06/2022]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/b138954. ISBN 1852338822.
- Makridakis, S.G.; Wheelwright, S.C.; Hyndman, R.J. Forecasting: methods and applications. 3rd ed. New York [etc.]: John Wiley & Sons, cop. 1998. ISBN 0471532339.

- Diebold, Francis X. Elementos de pronósticos. México [etc.]: International Thomson, cop. 1999. ISBN 9687529741.
- Johnson, R.A.; Wichern, D.W. Applied multivariate statistical analysis. 6th ed. Englewood Cliffs, N.J: Prentice-Hall, cop. 2007. ISBN 9780131877153.

RESOURCES

Other resources:
Course notes, exercises, videos and quizzes available in Atenea (in catalan)