

PROBABILITY AND STATISTICS

**Bachelor in Philosophy, Politics, Law and Economics PPLE
SEP-2023 PRST-PP.2.M.A**

Area Operations and Business Analytics

Number of sessions: 15

Academic year: 23-24

Degree course: SECOND

Number of credits: 3.0

Semester: 1^o

Category: COMPULSORY

Language: English

Professor: **BORJA MESA**

E-mail: bmesa@faculty.ie.edu

Borja Mesa-Sánchez is an expert in the area of industrial organization. After finishing his Ph.D. in Quantitative Economics (with honors) at Universidad de Alicante he joined the Department of Economics of Universidad Carlos III de Madrid as a Post-doctoral fellow. He received a Master Degree in Quantitative Economics, and his Bachelor's Degree from Universidad Carlos III de Madrid (with honors). His research is in competition policy, game theory and industrial organization. He has published in international scientific journals ranked in JCR by ISI. Borja teaches at undergraduate and graduate level and he has been the advisor for many senior thesis. He has worked in the department of training in Reuters, and he has won two literary awards.

Experience

- Visiting Professor Brown University
- Post-doctoral fellow, Universidad Carlos III de Madrid
- Adjunct Professor, Saint Louis University
- Teaching Assistant, Universidad de Alicante
- Training Assistant, Reuters

Education

- Post-doctoral fellow, Universidad Carlos III de Madrid
- Ph.D in Quantitative Economics (with honors), Universidad de Alicante
- Stays in European University Institute and University of York
- MSc in Quantitative Economics, Universidad de Alicante
- Bs Economics (with honors), Universidad Carlos III de Madrid

bmesa@faculty.ie.edu

SUBJECT DESCRIPTION

In 2020, the whole world was immobilized by the spread of a virus known as SARS-CoV-2: pain and shock prevailed everywhere. Never before people realized the importance of data, not any kind of kind; a reliable one. How many of us have experienced anguish and despair while listening to governments' daily reports on the number of people infected/affected/dead and cured? How many of us were startled at the wide divergences and sometimes inconsistencies in the numbers reported, especially the ones related to the prevalence and death rates? How many wondered why the Spanish and Italian death rate, close to ten per cent, is much higher than that of Germany? No doubt, there are many factors behind such a complex phenomenon, but the truth of the matter is that we will not be able to disentangle such factors until we have reliable data available.

What is reliable data? How important is it in representing a real-life situation? What do experts mean when they talk about representative samples? What is a sample and is there more than one way to take a sample? What insights can we draw from data? What are the numbers that matters and how can we use them to infer conclusions about our population? Answers to these questions and many more will be given in this introductory course on probability and statistics.

Statistics is the science of data, a discipline grounded in mathematics that converts raw data into actionable information. It uses mathematical tools to construct formal models, to summarize and process data and to reduce uncertainty in different environments: financial markets, the insurance industry, biomedicine, consumer behavior, presidential elections, gambling industry, physics, etc.

In this introductory course, we will focus on understanding some of the foundational concepts in statistics. In the first part of the course, students will learn how to collect and summarize data, how to describe data graphically and numerically, understand patterns of randomness that can affect business activities and relate them to known probability distributions. The second part of the course deals with inferential statistics, concepts like confidence intervals or hypothesis testing will be widely analyzed.

This course will be the basis for further subjects such as Econometrics or Social Analysis.

LEARNING OBJECTIVES

The objective of this course is to provide students with the tools to organize and understand data and to make use of this information in social sciences applications. At the end of the course you should be able to:

- Describe data by means of graphs or figures, understanding in which contexts each of these descriptive tools is useful.
- Understand patterns of randomness and relate them to known probability distributions.
- Understand the differences between population and sample distributions.
- Read the most common distribution tables.
- Derive confidence intervals for a parameter.
- Use confidence intervals to perform hypothesis testing.
- Use RStudio to perform statistical analysis.

Additionally, the course will focus on the acquisition or reinforcement of generic skills:

- The ability to summarize and present information in a meaningful way.

- The ability to build an abstract model to address social sciences problems.
- The ability to quickly identify the tools that are useful in social sciences.
- The ability to use Statistical software.

TEACHING METHODOLOGY

IE University teaching method is defined by its collaborative, active, and applied nature. Students actively participate in the whole process to build their knowledge and sharpen their skills. Professor's main role is to lead and guide students to achieve the learning objectives of the course. This is done by engaging in a diverse range of teaching techniques and different types of learning activities such as the following:

Learning Activity	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	33.33 %	25.0 hours
Discussions	6.67 %	5.0 hours
Exercises in class, Asynchronous sessions, Field Work	26.67 %	20.0 hours
Group work	13.33 %	10.0 hours
Individual studying	20.0 %	15.0 hours
TOTAL	100.0 %	75.0 hours

PROGRAM

MODULE 1. DESCRIPTIVE STATISTICS

SESSIONS 1 - 2 (LIVE IN-PERSON)

SESSIONS 1-2 [THEORY]: PRESENTATION, TOPIC 1, AND TOPIC 2

Topic 1: Introduction to Data

Topic 2: Summarizing Data

Fundamentals Elements of Statistics. Types of Data. Data Collection Principles. Summary Statistics: Mean, Median, Variance, Standard Deviation, Skewness, Kurtosis, Quartiles, Percentiles, Covariance, Correlation. Charts: Histograms, Scatter Plots, Boxplots. Chebyshev's Rule and the Empirical rule. The sample z-score. Other concepts

Activities: Lecture, practical examples, class discussion

Book Chapters: Chapters 1 and 2 (OIS) (mandatory, before the sessions) in OpenIntro Statistics (See Bibliography)

Book Chapters: Chapters 1 and 2 (MBS) in Statistics for Business and Economics (See Bibliography)

Book Chapters: Chapters 1 and 2 (NCT) in Statistics for Business and Economics (See Bibliography)

SESSION 3 (ASYNCHRONOUS)

SESSION 3 [LAB]: DESCRIPTIVE STATISTICS WITH R.

Using Technology: R Studio software

Activities: Descriptive Statistics with R Studio. Data sets manipulation. Macromagnitudes. GDP, inflation, unemployment. Okun's law. Phillips curve. Solving exercises to implement the concepts learnt in the theory sessions. Video lectures, practical cases, forum discussion, test.

Topics: Basic introduction to R

Install R and R studio

Descriptive Statistics with R Studio

SESSION 4 (LIVE IN-PERSON)

SESSION 4 [LAB]: DESCRIPTIVE STATISTICS WITH R.

Using Technology: R Studio software

Activities: Descriptive Statistics with R Studio. Data sets manipulation. Macromagnitudes. GDP, inflation, unemployment. Okun's law. Phillips curve. Solving exercises to implement the concepts learnt in the theory sessions. Video lectures, practical cases, forum discussion, test.

Topics: Basic introduction to R

Solving a Problem Set with R

MODULE 2. RANDOM VARIABLE AND PROBABILITY DISTRIBUTIONS

SESSIONS 5 - 7 (LIVE IN-PERSON)

SESSIONS 5-7 [THEORY]: TOPIC 3, TOPIC 4, and TOPIC 5

Topic 3: Probability

Topic 4: Discrete Random Variable. Discrete Probability Distributions

Topic 5: Continuous Random Variable. Continuous Probability Distributions

Basic concepts about Probability. Conditional Probability. Bayes' Theorem. Definition of random variables: Expected value, variance. Discrete random variables: probability functions and properties. Binomial, and Poisson distributions. Continuous random variables: probability functions and properties. Normal, and Exponential probability distributions

Activities: Lecture, practical examples, class discussion

Book Chapters: Chapter 3. Section 4.1, 4.2.1, 4.3, 4.5 (OIS) (mandatory, before the sessions) in OpenIntro Statistics (See Bibliography)

Book Chapters: Chapters 3 and 4 (MBS) in Statistics for Business and Economics (See Bibliography)

Book Chapters: Chapters 3, 4 and 5 (NCT) in Statistics for Business and Economics (See Bibliography)

SESSION 8 (LIVE IN-PERSON)

SESSION 8 [LAB]: PROBABILITY DISTRIBUTIONS WITH R.

Using Technology: R Studio software

Activities: Empirical Distributions with R Studio. Macromagnitudes. GDP, inflation, unemployment. Solving exercises to implement the concepts learnt in the theory sessions. Video lectures, practical cases, forum discussion, test.

Discrete Probability Distributions with R Studio

Continuous Probability Distributions with R Studio

MODULE 3. FOUNDATIONS FOR INFERENCE

SESSION 9 (LIVE IN-PERSON)

SESSION 9 [LAB]: PROBABILITY DISTRIBUTIONS WITH R.

Using Technology: R Studio software

Activities: Empirical Distributions with R Studio. Macromagnitudes. GDP, inflation, unemployment. Okun's law. Phillips curve. Solving exercises to implement the concepts learnt in the theory sessions. Video lectures, practical cases, forum discussion, test.

Topics: Probability Distributions

Solving a Problem Set with R

SESSIONS 10 - 11 (LIVE IN-PERSON)

SESSIONS 10-11 [THEORY]: TOPIC 6 and TOPIC 7

Topic 6: Sampling Distributions and the Central Limit Theorem

Topic 7: Confidence Intervals

Point estimates and sampling variability. Brief Introduction to Sampling Methods. Random Samples. Distribution of Sample Means. The Central Limit Theorem. Sampling Distributions of Sample Proportions. Confidence intervals for the mean and the proportion. Hypothesis testing: main concepts. Applying confidence intervals to hypothesis testing

Activities: Lecture, practical examples, class discussion

Book Chapters: Chapter 5 (section 5.3.4 is not included), sections 6.1.1, 6.1.2, 6.1.4, 6.1.5, and section 7.1 (section 7.1.5 is not included) (OIS) (mandatory, before the sessions) in OpenIntro Statistics (See Bibliography)

Book Chapters: Chapters 5 and 6 (MBS) in Statistics for Business and Economics (See Bibliography)

Book Chapters: Chapters 6 and 7 (NCT) in Statistics for Business and Economics (See Bibliography)

SESSION 12 (LIVE IN-PERSON)

SESSION 12 [LAB]: SAMPLING DISTRIBUTIONS WITH R STUDIO.

Using Technology: R Studio software

Activities: Sampling Distributions and Confidence intervals with macromagnitudes. GDP, inflation, unemployment. Okun's law. Phillips curve. Solving exercises to implement the concepts learnt in the theory sessions. Video lectures, practical cases, forum discussion, test.

Topics: Sampling Distributions. Confidence Intervals.

SESSION 13 (LIVE IN-PERSON)

SESSION 13 [LAB]: SAMPLING DISTRIBUTIONS WITH R STUDIO.

Using Technology: R Studio software

Activities: Sampling Distributions and Confidence intervals with macromagnitudes. GDP, inflation, unemployment. Okun's law. Phillips curve. Solving exercises to implement the concepts learnt in the theory sessions. Video lectures, practical cases, forum discussion, test.

Topics: Sampling Distributions. Confidence Intervals

Solving a Problem Set with R

SESSION 14 (ASYNCHRONOUS)

SESSION 14 [LAB]: CASE STUDY WORKGROUP

Using Technology: R Studio software

Activities: Students are expected to solve a case study using R Studio and applying the concepts learnt throughout the course. Students will work in groups

Topics: all topics covered in the course

SESSION 15 (LIVE IN-PERSON)

SESSION 15: FINAL EXAM

The final exam is scheduled on the last session of the course. The final exam will be computer-based. Therefore, you must bring your own computer and ensure that you will be able to connect to the Internet. The exam is open book. The final exam covers all the topics seen in class. The final exam will include material from the PowerPoint slides, the book, and the problem sets. It is highly recommendable to delve deeply into the topics using the book. No questions are allowed during the exam.

In order to pass the course, you need a minimum grade of 3.5 in the final exam. If your grade in the final exam does not reach the threshold value of 3.5, you will fail the course, even in the case in which your weighted average (computed using the table above) exceeds 5.0.

EVALUATION CRITERIA

Ordinary evaluation

Your final grade in the course will be based on a combination of different criteria that are described in the following table:

Grading System:

criteria	percentage	Learning Objectives	Comments
Class Participation	15 %		See comments below on class participation
Workgroups	20 %		Case study to be solved in groups
Individual Work	30 %		Problem Sets
Final Exam	35 %		Omnicomprehensive final exam

RE-SIT / RE-TAKE POLICY

Class participation (15%)

Three main criteria will be used in reaching a judgment about your class participation:

Participation in class will be evaluated positively if students: (1) attain a threshold quantity of contributions that is sufficient for making a reliable assessment of comment quality. Additionally, (2) participation will be evaluated in quality terms. A high-quality comment reveals a depth of insight, rigorous use of case evidence, consistency of argument, and realism. A high-quality presentation of ideas must consider the relevance and timing of comments and the flow and content of the ensuing class discussion. It demands comments that are concise and clear, and that are conveyed with a spirit of involvement in the discussion at hand. Finally, (3) general attitude and behavior in class will also be considered. Students affecting the class environment in a negative way will lose points in the assistance grade.

Case study (Workgroups) (20%)

The elements for effective teamwork can help you to achieve success in all areas of life. It is crucial that you learn how to collaborate with other students even when you have not chosen the students whom to work with. This exam mainly consists in solving and discussing some questions about a case study in Statistics with real data using the statistical software R Studio. This exam is in groups (4-5 students per group). More details about this exam will be provided in advance in class.

Labs (Problem Sets) (30%)

Students will solve 3 problem sets throughout the course, one problem set per module. Answers will be submitted through BlackBoard.

Final Exam (35%)

The final exam is scheduled on the last session of the course. The final exam will be computer-based. Therefore, you must bring your own computer and ensure that you will be able to connect to the Internet. The exam is open book. The final exam covers all the topics seen in class. The final exam will include material from the PowerPoint slides, the book, and the problem sets. It is highly recommendable to delve deeply into the topics using the book. No questions are allowed during the exam.

In order to pass the course, you need a minimum grade of 3.5 in the final exam. If your grade in the final exam does not reach the threshold value of 3.5, you will fail the course, even in the case in which your weighted average (computed using the table above) exceeds 5.0.

Retake exam

Each student has four chances to pass any given course distributed in two consecutive academic years (regular period and July period). It is mandatory to attend 100% of the classes. Students who do not comply with at least 70% attendance will lose their 1st and 2nd chance, and go directly to the 3rd one (they will need to enroll again in this course the next academic year). Grading for retakes will be subject to the following rules:

- Those students who failed the subject in the first regular period will have to do a retake in July (except those not complying with attendance rules who are banned from this possibility).
- Dates and location of the July retakes will be posted in advance and will not be changed. Please take this into consideration when planning your summer.
- The maximum grade that a student may obtain in the 2nd exam session is 8 out of 10. Those students in the 3rd call will be required to attend 50% of the classes. If due to schedule overlap, a different option will be discussed with the professor in order to pass the subject.

BIBLIOGRAPHY

Compulsory

- Diez, David; Barr Christopher; Cetinkaya-Rundel, Mine. *OpenIntro Statistics*. Any Edition (4th Edition). OpenIntro. ISBN 1943450072 (Digital)

Available for free online

<https://www.openintro.org/book/os/>

Recommended

- McClave, James T, Paul George Benson, and Terry Sincich. (2022). *Statistics for Business and Economics*. 14 Edition, Global ed. Pearson. ISBN 9781292413396 (Digital)

This reference is included for two reasons. The first is to give the student the possibility to study the concepts using another reference. The second reason is that some aspects covered in the sessions are not included in the main reference.

- Newbold, Paul. Carlson, William L., Thorne, Betty M. (2013). *Statistics for Business and Economics*. 8th edition. Pearson Education Limited. ISBN 9780132745659 (Digital)

This reference is included for two reasons. The first is to give the student the possibility to study the concepts using another reference. The second reason is that some aspects covered in the sessions are not included in the main reference.

- James, G., Witten, D., Hastie, T., Tibshirani, R. (2017). *An Introduction to Statistical Learning with Applications in R*. 8th edition. Springer-Verlag New York. ISBN 9781461471 (Digital)

BEHAVIOR RULES

Please, check the University's Code of Conduct [here](#). The Program Director may provide further indications.

ATTENDANCE POLICY

Please, check the University's Attendance Policy [here](#). The Program Director may provide further indications.

ETHICAL POLICY

Please, check the University's Ethics Code [here](#). The Program Director may provide further indications.