FUNDAMENTALS OF PROBABILITY & STATISTICS

IE University
Professor: **CRISTINA GARCÍA DE LA FUENTE**

E-mail: cgdelafuente@faculty.ie.edu

Academic year: 23-24
Degree course: FIRST
Semester: 1º
Category: BASIC
Number of credits: 6.0
Language: English

PREREQUISITES
Mathematics for Economists, Programming for Economists I

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 of any 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 matter 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, Data Science for Economists or Econometrics II: Identification & Time Series.

OBJECTIVES AND SKILLS

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.
- Make inference by understanding the concept of null and alternative hypotheses
- Use statistical methods for decision-making in social sciences.

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.

METHODOLOGY

Sessions are divided in theory sessions (lecture-based) and practice sessions (activity-based):

Lectures: The main resources in these classes will be projections of slides, providing files and electronic information and the use of the Internet as a support tool, work, and communication. The student must supplement these lectures with recommended readings by the professor. These will be available to the students via Campus Online.

Practical: During these sessions, the students will work on exercises, or a project; which are tightly connected with the concepts explained during the theory sessions. Most of the problems and projects will require computations, so students should be familiarized with software R or R-Studio. Students must prepare and solve the problem sets assigned prior to the beginning of the practice class by using the theory lectures and the recommended bibliography. Students are expected to participate actively in class, expressing their difficulties and proposing solutions. The professor will always help the students by clarifying the problem sets solutions; assuming the students have previously tried to work on their own. Otherwise, this will be counter-productive.

Teaching methodology	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	30.0 %	45 hours
Discussions	3.33 %	5 hours
Exercises	20.0 %	30 hours
Group work	26.67 %	40 hours
Other individual studying	20.0 %	30 hours
TOTAL	100.0 %	150 hours

PROGRAM

MODULE 1. DESCRIPTIVE STATISTICS

Sessions 1 to 6

SESSION 1 (LIVE IN-PERSON)

[THEORY]: PRESENTATION & TOPIC 1

Topic 1: Introduction to Data
Data Basics. Sampling Principles. Experiments.

Statistics for Business and Economics [NCT]: Chapter 17

SESSION 2 (LIVE IN-PERSON)

[THEORY]: TOPIC 2

Topic 2: Summarizing Data

Summary Statistics: Mean, Median, Variance, Standard Deviation, Skewness, Quartiles, Percentiles, Covariance, Correlation, Contingency Tables.

Charts: Histograms, Scatter Plots, Boxplots.

Statistics for Business and Economics [NCT]: Chapters 2 and 3

SESSION 3 (LIVE IN-PERSON)

[THEORY]: TOPIC 2

Topic 2: Summarizing Data

Summary Statistics: Mean, Median, Variance, Standard Deviation, Skewness, Quartiles, Percentiles, Covariance, Correlation, Contingency Tables.

Charts: Histograms, Scatter Plots, Boxplots.

Statistics for Business and Economics [NCT]: Chapters 2 and 3

SESSION 4 (LIVE IN-PERSON)

[LAB]: DESCRIPTIVE STATISTICS WITH R

Using Technology: R software

- Descriptive Statistics with R: sum(), mean(), sd(), cor.test(), fivenum(), hist(), boxplot(), plot().

SESSION 5 (LIVE IN-PERSON)

[THEORY]: TOPIC 2

Topic 2: Summarizing Data

Summary Statistics: Mean, Median, Variance, Standard Deviation, Skewness, Quartiles, Percentiles, Covariance, Correlation, Contingency Tables.

Charts: Histograms, Scatter Plots, Boxplots.

Statistics for Business and Economics [NCT]: Chapters 2 and 3

SESSION 6 (LIVE IN-PERSON)

[PRACTICE]: PROBLEM SET

Solving exercises to implement the concepts learnt in the theory sessions.

MODULE 2. DISTRIBUTIONS OF RANDOM VARIABLES

Sessions 7 to 10

SESSION 7 (LIVE IN-PERSON)

[THEORY]: TOPIC 3

Topic 3: Discrete Probability Distributions
Definition of random variables: Expected value, variance. Discrete random variables: probability functions and properties. Binomial, and Poisson.

Statistics for Business and Economics [NCT]: Chapter 4

SESSION 8 (LIVE IN-PERSON)

[LAB]: PROBABILITY DISTRIBUTIONS WITH R

Using Technology: R software

   - Probability Distributions with R: dbinom(), pbinom(), dpois(), ppois(), punif(), qunif(), runif(), pnorm(), qnorm(), rnorm(), pexp(), qexp(), pt(), qt().

SESSION 9 (LIVE IN-PERSON)

[THEORY]: TOPIC 4

Topic 4: Continuous Probability Distributions

Continuous random variables: probability functions and properties. Uniform, Normal, Exponential.

Statistics for Business and Economics [NCT]: Chapter 5

SESSION 10 (LIVE IN-PERSON)

[PRACTICE]: PROBLEM SET

Solving exercises to implement the concepts learnt in the theory sessions.

SESSION 11 (LIVE IN-PERSON)

[PRACTICE]: Review session.

SESSION 12 (LIVE IN-PERSON)

QUIZ 1

Quiz questions will be quite similar to those in problem sets. Answers will be submitted through BlackBoard. The exam is open book.

MODULE 3. SAMPLING DISTRIBUTIONS. CONFIDENCE INTERVALS

Sessions 11 to 15

SESSION 13 (LIVE IN-PERSON)

[THEORY]: TOPIC 5

Topic 5: Sampling Distributions. The Central Limit Theorem

Brief Introduction to Sampling Methods. Random Samples. Distribution of Sample Means. The Central Limit Theorem. Sampling Distributions of Sample Proportions.

Statistics for Business and Economics [NCT]: Chapter 6

SESSION 14 (LIVE IN-PERSON)

[LAB]: SAMPLING DISTRIBUTIONS AND CONFIDENCE INTERVALS WITH R

Using Technology: R software

SESSION 15 (LIVE IN-PERSON)

[THEORY]: TOPIC 6

Topic 6: Inferences Based on a Single Sample: Estimation with Confidence Intervals.

Point estimation of parameters. Confidence interval for a population mean (large samples, normal distribution). Confidence interval for a population mean (small samples, t distribution). Large sample confidence interval for a population proportion. Determining the sample size.

Statistics for Business and Economics [NCT]: Chapter 7

MODULE 4. HYPOTHESIS TESTING

Sessions 16 to 24

SESSION 16 (LIVE IN-PERSON)

[THEORY]: TOPIC 7

Topic 7: Inferences Based on a Single Sample: Hypothesis Testing.

Hypothesis Testing One Sample Means. Large samples (Variance Known). Small Samples (Variance Unknown): The t distribution. Hypothesis Testing One Sample Proportions.

Statistics for Business and Economics [NCT]: Chapter 9

SESSION 17 (LIVE IN-PERSON)

[THEORY]: TOPIC 7

Topic 7: Inferences Based on a Single Sample: Hypothesis Testing.

Hypothesis Testing One Sample Means. Large samples (Variance Known). Small Samples (Variance Unknown): The t distribution. Hypothesis Testing One Sample Proportions.

Statistics for Business and Economics [NCT]: Chapter 9

SESSION 18 (LIVE IN-PERSON)

[LAB]: INFERENTIAL STATISTICS WITH R

Using Technology: R software

- t.test(), prop.test(), power.t.test()

SESSION 19 (LIVE IN-PERSON)

[THEORY]: TOPIC 7

Topic 7: Inferences Based on a Single Sample: Hypothesis Testing.

Hypothesis Testing One Sample Means. Large samples (Variance Known). Small Samples (Variance Unknown): The t distribution. Hypothesis Testing One Sample Proportions.

Statistics for Business and Economics [NCT]: Chapter 9

SESSION 20 (LIVE IN-PERSON)

[PRACTICE]: PROBLEM SET

Solving exercises to implement the concepts learnt in the theory sessions.

SESSION 21 (LIVE IN-PERSON)

[PRACTICE]: Review sesson.

SESSION 22 (LIVE IN-PERSON)

QUIZ 2

Quiz questions will be quite similar to those in problem sets. Answers will be submitted through BlackBoard. The exam is open book.

SESSION 23 (LIVE IN-PERSON)

[THEORY]: TOPIC 8

Topic 8: Inferences Based on Two Samples

Two-Paired Sample Tests. Two-Independent Sample Tests Difference of Population Means.Difference of Population Proportions

Statistics for Business and Economics [NCT]: Chapter 10

SESSION 24 (LIVE IN-PERSON)

[THEORY]: TOPIC 8

Topic 8: Inferences Based on Two Samples

Two-Paired Sample Tests. Two-Independent Sample Tests Difference of Population Means.Difference of Population Proportions

Statistics for Business and Economics [NCT]: Chapter 10

SESSION 25 (LIVE IN-PERSON)

[LAB]: INFERENTIAL STATISTICS WITH R

Using Technology: R software

- t.test(), prop.test()

SESSION 26 (LIVE IN-PERSON)

[THEORY]: TOPIC 8

Topic 8: Inferences Based on Two Samples

Two-Paired Sample Tests. Two-Independent Sample Tests Difference of Population Means.Difference of Population Proportions.

Statistics for Business and Economics [NCT]: Chapter 10

MODULE 5. LINEAR REGRESSION

Sessions 26 to 29

SESSION 27 (LIVE IN-PERSON)

[THEORY]: TOPIC 9

Topic 9: Linear Regression.

Simple linear regression.

Statistics for Business and Economics [NCT]: Chapter 11

SESSION 28 (LIVE IN-PERSON)

[PRACTICE]: PROBLEM SET

Solving exercises to implement the concepts learnt in the theory sessions.

Statistics for Business and Economics [NCT]: Chapter 11

SESSION 29 (LIVE IN-PERSON)

[PRACTICE]: Review session.

SESSION 30 (LIVE IN-PERSON)

FINAL EXAM

There will be one final exam, 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 and use the resources needed to perform statistical analysis as seen in class. The exam is open book.

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.

BIBLIOGRAPHY

Compulsory

- Newbold, Paul. Carlson, William L., Thorne, Betty M.. (2018). *Statistics for Business and Economics.* 13th edition. Pearson Education Limited. ISBN 9781292227085 (Printed)

Recommended

- Diez, David; Barr Christopher; Cetinkaya-Rundel, Mine. (2019). *OpenIntro Statistics.* 4th Edition. OpenIntro, Inc. ISBN 1943450072 (Digital)

  This book is available in an online version only and can be acquired via: https://www.openintro.org/book/os/

- 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 (Printed)

- Mcclave, J.T, Benson, P.G., & Sincich, T.. (2021). *Statistics for Business and Economics.* global edition. Pearson Education Limited. ISBN 9781292413396 (Printed)

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	Comments
Class Participation	10 %	See comments below on class participation
Workgroups	20 %	Group project (see separate document)
Intermediate Tests	20 %	Quizzes
Individual Work	15 %	Problem Sets
Final Exam	35 %	Omnicomprehensive final exam

Class participation (10%)

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

1- Attendance: Attendance to class is compulsory. (1) Students must comply with the 70% attendance rule. Otherwise, they will lose their 1st and 2nd chance and go directly to the 3rd one (they will need to enrol again in this course next academic year). (2) Punctuality will be taken into consideration when grading this assistance item and the teacher reserves the right to allow attendance to class to those students not being on time. 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.

2- Active 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.

Group project (15%) (to be discussed in class)

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. Each group will be composed of 4-5 students and must prepare a project due at the end of the course (more details about the final deadline will be periodically provided during the course.) The group project will consist in the identification of a real-world problem, taken from social sciences or any other field of interest, the collection of relevant data, the statistical analysis of the data, the development of a statistical model, and the final interpretation of the obtained results. Submissions will be delivered using Turnitin, following the appropriate link provided on campus online. No work will be accepted if submitted otherwise. At the beginning of the course, the professor will upload a specific document called project instructions where a step-by-step description of what the students need to do.

Quizzes (25%)

There will be two quizzes. Quizzes contain theory and practice questions. Students will be requested to submit the answers through BlackBoard. Quizzes will include material from the slides, the required textbook, and the problem sets. It is highly recommendable to delve deeply into the topics using the book. Since quizzes are computer-based, students must bring their own computer and ensure that they will be able to connect to the Internet and use the resources needed to perform statistical analysis as seen in class. The exam is open book.

Problem Sets (15%)

Students will solve several problem sets throughout the course. Answers will be submitted through BlackBoard. The problem set with the worst grade will not be considered for calculating the final grade.

Final Exam (35%)

There will be one final exam, 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 and use the resources needed to perform statistical analysis as seen in class. The exam is open book.

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.

PROFESSOR BIO

Professor: **CRISTINA GARCÍA DE LA FUENTE**

E-mail: cgdelafuente@faculty.ie.edu

Cristina de la Fuente, PhD.

I studied Applied Mathematics in Mexico City, at my beloved alma mater, ITAM. When I finished my bachelor's degree, I came to Spain to study a master's and a PhD, both in Mathematical Engineering at the Statistics Department of the Universidad Carlos III de Madrid.

I have always been a teacher. First, I tutored young kids during my bachelor's degree. Then, I started teaching at the UC3M while I was a student there. Since I graduated, I have been a Math and Statistics professor in public and private, Spanish and international universities. Teaching really is my calling in life.

OTHER INFORMATION

If you have a question(s) that was not answered in class, you are welcome to ask your question(s) via email. Although I will make every effort to respond to your question(s) as quickly and thoroughly as possible, please recognize that I may not be available when you send an email. If your question cannot be properly answered via email and/or you would prefer to meet in person, please make an appointment to meet with me via Zoom.

E-mail: cgdelafuente@faculty.ie.edu

Office Hours: Upon appointment

ATTENDANCE

Attendance at all scheduled classes is mandatory and essential for success in the course. If you miss class for any reason, you are responsible for getting notes from classmates. If you have questions about any assignment please send me an email. Under most circumstances, students who miss a class in which a presentation, quiz, or final exam is held will not be granted an exception or given an opportunity to do a make-up assignment or exam. However, if illness or other circumstances prevent you from adhering to the assignment/presentation due dates stated in this syllabus, please inform your Program Management as soon as possible and provide, if necessary, official documentation (e.g., from a medical doctor, counsellor). They will inform the professor who will decide if an exception may be granted.

Student Privacy Statement

At times, students may disclose personal information through class discussions. It is expected that all members of the class will respect the privacy of their classmates. This means that the information that is closed in the class will not be repeated or discussed with other students outside of the course.

Decisions about Grades

Decisions about grades are made very carefully, and are final at the end of the course. If you have questions regarding a certain grade or you would like to receive personal feedback, you must request a meeting with me to discuss grades on specific assignments before the last class of the course. Any disputes regarding grades must be resolved before the final exam. "Extra credit" or make-up assignments will only be allowed under extenuating circumstances at the sole discretion of the course professor.

ACADEMIC INTEGRITY

Unless you are specifically instructed to work with other students in a group, all of your assignments, papers, projects, presentations, and any work must reflect your own work and thinking. What is academic integrity? When you do the right thing even though no one is watching. The core values of integrity, both academic and otherwise include: honesty, fairness, respect, responsibility, and trust. Academic integrity requires that all students within Instituto de Empresa (IE) act in accordance with these values in the conduct of their academic work, and that they follow the rules and regulations concerning the accepted conduct, practices and procedures of academic research and writing. Academic integrity violations are defined as Cheating, Plagiarism or other violations of academic ethics.

Cheating and plagiarism are very serious offenses governed by the IE student code of conduct. Any student found cheating or plagiarizing on any assignment or component of this course will at a minimum receive a "0" on the affected assignment. Moreover, the student will also be referred to the University Judicial System for further action. Additional penalties could include a note on your transcript, failing the class, or expulsion from the university. It is important to note that, while the list below is comprehensive, it should not be considered exhaustive.

Cheating includes:

- An act or attempt to give, receive, share, or utilize unauthorized information or unauthorized assistance at any time for assignments, papers, projects, presentations, tests or examinations. Students are permitted to mentor and/or assist other students with assignments by providing insight and/or advice. However, students must not allow other students to copy their work, nor will students be permitted to copy the work of other students. Students must acknowledge when they have received assistance from others.
- Failure to follow rules on assignments, papers, projects, presentations, tests or examinations as provided by the course professor and/or as stipulated by IE.
- Unauthorized co-operation or collaboration.
- Tampering with official documents, including electronic records.
- The impersonation of a student on presentations, exercises, tests or an examination. This

includes logging onto any electronic course management tool or program (e.g. BlackBoard) using someone else's login and password.

Plagiarism includes:

- Using the work of others and attempting to present it as your own. For example, using phrases or passages from books, articles, newspapers, or the internet and not referencing them properly in your document. This includes using information from others without citing it, misrepresentation of cited work, and misuse of quotation marks.
- Submitting an assignment or paper that is highly similar to what someone else has written (i.e., minimal changes in wording, or where the sentences are similar, but in a different order).
- You don't have to commit "word for word" copying to plagiarize – you can also plagiarize if you turn in something that is "thought for thought" the same as someone else.

Other violations of academic ethics include:

- Not acknowledging that your work or any part thereof has been submitted for credit elsewhere.
- Misleading or false statements regarding work completed.
- Knowingly aiding or abetting anyone in committing any form of an academic integrity violation.

CODE OF CONDUCT IN CLASS

1. Be on time: Students arriving more than 5 minutes late will be marked as "Absent".

Only students that notify in advance in writing that they will be late for a specific session might be granted an exception (to the discretion of the professor).

2. Respect your classmates. Classroom discussion is an important part of the learning process. Therefore, it is vital to maintain a classroom environment that is respectful and free of discrimination and/or recrimination from peers. Please keep in mind that at times, students may disclose personal information through class discussions. It is expected that all members of the class will respect the privacy of their classmates. However, please remember that class is NOT a protected, confidential environment, and the professor cannot guarantee that other students/peers will maintain your confidential information should you choose to share it.

3. If applicable, bring your name card and strictly follow the seating chart. It helps faculty members and fellow students learn your names.

4. Don't leave the room during the lecture: Students are not allowed to leave the room during the lecture. If a student leaves the room during the lecture, he/she will not be allowed to re-enter and, therefore, will be marked as "Absent".

Only students that notify that they have a special reason to leave the session early will be granted an exception (to the discretion of the professor).

5. Do not engage in side conversation. As a sign of respect toward the person presenting the lecture (the teacher as well as a fellow student), side conversations are not allowed. If you have a question, raise your hand and ask it. If you do not want to ask it during the lecture, feel free to approach your teacher after class.

If a student is disrupting the flow of the lecture, he/she will be asked to leave the classroom and, consequently, will be marked as "Absent".

6. Use your laptop for course-related purposes only. The use of laptops during the lectures should be authorized by the professor. The use of Facebook, Twitter, or accessing any type of content not related to the lecture is penalized. The student will be asked to leave the room and, consequently, will be marked as "Absent".

7. No cellular phones: IE University implements a "Phone-free Classroom" policy and, therefore, the use of phones, tablets, etc. is forbidden inside your classroom. Failing to abide by this rule entails expulsion from the room and will be counted as one Absence.

8. Escalation policy: 1/3/5. Items 4, 5, and 6 above entail expulsion from the classroom and the consequent marking of the student as "Absent." IE University implements an "escalation policy": The first time a student is asked to leave the room for disciplinary reasons (as per items 4, 5, and 6 above), the student will incur one absence, the second time it will count as three absences, and from the third time onward, any expulsion from the classroom due to disciplinary issue will entail 5 absences.