Course Unit File

Statistical Analysis

Level:	**Second cycle**	**Semester:**	**Fall**
Curricular Unit Code:	**200185**	**Duration:**	**S1**
Year:	**1**	**Academic year:**	**2021-22**

Total of hours

Total of Hours Spent	Contact Hours	Nr. of hours per week	ECTS
210	; 30 TP	2	7.5

T - Theoretical; TP - Theoretical and practical ; PL - Practical and Laboratorial; TC - Field Work; S - Seminar; E - Internship; OT - Tutorial Orietation; O - Other

Responsible teaching staff

Ana Cristina Marinho da Costa

Other teaching staff

-

Office hours

Friday, 17h00 - 18h00 (please confirm meeting by email or at the end of a class)

General objectives

This curricular unit aims at supplying to the students the theoretical and practical knowledge about methodologies on parametric and nonparametric statistical inference. Students will explore the core principles of statistics, from both the conceptual and applied perspectives. The students will acquire competences related to random variables, estimators, sampling distributions, point and interval estimation and hypothesis testing. Additionally, some issues of asymptotic distributions are addressed. Moreover, the analysis of variance is introduced, as well as several nonparametric statistical tests. The students will clearly understand the conditions of applicability of each procedure. The concepts and principles will be illustrated using real-world concepts applicable to many industries, including medical, business, sports, insurance, etc.

Enrolment Requirements

Teaching language: English.

Learning outcomes of the curricular unit (LO)

Upon completion of this course, students should be able to:
1. Describe a random variable; understand probabilistic models and probabilities calculation
2. Characterise the Binomial and Normal distributions; compute probabilities
3. Understand the concept of sampling distribution, explain and apply the Central Limit Theorem
4. Explain the impact of the sample size on the sampling distribution
5. Understand and investigate the properties of estimators
6. Build and interpret confidence intervals
7. Calculate the sample size given the precision of the point estimate
8. Formulate the hypothesis of the statistical test and decide based on an appropriate test
9. Explain the two types of error in statistical tests
10. Calculate and interpret the p-value
11. Make inferences about the mean, proportion, difference between means and proportions, variance and ratio of variances
12. Verify the assumptions of ANOVA, obtain the ANOVA table, apply the F-test and post-hoc tests
13. Distinguish parametric & nonparametric test procedures
14. Perform nonparametric hypothesis tests

Syllabus

The curricular unit is organized in 8 Learning Units (LU):
1. RANDOM VARIABLES
- Probabilistic models
- Discrete r.v.
- Continuous r.v.
2. PROBABILITY DISTRIBUTIONS
- Binomial, Poisson, Normal
- Approximation of Binomial to Normal
- t, Chi-square, F
3. SAMPLING DISTRIBUTIONS
- Sampling statistics & sampling distributions
- Distribution of sampling mean and sampling proportion
4. POINT ESTIMATION
- Unbiasedness, efficiency, consistency
5. INTERVAL ESTIMATION
- CI for mean, proportion, variance
- CI for difference between means and between proportions
- Sample size determination
6. HYPOTHESIS TESTING
- Concepts
- Tests for mean,proportion, variance, difference between means and between proportions, ratio between variances
- Tests for correlation coefficient
7. ANALYSIS OF VARIANCE
- One-way ANOVA with fixed effects
- Multiple comparison tests
- Tests to the equality of k variances
8. NONPARAMETRIC TESTS
- Introduction
- Distribution fitting tests
- Comparing independent and paired-samples
- Spearman's rank correlation test

Curricular unit Planning

WEEK	PLANNED CONTENT
1	Information about the curricular unit including objectives, contents, bibliography and evaluation methods. LU1: Random variables.
2	LU1: Random variables. LU2: Probability distributions: Binomial and Poisson.
3	LU2: Probability distributions: Normal, Chi-Square, Students' t and F.
4	LU3: Sampling statistics and sampling distributions. Sampling distribution of the mean; Central Limit Theorem; sampling distribution of the proportion.
5	LU4: Point estimation: properties of estimators.
6	LU5: Confidence interval estimation.
7	LU5: Confidence interval estimation: sample size determination. LU6: Hypothesis testing: notation and concepts.
	NO CLASS (break for trimester exams)
8	Clarification of doubts.
	TEST (LU1 - LU5) Saturday (schedule to be defined)
9	LU6: Hypothesis testing for the mean, for the difference between means, for the proportion, and for the difference between proportions.
10	LU6: Hypothesis testing for the variance, for the ratio between variances, and for the correlation coefficient.
11	LU7: Analysis of variance.
12	LU7: Analysis of variance. Practical examples in Excel.
13	LU8: Nonparametric testing: introduction; distribution fitting tests.

14	LU8: Nonparametric testing: comparing independent samples; comparing paired samples; Spearman's rank correlation test.
28 December	**Assignment Report must be uploaded to the Moodle platform (one PDF file)**
January	**EXAM (about all topics, but more focused on LU6 - LU8)** (see the masters' exams calendar)

Demonstration of the syllabus coherence with the curricular unit's learning objectives

The learning units (LU) cover the learning outcomes (LO) as follows:
- LO 1 is addressed in the LU1;
- LO 2 is addressed in the LU2;
- LO 3 and 4 are addressed in LU3;
- LO 5 is addressed in the LU4;
- LO 6, 7 and 11are addressed in LU5;
- LO 8, 9, 10 and 11 are addressed in LU6;
- LO 12 is addressed in LU7;
- LO 11, 13 and 14 are addressed in LU8.

Teaching methodologies

The curricular unit is based on theoretical and practical lessons. A variety of instructional strategies will be applied, including lectures, slide show demonstrations, step-by-step applications (with and without software), questions and answers. The sessions include presentation of concepts and methodologies, solving examples, discussion and interpretation of results. The practical component is geared towards solving problems and exercises, including discussion and interpretation of results. A set of exercises to be completed independently in extra-classroom context is also proposed.

Evaluation

REGULAR PERIOD (1st call): test (30%; about LU1-LU5), exam (50%; about all topics, but more focused on LU6-LU8), Assignment Report (20%; mandatory for approval).

RESIT PERIOD (2nd call): final exam (100%).

RULES:
To complete the test/exam, students must provide themselves with the form and statistical tables disclosed in Moodle, and also with a scientific calculating machine. Graphic calculating machines are not allowed.

The Assignment Report is mandatory for approval in the regular examination period. The report must be prepared individually, in Portuguese or English, using a set of artificial data. Detailed information and the data set of the assignment are disclosed in Moodle. Reports submitted after the deadline will have a penalty of 0.5 points for each day of delay. The maximum delay allowed is 3 days. Reports not submitted in the Moodle platform will be rejected.

Demonstration of the coherence between the teaching methodologies and the learning outcomes

The presentation of theoretical concepts and methodologies, followed by application exercises will provide students with the knowledge, skills and abilities listed as learning objectives at the beginning of the semester. For each learning unit (LU), detailed learning objectives and recommended readings on different resources are listed. The assignment is supervised by the teacher and is subject to a report. Students can use any software to perform the data analysis (e.g., Excel, Xlstat, R, SAS, JMP, SPSS, etc.). The assignments are stimulating and relevant in understanding the topics of study. Learning objectives associated with the last three units are evaluated in the assignment. The test and final exam require students to integrate the entire contents of the course and prove their mastery of it.

Bibliography

Newbold, P., Carlson, W. L., Thorne, B. (2013). *Statistics for Business and Economics* . 8th edition, Boston: Pearson, https://ebookcentral.proquest.com/lib/novaims/detail.action?docID=5174169 (full text available after login in NOVA IMS network or through VPN connection).
Conover, W. J. (1999). *Practical Nonparametric Statistics* . 3rd edition. New York: Wiley.
Extra reading bibliography:
Mariappan, P. (2019). *Statistics for Business* . New York: Chapman and Hall/CRC, https://doi.org/10.1201/9780429443244 (full text available after login in NOVA IMS network or through VPN connection).
Wilks, S. (1948). *Elementary Statistical Analysis* . Princeton, New Jersey: Princeton University Press. https://www.jstor.org/stable/j.ctt183q2d4 (full text available after login in NOVA IMS network or through VPN connection).
Hogg, R. V., Tanis, E. A. (2001). *Probability and Statistical Inference* . 6th edition, New Jersey: Pearson/Prentice-Hall.
Murteira, B., Ribeiro, C.S., Silva, J.A., Pimenta, C. (2010). *Introdução à Estatística* . Lisboa: Escolar Editora.
Afonso, A., Nunes, C. (2011). *Estatística e Probabilidades. Aplicações e Soluções em SPSS* . Lisboa: Escolar Editora.