Universitat Autònoma de Barcelona

2020/2021

Statistics

Code: 100105
ECTS Credits: 7

Degree	Type	Year	Semester
2500149 Mathematics	OB	3	2

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Alejandra Cabaña Nigro
Email: AnaAlejandra.Cabana@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: No
Some groups entirely in Spanish: No

Other comments on languages

Class materials will be mostly written in English

Prerequisites

Linear algebra. Mathematical analysis. Probability.

Objectives and Contextualisation

In this course, the concept of Inference, in its inductive version, must be fundamentally learned.

The concepts of Modeling, Estimation (by point and intervals) and Goodness of fit must be introduced. And the linear regression techniques.

The students will have to learn:

1. The descriptive and exploratory statistics that will allow to extract and summarize efficiently information of the data.

2. Statistical Inference: how the Statistics quantifies the uncertainty of the information extracted from the data.

3. The modeling of populations, the estimation of parameters, especially maximum likelihood, and the planning and resolution of contrasts of hypotheses (parametric and non-parametric).

3. Basic properties of optimal estimators: invariance, sufficiency, efficiency, bias, variance and asymptotic properties.

4. Establish and solve applied problems. With the examples, the resolution of problems and the practices with statistical software (R), the student will work with concrete models and with real data: inferential for the most important parameters of one and two normal populations. Adjustment tests, inferential methods for the linear model.

Competences

- Apply critical spirit and thoroughness to validate or reject both ones own arguments and those of others.
- Distinguish, when faced with a problem or situation, what is substantial from what is purely chance or circumstantial.
- Recognise the presence of Mathematics in other disciplines.
- Students must be capable of applying their knowledge to their work or vocation in a professional way and they should have building arguments and problem resolution skills within their area of study.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
- Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
- Students must have and understand knowledge of an area of study built on the basis of general secondary education, and while it relies on some advanced textbooks it also includes some aspects coming from the forefront of its field of study.
- Take sex- or gender-based inequalities into consideration when operating within one's own area of knowledge.
- Use computer applications for statistical analysis, numeric and symbolic calculus, graphic display, optimisation or other purposes to experiment with Mathematics and solve problems.
- When faced with real situations of a medium level of complexity, request and analyse relevant data and information, propose and validate models using the adequate mathematical tools in order to draw final conclusions
- Work in teams.

Learning Outcomes

1. Apply critical spirit and thoroughness to validate or reject both ones own arguments and those of others.
2. Descriptively synthesise and analyse datasets.
3. Formulate and solve hypothesis contrast problems in one or two populations
4. Identify the main inequalities and discriminations in terms of sex/gender present in society.
5. Students must be capable of applying their knowledge to their work or vocation in a professional way and they should have building arguments and problem resolution skills within their area of study.
6. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
7. Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
8. Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
9. Students must have and understand knowledge of an area of study built on the basis of general secondary education, and while it relies on some advanced textbooks it also includes some aspects coming from the forefront of its field of study.
10. Understand the basic properties of point and interval estimators.
11. Use large datasets with the help of a statistical package.
12. Use the maximum verisimilitude, Bayes and least square methods to construct estimators
13. Work in teams

Content

The subject is structured in four chapters:

Topic 1: Introduction to Inference.

Summary of the fundamental tools of probability: LLN and CLT.

Binomial and normal. Comparison of two proportions. Pearson test.

Simulation and goodness of fit.

Topic 2: Modeling and estimation.

Normal, gamma, Pareto, Poisson, negative binomial distributions, ...

Estimation methods: moments, maximum likelihood, minimum least squares.

Generating functions.

Topic 3: Assessment of estimators and asymptotic theory.

Bias, mean quadratic error, consistency, asymptotic normality, ...

Cramér-Rao inequality. Fisher information. Efficiency

Asymptotic distribution of the maximum likelihood estimator .

Likelihood ratio test. Scoring and Wald test.

Fisher Theorem. Student Laws, Pearson's $\chi 2$ and Fisher's F.

Contrasts. Null and alternative hypothesis. Type of error

Comparison of two populations and analysis of the variance.

Bootstrap

Topic 4: Linear regression.

Linear regression. Estimate and contrasts.

Unless the requirements enforced by the health authorities demand a prioritization or reduction of these contents.

Methodology

We have theoretical classes, problems and practices.

The new subject will be introduced primarily in theory classes, but it will be necessary to extend the teacher's explanations with the student's autonomous study, with the support of the reference bibliography. Student participation will be valued at the exhibitions of the student. teacher There will be partial control of theory and problems in mid-April. The Virtual Campus will upload material to review the notes collected in class.

The class of problems will be devoted to the resolution oriented to some problems proposed. Students' participation in the problem classes will be especially encouraged.

Practical classes will introduce the use of R software with statistical applications. You will see descriptive and inferential methodologies. You will have to deliver some practical work.

The proposed teaching methodology may experience some modifications depending on the restrictions to face-to-face activities enforced by health authorities.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Master classes: theory	28	1.12	10, 3, 7, 5, 2, 13, 11, 12
Practical work with computer tools	14	0.56	10, 7, 5, 2, 12
Problem classes	14	0.56	10, 3, 7, 12
Type: Supervised			
Tutorials	5	0.2	
Type: Autonomous			
Practical work with computer tools	25	1	
Problem solving (workshops and classes)	20	0.8	10, 3, 7, 5, 2, 13, 11, 12
Study and think problems	39	1.56	10, 3, 7, 5, 2, 13, 11, 12

Assessment

The assessment is carried out continuously throughout the course.

Continuous assessment has several fundamental objectives: Monitor the teaching and learning process, allowing both the student and the teacher to know the degree of achievement of the competencies and correct, if possible, the deviations that occur. Encourage the students' continued effort in the face of over-effort, often useless, last-minute. Verify that the student has achieved the competences determined in the syllabus. For this, the accreditation of a minimum level in all assessment activities will be requested (3 out of 10).

To carry out this evaluation, the following instruments are available: The documentation provided by the students for their practical work (practical dossiers), the results obtained in the laboratory sessions. The grade obtained in this assessment represents 60% of the final mark of the subject.

The continuous assessment is complemented by a final written test. The qualification thus obtained will represent 40% of the final mark of the subject (it allows to recover a part of a first fundamental proof).

The recovery exam will be directed to students who have not passed the minimum level yet to the passed.
Student's assessment may experience some modifications depending on the restrictions to face-to-face activities enforced by health authorities.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Documentation delivered by students	30 %	18	0.72	1, 10, 4, 3, 9, 8, 7, 5, 6, 2, 13, 11, 12
Final Exam	40%	7	0.28	10, 3, 7, 5, 2, 12
Partial Exam-1	30%	5	0.2	10, 3, 7, 5, 2, 11, 12

Bibliography

Fundamental

1. Casella, G..and Berger, R. (2002) . *Statistical Inference, 2º ed*. Wadsworth, Belmont, CA.
2. Casella, G., Berger, R. and Santana, D. (2002). Solutions Manual for Statistical Inference, Second Edition.
3. Luis Ruiz Maya Pérez, Francisco Javier Martín-Pliego López. (2006). Estadística. II, Inferencia. Editoria AC.
4. Millar, R. (2011). *Maximum Likelihood Estimation and Inference*. Wiley.
5. D. Peña. (2002). "Fundamentos de Estadística". Alianza Editorial.
6. D. Peña. (2002). "Regresión y diseño de experimentos". Alianza Editorial.

Complement

1. Das Gupta ("2008) "Asymptotic Theory of Statistics and Porbability", Springer.
2. J.A.Rice (2007), Mathematicla Statistics and data analysis, 3rd Ed, Duxbury/Thomson
3. Versani, J. "Using R for introductory Statistics", Taylor and Francis.
4. M. Kendall and A. Stuart (1983). "The Advanced Theory of Statistics".Griffin and Co. Limited, London.
5. Lehman, E.L. and Romano (2005, 3rd Ed.), J.P, "Testing Statistical Hypotheses", Springer
6. C.R. Rao (1973). "Linear Statistical Inference and its Applications". Wiley, London.
7. M.L. Rizzo (2007). "Statistical computing with R". Computer Science and Data Anal Cysis Series". Chapman & Hall / CRC
8. Williams, D. (2001) "Weighihg the Odds", Cambridge University Press.