UAB
Universitat Autònoma de Barcelona

Modelling and Inference

2026/2027

Code: 104392
Credits: 6

Degree programme	Type	Course
Computational Mathematics and Data Analytics	OB	2

Contact lecturer

Name : Víctor Navas Portella

Email : victor.navas@uab.cat

Group languages

You can consult this information at the end of the document.

Prerequisites

Good knowledge of the contents of the courses taken during the first year of the bachelor's degree is considered very important, especially those of probability and calculus.

Objectives

This is the first course in the Bachelor's degree that focuses on Statistical Inference, a branch of statistics that uses data from a "representative" sample to acquire information about a population. The course is required throughout the Bachelor's degree, as it covers different concepts and techniques that serve as the basis for many of the topics introduced in upcoming courses within the Bachelor's. In particular, the course will start with a brief introduction to statistics, followed by a chapter on parameter estimation (both point and based on confidence intervals), and finally chapters on frequentist-based significance tests and an introduction to classical linear regression models. Special emphasis will be placed on the statistical methods that can be used to compare machine learning algorithms.

To protect everyone's safety, in-person teaching and evaluable activities will be adjusted in accordance with health authority recommendations.

Learning outcomes

- CM14 (Implement strategies to confirm or refute hypotheses.) Implement strategies to confirm or refute hypotheses.
- CM15 (Manage the information for validation through statistical processing.) Manage the information for validation through statistical processing.
- CM16 (Assess, using the data obtained, inequalities on the grounds of sex/gender.) Assess, using the data obtained, inequalities on the grounds of sex/gender.
- KM12 (Identify statistical inference as a tool for forecasting and prediction.) Identify statistical inference as a tool for forecasting and prediction.
- KM13 (Describe the basic properties of point and interval estimators.) Describe the basic properties of point and interval estimators.
- KM14 (Identify the usefulness of Bayesian methods, applying them appropriately.) Identify the usefulness of Bayesian methods, applying them appropriately.
- SM14 (Use the properties of density and distribution functions.) Use the properties of density and distribution functions.
- SM15 (Use suitable statistical software to manage databases, to obtain summary indices of the study variables and to analyse data using inference techniques.) Use suitable statistical software to manage

databases, to obtain summary indices of the study variables and to analyse data using inference techniques.

Contents

Topic 0. Probability Preliminaries (Review)

Topic 1. Introduction to Statistics and Point Estimation

1. Descriptive statistics and inferential statistics.
2. Sampling and common statistics: sample moments and order statistics.
3. Sampling distributions in normal populations.
4. Fundamental concepts: Likelihood Function and Fisher Information.
5. Estimators: Properties (bias, consistency, efficiency, Cramér-Rao bound).
6. Methods for obtaining estimators (Method of Moments, Maximum Likelihood) and asymptotic properties.
7. Delta Method.

Topic 2. Estimation by confidence intervals.

1. Concept of a confidence interval.
2. Construction of confidence intervals.
3. Confidence intervals for the parameters of a single population.
4. Confidence intervals for the parameters of two populations.

Topic 3: Significance tests.

1. Introduction and concepts (Hypotheses, Type I and Type II errors, power, p-value).
2. Foundations: Neyman-Pearson Lemma and Likelihood Ratio Test (LRT).
3. Tests for the parameters of a single population.
4. Tests for the parameters of two populations.
5. Chi-square tests.
5. Non-parametric tests

IMPORTANT: In teaching, the gender perspective involves reviewing androcentric biases and questioning the assumptions and hidden gender stereotypes. This revision involves including the contents of the subjectthe knowledge produced by scientific women, often forgotten, seeking the recognition of their contributions,as well as that of their works in the bibliographical references.

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Practical classes	10	0.4	CM14, CM15, CM16, KM12, KM13, KM14, SM14, SM15
Problems class	12	0.48	CM14, CM15, CM16, KM12, KM13, KM14, SM14
Theory classes	27	1.08	CM14, CM15, CM16, KM12, KM13, KM14, SM14
Exams	15	0.6	CM14, CM15, CM16, KM12, KM13, KM14, SM14
Workshop resolution	23	0.92	CM14, CM15, CM16, KM12, KM13, KM14, SM14, SM15
Problems resolution	33	1.32	CM14, CM15, CM16, KM12, KM13, KM14, SM14

The course is organized into lecture, exercise and lab sessions.

In lectures, we will introduce the concepts and techniques outlined in the course program. Given that the content is mostly based on the standard topics of an introduction to statistical inference course, the recommended bibliography can be used to follow the course. Lecture slides and related material will be available in Moodle. The exercise sessions are intended to work through and understand statistical concepts. Each exercise will be available in Moodle. The goal of the lab sessions is to learn how to apply the methods given in lectures using the statistical software R, as well as how to evaluate the findings.

IMPORTANT: To work more comfortably with R, it is recommended to use the RStudio interface: it is free, \"Open source\" and works with Windows, Mac and Linux. *https://www.rstudio.com/*

OBSERVATION: The gender perspective in teaching goes beyond the contents of the subjects, since it also implies a revision of the teaching methodologies and of the interactions between the students and the teaching staff, both in the classroom and outside. In this sense, participatory teaching methodologies, where an egalitarian, less hierarchical environment is generated in the classroom, avoiding stereotyped examples in gender and sexist vocabulary, with the aim of developing critical reasoning and respect for the diversity and plurality of ideas, people and situations, tend to be more favorable to the integration and full participationof the students in the classroom, and therefore their effective implementation in this subject will besought.

Annotation: within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continuous assessment activities

Title	Weight	Hours	ECTS	Learning outcomes
Final exam	0,50	10	0.4	CM14, CM15, CM16, KM12, KM13, KM14, SM14
Grading exercises	0,15	12	0.48	CM14, CM15, CM16, KM12, KM13, KM14, SM14, SM15

Mid-term exam	0,35	8	0.32	CM14, CM15, CM16, KM12, KM13, KM14, SM14

During the lecture sessions the basic concepts of the subject will be introduced and a wide set of examples will be presented. In the problems and practices sessions, exercises will be solved and practices with R will be carried out. Classroom attendance is recommended to have an idea about the course in general, as well as the exercises and practices.

Assessment:

The student's grade will be the weighted average of the following activities:

PAC1: partial exam, which accounts for 35% of the grade.

PAC2: Evaluation of the computational practice part accounts for 15% of the grade.

Final exam: which will consist of some conceptual questions in the form of short questions and some problems in which you will have to solve a series of exercises similar to those that have been worked on in class sessions. This exam represents the remaining 50% of the grade.

Important: if the grade for any of these activities does not reach 3 out of 10, it will count as 0 in the calculation of the final grade.

Recovery: if this grade does not reach 5, the student has the right to another opportunity to pass the subject through the recovery exam. In this exam, 85% of the grade corresponding to the final exam and PAC1 can be recovered. The practical part with R (PAC2) is not recoverable. In no case can the recovery exam be used to raise the grade if the student has already passed the subject with the first exam.

Single evaluation:

The student who has taken advantage of the single evaluation modality must take a final exam that will consist of some conceptual questions in the form of short questions and some problems in which they will have to solve a series of exercises similar to those that have been worked on in class sessions. Once completed, you will deliver, in addition to the exam, the exercises relatedto the R computer practices completed throughout the course.

The student's grade will be the weighted average of the two previous activities, where the final exam will account for 85% of the mark, and the evaluation of the answer sheets of the computer practices with R the remaining 15%.

Important: if the grade for any of these activities does not reach 3 out of 10, it will count as 0 in the calculation of the final grade.

If this qualification does not reach 5, the student has the right to another opportunity to pass the subject through the recovery exam that will be held on the date set by the coordination of the degree. In this exam it will be possible to recover 85% of the mark corresponding to the final exam. The practical part with R is not recoverable. In no case can the recovery exam be used to raise the grade if the student has already passed the subject with the first exam.

Use of Artificial Intelligence (AI)

For this course, the use of AI technologies is permitted exclusively for support tasks, such as:

- Literature search.
- Code debugging, text proofreading, and translations

Students must:

- Clearly identify which parts have been generated with the assistance of this technology.
- Specify the tools used.
- Include a critical reflection on how these tools have influenced both the process and the final outcome of the activity.

A lack of transparency in the use of AI in gradable activities will be considered a breach of academic honesty.

The teaching staff reserves the right to call any student for a validation interview regarding the content and development process of any course activity.

The inappropriate or fraudulent use of AI may result in a partial or total penalty on the activity's grade, or more severe sanctions in more serious cases.

Bibliography

1. Daalgard, P.: Introductory Statistics with R. Springer. 2008.
2. Daniel, W.W.: Biostatistics. Wiley. 1974.
3. DeGroot, M. H.: Schervish, M.J. Probability and Statistics. Pearson Academic. 2010.
4. Delgado, R.: Probabilidad y Estadística con aplicaciones. 2018. https://www.amazon.es/Probabilidad-Estad%C3%ADstica-aplicaciones-Rosario-Delgado/dp/1983376906
5. Heumann, C., Schomaker, M., Shalbh: Introduction to Statistics and Data Analysis: With Exercises, Solutions and Applications in R. Second Edition. Springer. 2023. Available on-line through UAB: https://link.springer.com/book/10.1007/978-3-031-11833-3
6. Plaue, M.: Data Science: An Introduction to Statistics and Machine Learning. Springer. 2023. Available on-line through UAB: https://link.springer.com/book/10.1007/978-3-662-67882-4
7. R Tutorial. An introduction to Statistics. https://cran.r-project.org/manuals.html
8. Salsburg, D. The Lady tasting tea : how statistics revolutionized science in the twentieth century. 2002. ISBN-13: 978-0805071344
9. Silvey, S.D.: Statistical Inference. Chapman&Hall. 1975.

https://app.jove.com/science-education/v/12796/introduction-to-statistics

Software

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Course groups and languages

The information provided is provisional until November 30. After this date, you will be able to consult the language of each group through this link. To access the information, you will need to enter the course CODE

Type of teaching	Group	Language	Semester	Shift

(TE) Theory	1	Catalan	first semester	morning-mixed
(PLAB) Practical laboratories	1	Catalan	first semester	morning-mixed
(SEM) Seminars	1	Catalan	first semester	morning-mixed