

**Informatics Tools for Statistics**

Code: 104849  
ECTS Credits: 6

| Degree                     | Type | Year | Semester |
|----------------------------|------|------|----------|
| 2503852 Applied Statistics | FB   | 1    | 1        |

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: Joaquim Roe Vellve  
Email: Joaquim.Roe@uab.cat

## Use of Languages

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

## Other comments on languages

This document is a translation of the Catalan original. In case of inaccuracies or errors, the Catalan version is the valid, official, guide for the course.

## Teachers

Aureli Alabert  
Sundus Zafar

## Prerequisites

Because it is an undergraduate and in the first semester, it has no prerequisite.

## Objectives and Contextualisation

The main objectives of the subject are the following:

- Familiarize yourself with the use of an algebraic and calculating manipulator. This manipulator must be considered as a common tool when studying the rest of the subjects.
- Learn how to structure and write scientific texts with the LaTeX word processor.
- Familiarize yourself with the concept of statistical package. In particular, create and transform databases and get used to work environments in graphic mode and command line.
- Learn how to use a command line operating system, taking advantage of the power of current computers to merge, separate or extract data from files or file sets.
- Introduce yourself to the formalization of algorithms using a programming language.

## Competences

- Make efficient use of the literature and digital resources to obtain information.
- Select the sources and techniques for acquiring and managing data for statistical processing purposes.

- 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 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.
- Use quality criteria to critically assess the work done.
- Use software for statistical analysis, numerical and symbolic analysis, graphic visualisation, optimisation or others, to solve problems.
- Work cooperatively in a multidisciplinary context, respecting the roles of the different members of the team.

## Learning Outcomes

1. Correctly identify the types of data and measurements.
2. Critically assess the work done on the basis of quality criteria.
3. Identify the advantages and disadvantages of the internet as a major source of information in statistics.
4. Make effective use of references and electronic resources to obtain information.
5. Show skill in operations with symbolic calculus packages.
6. 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.
7. 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.
8. Use programming languages related to statistical applications competently.
9. Use science editors to present papers, problems, reports and scientific texts in general.
10. Work cooperatively in a multidisciplinary context, accepting and respecting the roles of the different team members.

## Content

1. Brief introduction to computing. Computing resources at the University available to use in the course.
2. Textprocessor (LaTeX): Localization of the software. Structure of a TeX file.
3. Statistical packs (R): Work environments. Declaration of variables. Creation, obtaining and manipulation of databases. Descriptive tools Graphic environment.
4. Computer Algebra System (Sage): numerical and algebraic calculations. Function graphs. Resolution of equations. Definition of functions. Lists, sets and successions. Logical programming, iterations and procedures.
5. Operating system (Bash): graphic environment. The console First instructions and obtaining help. Manipulation of files. Data manipulation: regular expressions, grep and awk.
6. Programming (Python): Introduction to Python.

## Methodology

The practical classes are held in computer rooms or classrooms prepared for the use of laptops.

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.

## Activities

| Title             | Hours | ECTS | Learning Outcomes       |
|-------------------|-------|------|-------------------------|
| Type: Directed    |       |      |                         |
| Practice sessions | 52    | 2.08 | 2, 9, 1, 3, 6, 10, 8, 4 |

Type: Autonomous

|                                             |    |      |                     |
|---------------------------------------------|----|------|---------------------|
| LaTeX document preparation                  | 10 | 0.4  | 2, 9, 3, 10, 4      |
| Preparación del examen de sistema operativo | 19 | 0.76 | 2, 1, 3, 6, 8, 4    |
| Preparation of the computer algebra system  | 20 | 0.8  | 2, 1, 3, 7, 6, 8, 4 |
| Writing a Python program                    | 20 | 0.8  | 2, 1, 3, 6, 8, 4    |
| Writing the statistical package program     | 20 | 0.8  | 2, 1, 3, 6, 8, 4    |

## Assessment

Continuous evaluation activities provide a grade; to pass the subject this grade has to be greater or equal than 5, and the mark obtained on each subject block has to be greater or equal than 3. A time of 4 hours is reserved to re-evaluate any exam that the student failed.

## Assessment Activities

| Title                        | Weighting | Hours | ECTS | Learning Outcomes    |
|------------------------------|-----------|-------|------|----------------------|
| Computer algebra system exam | 0.26      | 3     | 0.12 | 2, 5, 9, 1, 6, 8, 4  |
| Final exam                   | 0.43      | 4     | 0.16 | 2, 1, 3, 6, 8, 4     |
| LaTeX document delivery      | 0.13      | 0     | 0    | 2, 9, 3, 7, 6, 10, 4 |
| Operating system exam        | 0.17      | 2     | 0.08 | 2, 1, 3, 6, 8, 4     |
| Python program delivery      | 0.22      | 0     | 0    | 2, 1, 3, 6, 8, 4     |
| Statistic package program    | 0.22      | 0     | 0    | 2, 1, 3, 6, 8, 4     |

## Bibliography

As all work is done on computers, the main source of information will be the *help* of the programs that are used. In addition, as a complementary bibliography we recommend the following online resources.

- Tobias Oetiker, Hubert Partl, Irene Hyna and Elisabeth Schlegl. The not so short introduction to LaTeX2E (or LaTeX in 139 minutes). <https://tobi.oetiker.ch/lshort/lshort.pdf>
- W.N. Venables, D.M. Smith and the R Development Core Team: An introduction to R. <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>
- GNU Bash manual, <https://www.gnu.org/software/bash/manual/>
- Guido van Rossum, Fred L. Drake, Jr., editors, The Python Language Reference, <http://marvin.cs.uidaho.edu/Teaching/CS515/pythonReference.pdf>
- The Sage Reference Manual, <https://doc.sagemath.org/html/en/reference/>

## Software

SageMath, R, Python, LaTeX and GNU/Linux.