UAB
Universitat Autònoma de Barcelona

Statistics and Data Analysis

Code: 44079
ECTS Credits: 9

2025/2026

Degree	Type	Year
High Energy Physics, Astrophysics and Cosmology	OB	0

Contact

Name: Francisco Javier Rico Castro

Email: franciscojavier.rico@uab.cat

Teachers

Abelardo Moralejo Olaizola

Carles Sánchez Alonso

Jorge Carretero Palacios

Pau Tallada Crespí

Martin Borstad Eriksen

Francesc d'Assis Torradeflot Curero

Teaching groups languages

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

Prerequisites

It is needed a personal computer with a running installation of Python 3.

Install Python 3 with the Anaconda installer. In this way, your Python distribution will contain all the associated packages needed for this course.

Follow these steps:

1. Download the Anaconda installer for Python 3 here https://www.anaconda.com/download/

2. Follow the installation instructions - both GUI or terminal versions work fine. If prompted, select the option to add the new anaconda directory to your path.

The use of GNU/Linux is highly recommended.

Objectives and Contextualisation

In this course we will learn how to distill scientific knowledge from experimental data, a process that relies on statistical methods. We will learn the basics concepts of Probability and Statistics (in their Frequentist and Bayesian frameworks). In addition, we will study and practice several particular statistical methods and data analysis techniques usually used in the fields of High Energy Physics, Astrophysics and Cosmology. To that aim, we will learn and practice the use of modern statistics and analysis software tools.

Competences

- Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
- Use mathematics to describe the physical world, select the appropriate equations, construct adequate models, interpret mathematical results and make critical comparisons with experimentation and observation.
- Use the adequate software, programming languages and computer packages to research problems related to high energy physics, astrophysics and cosmology.
- Work in a group and take on responsibility, interacting professionally and constructively with other people with complete respect for their rights.

Learning Outcomes

1. Apply data analysis techniques to problems in the areas of particle physics, astrophysics and cosmology, as well as other close but different areas.
2. Learn how statistical analysis software works.
3. Use Monte Carlo techniques to model real problems of physics.
4. Work in small groups to solve problems of data analysis.

Content

Part 1: Basic concepts on probability, statistics and Monte Carlo techniques

Part 2: Python for Statistics and Data Analysis

Part 3: Parameter estimation, Hypothesis test and Unfolding

Part 4: Bayesian Statistics

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Lectures	56	2.24	1, 2, 3
Study of theory and practical examples	64	2.56	1, 2, 4, 3
Type: Autonomous			
Discussion, workgroups, problem solving	60	2.4	1, 2, 4, 3

- Theory lectures including practical examples in the fields of High Energy Physics, Astrophysics and Cosmology
- Homework exercises to be solved by students alone or in small groups
- Discussion of problems during classes and tutorials
- Hands-on sessions on software tools for statistics and data analysis (in Python programming language)
- Explanation and discussion of sample code/algorithms in Python programming languages during classes and tutorials

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

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Attendance and active participation to the lectures	5%	0	0	1, 2, 3
Resolution of a final, synthesis exam	45%	5	0.2	1, 2, 3
Resolution of class exercises	50%	40	1.6	1, 2, 4, 3

The evaluation will take into account:

- Attendance and active participation to the lectures
- Resolution of specific exercises along the course
- Resolution of a final exam

For those students not passing the course after the regular evaluation procedure, there will be a *recuperation* evaluation round consisting on a synthesis exam. There will be no threshold mark to be eligible for the recuperation evaluation round, other than the general requirement of having been evaluated at least for a 66% of the total qualification activities in the first round.

This subject/module does not foresee the single assessment system.

Bibliography

- G. Bohm and G. Zech; "Introduction to Statistics and Data Analysis for Physicists", 3rd Edition, 2017, Verlag Deutsches Elektronen-Synchrotron (available on-line https://s3.cern.ch/inspire-prod-files-d/da9d786a06bf64d703e5c6665929ca01)
- F. James; "Statistical Methods in Experimental Physics", 2nd Edition, 2006, World Scientific
- G. Cowan; "Statistical Data Analysis", 1998, Oxford University Press
- A. Gelman, J. B. Carlin, H. S. Stern, et al. "Bayesian Data Analysis", 3rd Edition, 2013, CRC Press

Software

We will introduce and make use of the Python programming language (see the "Prerequisists" section for installation details).

In particular, we will study and use the following Python libraries: numpy, pandas, matplotlib, scipy and scikit learn.

Groups and Languages

Please note that this information is provisional until 30 November 2025. You can check it through this link. To consult the language you will need to enter the CODE of the subject.

Name	Group	Language	Semester	Turn
(TEm) Theory (master)	1	English	first semester	morning-mixed