

**Statistics and Data Analysis**

Code: 44079  
ECTS Credits: 9

Degree	Type	Year	Semester
4313861 High Energy Physics, Astrophysics and Cosmology	OB	0	1

**Contact**

Name: Francisco Javier Rico Castro  
Email: FranciscoJavier.Rico@uab.cat

**Use of languages**

Principal working language: english (eng)

**Teachers**

Ramon Miquel Pascual  
Abelardo Moralejo Olaizola  
Jorge Carretero Palacios  
Pau Tallada Crespi  
Santiago Serrano Elorduy  
Francesc d'Assis Torradeflot Curero

**Prerequisites**

For the Python Bootcamp (part 2), it is highly needed to bring a personal laptop with a running installation of Python 3.6.

For that, install Python 3.6 with the Anaconda intaller. In this way, your Python distribution will contain all the associated packages needed for this course.

Follow these steps:

1. Download Anaconda installer for Python 3.6 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 Linux or Mac 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.

## Skills

- 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 to use the Root statistical analysis tool.
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

## Methodology

- 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

## Activities

Title	Hours	ECTS	Learning outcomes
<b>Type: Directed</b>			
Lectures	56	2.24	1, 2, 3
Study of theory and practical examples	40	1.6	1, 2, 3, 4
<b>Type: Autonomous</b>			
Discussion, workgroups, problem solving	34	1.36	1, 2, 3, 4

## Evaluation

The evaluation will take into account:

- Attendance and active participation to the lectures
- Resolution, for each of the course parts, of specific take-home exercises
- Resolution of a final, synthesis take-home exam

For those students not passing the course after the regular evaluation procedure, there will be a *recuperation* evaluation round consisting also on specific take-home exercises for the different course parts, plus a final, 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.

## Evaluation activities

Title	Weighting	Hours	ECTS	Learning outcomes
Attendance and active participation to the lectures	5%	0	0	1, 2, 3
Resolution of Part 1 exercises	15%	15	0.6	1, 2, 3, 4
Resolution of Part 3 exercises	20%	20	0.8	1, 2, 3, 4
Resolution of Part 4 exercises	10%	10	0.4	1, 2, 3, 4
Resolution of a final, synthesis exam	50%	50	2	1, 2, 3, 4

## Bibliography

- G. Cowan; "Statistical Data Analysis", 1998, Oxford University Press
- K. A. Olive et al. (Particle Data Group); "Review of Particle Physics", Chin. Phys. C38 (2014) 090001 A. Bevan; "Statistical Data Analysis for the Physical Science", 2013, Cambridge
- F. James; "Statistical Methods in Experimental Physics", 2nd Edition, 2006, World Scientific
- L. Lyons, "Statistics for Particle and Nuclear Physicists", 1986, Cambridge University Press
- B. P. Roe, "Probability and Statistics in Experimental Physics", 1992, Springer
- A. G. Frodesen, et al., "Probability and statistics in particle physics", 1979, Columbia University Press D. Sivia and J. Skilling, "Data Analysis, A Bayesian Tutorial", 2nd ed., 2006, Oxford University Press A. Gelman, "Bayesian Data Analysis", 1995, CRC Press
- R. J. Barlow, "Statistics", 1989, J. Wiley
- W.T. Press et al., "Numerical Recipes: The Art of Scientific Computing", Cambridge University Press. E.T. Jaynes, "Probability Theory: The Logic of Science", Cambridge University Press.
- A. Stuart et al., "Kendall's Advanced Theory of Statistics", Vol 2A. Wiley.
- F. James, "Monte Carlo Theory and Practice", Rep. Prog. Phys. 43 (1980) 73.