

Degree programme	Type	Course
Applied Statistics	OB	2

Contact lecturer

Name : Angel Gonzalez Wong

Email : angel.gonzalez@uab.cat

Teaching staff

Carolina Soriano Tarraga

Gianluigi Caltabiano

Angel Gonzalez Wong

Group languages

You can consult this information at the [end](#) of the document.

Prerequisites

Basic knowledge of the English language is required as part of the teaching material is in that language.

Objectives

The great scientific and technological advances that have occurred in recent decades in the fields of biology and computer science, as well as the development of powerful computer systems, fast internet connections and large information databases throughout the world, have allowed researchers an easily access to an enormous amount of unprecedented biological data. This course introduces students to the field of Bioinformatics, a specialty that uses computer databases to store, retrieve, and aid in the understanding of biological information. This specialty is an emerging field of science that deals with the development of various analytical methods and tools for the study of biological data in an efficient and rigorous manner.

During the course, the basic concepts of Bioinformatics and Computational Biology will be introduced, as well as the main methods for analyzing information from genomic and protein sequences, as well as for managing pharmacological data. The practical sessions will complement this knowledge, allowing students to become familiar with the details and the use of the main tools and online resources of the specialty.

Learning outcomes

- CM14 (Propose the statistical model needed to analyse data sets belonging to real studies.) Propose the statistical model needed to analyse data sets belonging to real studies.
- KM17 (Recognise the statistical models for the analysis of data with different structures and

complexities that frequently appear in different fields of application.) Recognise the statistical models for the analysis of data with different structures and complexities that frequently appear in different fields of application.

- KM18 (Recognise the language of applications of economics and finances, biomedical science and engineering, provided by research and innovation in the field of statistics.) Recognise the language of applications of economics and finances, biomedical science and engineering, provided by research and innovation in the field of statistics.
- SM16 (Select appropriate sources of information for the statistical work.) Select appropriate sources of information for the statistical work.
- SM18 (Refine the information available for subsequent statistical processing.) Refine the information available for subsequent statistical processing.
- SM19 (Analyse complex data, whether this is due to their characteristics or their size.) Analyse complex data, whether this is due to their characteristics or their size.

Contents

- Introduction to Bioinformatics and Molecular Biology Concepts
- Bioinformatics Databases, Bioconductor, and the R Ecosystem
- Introduction to Genomics. Major Genomic Projects. Human Genome Project
- Genetic Variations and Phenotypes
- Comparison of Biological Sequences
- Sequence search and alignment
- Proteomics
- Concepts of Pharmacology and Chemoinformatics

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Practical classes	26	1.04	
Tutoring	10	0.4	
Theory classes	26	1.04	
Study	83	3.32	

The course is organized in sessions of two hours. Each session consists of a theoretical part (theory classroom) that will introduce the new concepts followed by a practical part (computer room) where the students will work on the implementation of concepts explained in the theoretical part. In each session the teacher will indicate the students some tasks to do autonomously, such as reading articles or sending reports. The material used by the teachers will be available on the Virtual Campus of the course.

In this course, the use of Artificial Intelligence (AI) technologies is allowed as an integral part of the development of the work, provided that the final result reflects a significant contribution from the student in terms of analysis and personal reflection. The student must clearly identify which parts have been generated using this technology, specify the tools used, and include a critical reflection on how these have influenced both the process and the final outcome of the activity. Lack of transparency in the use of AI will be considered a breach of academic integrity and may result in a penalty in the grade for the activity, or more serious sanctions in severe cases.

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
Presentation class exercises	30	1	0.04	CM14, KM17, KM18, SM16, SM18
Theoretical-practical exam 2	35	2	0.08	KM17, KM18, SM19
Theoretical-practical exam 1	35	2	0.08	KM17, KM18, SM19

- Theoretical-practical test 1 (35%)
- Theoretical-practical test 2 (35%)
- Practical exercises (30%)

Final recovery exam for those students who have not passed the subject through continuous assessment. To participate in this exam, students must have been previously evaluated in a set of activities that account for at least two-thirds of the total grade for the subject. The final exam will cover the entire course syllabus, and the minimum passing grade will be 5 points. This course does not include the single assessment system.

The commission of any irregularity in an assessment activity (academic fraud, plagiarism, or improper use of AI, unless such use is expressly authorized in the course guide) that may lead to a significant change in the grade will result in that activity being graded with a 0. If the course guide establishes that passing the subject requires obtaining a minimum mark in that assessment activity, or if multiple irregularities occur in the assessment activities of the same subject, the final grade for the subject will be 0. In addition, disciplinary proceedings may be initiated against any student who incurs any of these irregularities.

Bibliography

- Lesk A.M. *Introduction to Bioinformatics*. Oxford University Press 2005.
- Attwood, T.K., Parry-Smith, D.J., *Introducción a la Bioinformática*. Pearson Education, 2002.
- Foulkes A.S. *Applied Statistical Genetics with R. For Population-based Association Studies*. Springer Dordrecht Heidelberg London New York. ISBN 978-0-387-89553-6
- Gonzalez JR, Cáceres A. *Omic association studies with R and Bioconductor*. Chapman and Hall/CRC, ISBN 9781138340565, 2019.
- Selzer PM, Koch O, Marhöfer RJ. *Applied Bioinformatics: An Introduction*. Cham: Springer International Publishing, 2018.
- Lee JK. *Statistical Bioinformatics: A guide for life and biomedical science researchers*. Hoboken, N.J.: Wiley-Blackwell, 2010.
- Baxevanis AD, Bader GD, Wishart DS. *Bioinformatics: A practical guide to the analysis of genes and proteins*. Fourth edition. ed. Hoboken, NJ: Wiley, 2020.
- Specialized readings and articles available on the course's virtual campus.
- <https://www.bioconductor.org/>

Software

R: <https://www.r-project.org/>

Rstudio: <https://www.rstudio.com/>

Datawarrior: <https://openmolecules.org/datawarrior/>

PyMol: <https://pymol.org/2/>

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	afternoon
(PLAB) Practical laboratories	1	Catalan	first semester	afternoon