Universitat Autònoma de Barcelona

2022/2023

Bioinformatics

Code: 104872
ECTS Credits: 6

Degree	Type	Year	Semester
2503852 Applied Statistics	OB	2	2

Contact

Name: Angel Gonzalez Wong
Email: angel.gonzalez@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

Teachers

Leonardo Pardo Carrasco
Gianluigi Caltabiano
Angel Gonzalez Wong

Prerequisites

None

Objectives and Contextualisation

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.

Competences

- Analyse data using statistical methods and techniques, working with data of different types.
- Correctly use a wide range of statistical software and programming languages, choosing the best one for each analysis, and adapting it to new necessities.
- Critically and rigorously assess one's own work as well as that of others.
- Identify the usefulness of statistics in different areas of knowledge and apply it correctly in order to obtain relevant conclusions.

- Interpret results, draw conclusions and write up technical reports in the field of statistics.
- Make efficient use of the literature and digital resources to obtain information.
- Select and apply the most suitable procedures for statistical modelling and analysis of complex data.
- Select statistical models or techniques for application in studies and real-world problems, and know the tools for validating them.
- 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 be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
- 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.
- Work cooperatively in a multidisciplinary context, respecting the roles of the different members of the team.

Learning Outcomes

1. Apply statistical methods to the analysis of data on gene expression.
2. Critically assess the work done on the basis of quality criteria.
3. Draw conclusions that are consistent with the experimental context specific to the discipline, based on the results obtained.
4. Draw up technical reports that clearly express the results and conclusions of the study using vocabulary specific to the field of application.
5. Interpret statistical results in applied contexts.
6. Justify the choice of method for each particular application context.
7. Make effective use of references and electronic resources to obtain information.
8. Manage and exploit databases available in statistics institutes and other public bodies.
9. Reappraise one's own ideas and those of others through rigorous, critical reflection.
10. Recognize the advantages and drawbacks of the different statistical methodologies when studying data from a variety of disciplines.
11. Recognize the importance of the statistical methods studied within each particular application.
12. Recognize the statistical inference methods most commonly used in bioinformatics.
13. 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.
14. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
15. Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
16. 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.
17. Use different programmes, both open-source and commercial, associated with the different applied branches.
18. Work cooperatively in a multidisciplinary context, accepting and respecting the roles of the different team members.

Content

1. Introduction to Bioinformatics
2. Concepts of Molecular Biology
3. Bioinformatics Databases
4. Introduction to Genomics
5. Main Genomic Projects. The Human Genome Project

6. Genetic Variations and Phenotypes
7. Introduction to the Comparison of Biological Sequences
8. Sequence Alignments and Databases Searches
9. Concepts of Pharmacology. Chemoinformatics
10. Proteomics

Methodology

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.

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			
Practical classes	26	1.04	9, 2, 4, 18
Theory classes	26	1.04	1, 3, 5, 6, 15, 13, 14, 12, 11
Type: Supervised			
Tutoring	10	0.4	2, 6
Type: Autonomous			
Study	83	3.32	1, 3, 5, 6, 15, 13, 14, 12, 11, 7

Assessment

- Theoretical-practical test 1 (35%)

- Theoretical-practical test 2 (35%)

- Practical exercises (30%)

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Presentation class exercises	30	1	0.04	9, 2, 4, 8, 16, 10, 18, 17
Theoretical-practical exam 1	35	2	0.08	1, 3, 5, 6, 15, 13, 14, 12, 11, 7
Theoretical-practical exam 2	35	2	0.08	1, 3, 5, 6, 15, 13, 14, 12, 11, 7

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.

Software

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

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

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