Universitat Autònoma de Barcelona

2021/2022

Biostatistics

Code: 100766
ECTS Credits: 6

Degree	Type	Year	Semester
2500250 Biology	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: Giulia Binotto
Email: Giulia.Binotto@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

Maria Jolis Giménez
Florent Balacheff

Prerequisites

Although there are no official prerequisites, it is advisable for the student to review:
1) Combinatorics and Newton's binomial.
2) The probability and the statistics that have been studied in secondary school.
3) Elementary functions (exponential, logarithm) and series.

Objectives and Contextualisation

Contextualization:

This is a basic, instrumental type course that introduces probabilistic tools and basic statistics in Biology studies in order to analyze biological data from the description of natural phenomena or experiments. These tools will be used for other subjects of the degree and are essential for the future graduate in Biology training both for the pursuit of their profession and for research. Along with Mathematics, this is characterized by the fact that in addition to its own content, it helps the student to develop scientific rigor and logical thinking.

Training objectives of the subject: It is intended for the student to...

1. Be able to use fluently the language of the probability and the statistics used in Biology.
2. Learn how to explore descriptive methods with various sets of data, resulting from the observation of biological phenomena or experimentation.
3. Be able to raise the most suitable probabilistic models in different situations, and know how to use the probability rules to calculate the probability of the events of interest.
4. Know and understand the concept of random variable, know classical examples of random variables and in what situations are used for modeling.

5. Learn how to use the methods of statistical inference (confidence intervals and hypothesis tests) to reach conclusions on one or more populations based on partial information contained in random samples.
6. Know computer tools (R software and R Commander user graphical interface) for the statistical treatment of data.
7. Apply common sense and develop a critical spirit in dealing with the problems that will have to be solved, both at the time of its resolution and resolution, as well as at the time of drawing conclusions and making decisions.

Competences

- Act with ethical responsibility and respect for fundamental rights and duties, diversity and democratic values.
- Apply statistical and computer resources to the interpretation of data.
- Be able to analyse and synthesise
- Make changes to methods and processes in the area of knowledge in order to provide innovative responses to society's needs and demands.
- Obtain information, design experiments and interpret biological results.
- 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 develop the necessary learning skills to undertake further training with a high degree of autonomy.
- 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.
- Take account of social, economic and environmental impacts when operating within one's own area of knowledge.
- Take sex- or gender-based inequalities into consideration when operating within one's own area of knowledge.
- Understand, interpret and use mathematical and statistical tools to solve problems in biology.
- Work in teams.

Learning Outcomes

1. Analyse a situation and identify its points for improvement.
2. Apply statistical and computer resources to the interpretation of data.
3. Be able to analyse and synthesise.
4. Critically analyse the principles, values and procedures that govern the exercise of the profession.
5. Design experiments based on knowledge of statistics.
6. Identify and interpret the statistical tools that can be used to solve problems in biology.
7. Obtain information from experimental data, present it correctly and interpret it.
8. Propose new methods or well-founded alternative solutions.
9. 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.
10. 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.
11. Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
12. Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
13. 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.

14. Take account of social, economic and environmental impacts when operating within one's own area of knowledge.
15. Take sex- or gender-based inequalities into consideration when operating within one's own area of knowledge.
16. Use statistical tools to solve problems in biology.
17. Work in teams.

Content

1. Descriptive statistics.

- Data and random error. Measurement scales.
- Descriptive analysis of data from a single variable: frequency distributions, graphic representations, numerical summaries (position, dispersion and shape measurements).
- Descriptive analysis of data from two variables: correlation and regression line, tables of contingency.

2. Probability.

- Basic properties of probability. Conditional probability. Formula of total probabilities. Bayes Formula. Independence of events.
- Expectation and variance of a random variable.
- Discrete random variables. Bernoulli, Binomial and Hypergeometric distributions.
- Continuous random variables. Normal distribution. Approximation of the Binomial by the Normal distribution.
- Independence of random variables.

3. Statistical inference.

- Introduction to Statistics: population and sample, parameters and estimators.
- Distribution of the mean sample in the normal case with known variance: Z-statistic. Confidence interval for the mean of a normal population with known variance.
- Student's distribution. The case of the unknown variance: the T-statistic and the confidence interval for the mean of a normal population with unknown variance.
- Hypothesis test concept. Test for the mean and for the variance of a Normal population. Test for the proportion.
- Introduction to hypothesis tests. Hypothesis test for the mean of the normal with known variance and with unknown variance. Tests for the population proportion.
- Test of comparison of means and variances for two Normal populations. Test of comparison of proportions.
- The Shapiro-Wilk test of normality. Non-parametric tests for the comparison of means.
- Chi-square test for the goodness of fit and the independence.

Part of the topics will be developed in practice classes with statistics software.

*Unless the requirements enforced by the health authorities demand a prioritization or reduction of these contents.

Methodology

The center of the learning process is the work of the student. The student learns working, being the mission of the teaching staff help him/her in this task by providing information or showing him/her the sources where one can get it and directing your steps in a way that the learning process can be carried out effectively. In line with these ideas, and in accordance with the objectives of the subject, the course development is based on the following activities:

Theory classes:
The student acquires the scientific-technical knowledge of the subject assisting the theory classes, complementing them with self-study of the subjects explained in order to assimilate the concepts and the

procedures, to detect doubts and to realize summaries and schematics of the subject. In the theory classes, the professor introduces the basic concepts of the subject, showing their application. The classes are taught with blackboard and the support of ICT.

Problems and practices:
Problems and practices are sessions with a smaller number of students where the scientific-technical knowledge presented in the theory classes is worked on to complete their understanding and deepen it by solving problems and practical cases, with the appropriate software. Students will work individually or in groups, under the supervision of the professor, solving the proposed problems. This will be done both in class and autonomously by the student.

In the computer practice sessions, the student will learn to use computer tools for descriptive analysis of data sets and statistical inference.

*The proposed teaching methodology may experience some modifications depending on the restrictions to face-to-face activities enforced by health authorities.

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			
Problem classes and practices	22	0.88	15, 14, 4, 1, 2, 5, 6, 7, 8, 13, 12, 11, 9, 10, 3, 17, 16
Theory classes	30	1.2	15, 14, 4, 1, 5, 6, 7, 8, 13, 12, 10, 3, 16
Type: Supervised			
Individual Tutorials	8	0.32	15, 14, 4, 1, 2, 5, 6, 7, 8, 13, 12, 11, 9, 10, 3, 16
Type: Autonomous			
Study + work of problems and practices	83	3.32	15, 14, 4, 1, 2, 5, 6, 7, 8, 13, 12, 11, 9, 10, 3, 17, 16

Assessment

The evaluation of the subject consists of a part of continuous evaluation of the acquired competences: there will be two partial exams, each with a weight of 35%. These two partials will be the recoverable part of the subject.
The evaluation of the practices with computer will have a weight of 30% in the final evaluation of the subject. The mark of the practice part will be obtained from the delivery of some works.
To participate in the recovery examination, the students must have been previously evaluated in a series of activities whose weight equals to a minimum of 2/3 of the total grade of the subject. Therefore, the students will obtain the "Non-evaluable" qualification when the evaluation activities carried out have a weighting of less than 67% in the final grade.

*Student's assessment may experience some modifications depending on the restrictions to face-to-face activities enforced by health authorities.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Partial exams	70%	4	0.16	15, 14, 4, 1, 2, 5, 6, 7, 8, 13, 12, 11, 9, 10, 3, 16
Practice works	30%	0	0	15, 14, 4, 1, 2, 5, 6, 7, 8, 13, 12, 11, 9, 10, 3, 17, 16
Recovery exam	70%	3	0.12	15, 14, 4, 1, 2, 5, 6, 7, 8, 13, 12, 11, 9, 10, 3, 16

Bibliography

Bardina, X. Farré, M. *Estadística descriptiva.* Manuals UAB, 2009.
Besalú, M. Rovira C. *Probabilitats i estadística.* Publicacions i Edicions de la Universitat de Barcelona, 2013.
Delgado, R. *Probabilidad y Estadística para ciencias e ingenierías.* Delta, Publicaciones Universitarias. 2008.
Devore, Jay L. *Probabilidad y Estadística para ingeniería y ciencias.* International Thomson Editores. 1998.
Milton, J. S. *Estadística para Biología y Ciencias de la Salud.* Interamericana de España, McGraw-Hill, 2007 (3a ed. ampliada).
Remington, R. D. Schork, M. A. *Estadística Biométrica y Sanitaria.* Prentice/Hall Internacional, 1974.

Software

In the computer practice sessions, the student will learn to use the free software R with the graphical user interface R Commander (or an equivalent graphical interface), in order to apply the statistical tools for the descriptive analysis of data sets and statistical inference.