Universitat Autònoma de Barcelona

2023/2024

Biostatistics and Data Analysis

Code: 101917
ECTS Credits: 6

Degree	Type	Year	Semester
2501230 Biomedical Sciences	FB	1	2

Contact

Name: Maria Mercedes Campillo Grau

Email: mercedes.campillo@uab.cat

Teaching groups languages

You can check it through this link. To consult the language you will need to enter the CODE of the subject. Please note that this information is provisional until 30 November 2023.

Teachers

Leonardo Pardo Carrasco

Prerequisites

There are no official prerequisites, however prior knowledge of elementary mathematics including the concepts of differentiation and integration is recommended.

Objectives and Contextualisation

Biostatistics and Data Analysis aims to introduce the student to the fundamental knowledge and use of the basic tools of knowledge in accordance with the scientific method.

The course will address issues relating to research in the fields of Biology and Medicine with the mathematical method and, especially, from the theory of probabilities. This approach will allow the precise quantification of significant relationships between the various phenomena related to health and human pathology from the perspective of Biomedical research.

To achieve these objectives, students will need to work with various conceptual, methodological and instrumental tools necessary to develop a vision of Biomedicine in accordance with scientific rigor.

Competences

- Act with ethical responsibility and respect for fundamental rights and duties, diversity and democratic values.

- Apply knowledge acquired to the planning and implementation of research, development and innovation projects in a biomedical research laboratory, a clinical department laboratory or the biomedical industry.
- Describe biomedical problems in terms of causes, mechanisms and treatments.
- Make changes to methods and processes in the area of knowledge in order to provide innovative responses to society's needs and demands.
- Read and critically analyse original and review papers on biomedical issues and assess and choose the appropriate methodological descriptions for biomedical laboratory research work.
- 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.
- Work as part of a group with members of other professions, understanding their viewpoint and establishing a constructive collaboration.

Learning Outcomes

1. Act with ethical responsibility and respect for fundamental rights and duties, diversity and democratic values.
2. Design, plan and interpret different studies in order to tackle public health problems.
3. Determine the sample size needed to contrast the hypothesis.
4. Distinguish between the different sources of information on health problems.
5. Draw up and contrast hypotheses and identify the errors associated with them.
6. Estimate population parameters from the corresponding sample parameters.
7. Interpret problems and intervention measures in public health.
8. Make changes to methods and processes in the area of knowledge in order to provide innovative responses to society's needs and demands.
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. Understand and critique scientific articles on statistics.
16. Work as part of a group with members of other professions, understanding their viewpoint and establishing a constructive collaboration.

Content

UNIT 1. INTRODUCTION

1.1. Definition and objectives

1.2. Population and sample

1.3. Descriptive statistics, probability theory and inferencial statistics

UNIT 2. MONOVARIANT DESCRIPTIVE STATISTICS

2.1. Quantitative and qualitative variables. Absolute, relative and cumulative frequencies. Graphic representations

2.2. Continuous quantitative variables. Enumerative data: Frecuency tables. Graphic representations. Measures of central tendency: mean, median and mode. Measures of dispersion: range, variance, standard deviation and coefficient of variation. Morphological measures: bias and kurtosis

UNIT 3. BIVARIANT DESCRIPTIVE STATISTICS

3.1. Qualitative relationship between two variables: Contingency tables. Relationship between continuous quantitative and qualitative variables. Relationship between two continuous quantitative variables (correlation coefficient)

3.2. Matching data (repeated measurements)

UNIT 4. PROBABILITY THEORY

4.1. Experiment random sample space and event

4.2. Event operations: union, intersection, difference and contrary events. Incompatible events

4.3. Absolute and relative frequencies. Probability

4.4. Conditional probability. Independent events. Probability of union and intersection of events

4.5. Bayes Theorem

4.6. Measuring the frequency of a disease in the population. Incidence and prevalence

4.7. Evaluation of risk factors. Relative risk and odds ratio

4.8. Evaluation of diagnostic criteria. Sensitivity, specificity, positive and negative predictive values

UNIT 5. RANDOM VARIABLES

5.1. Discrete and continuous random variables

5.2. Probability density function, probability distribution function, expectation and variance of discrete and continuous random variables

5.3. Probability distributions from discrete random variables: Binomial and Poisson

5.4. Probability distributions from continuous random variables: normal, $\chi 2$ and Student's t

5.5. Central Limit Theorem. De Moivre theorem. Sampling distribution. Interval Probability

UNIT 6. ESTIMATION

6.1. Estimation methods: interval confidence. Differences between probability and confidence intervals

6.2. Estimated mean, variance and proportion of population. Determination of the sample size

UNIT 7. HYPOTHESIS TESTING

7.1. Null and alternative hypothesis. Errors type I and type II or α and β risk. One-tailed and two-tailed contrasts. Significance level. Sample Size

7.2. Testing about population mean, population variance and population proportion

7.3. Testing about of differences in mean, variance and proportions. F Fisher Snedecor distribution. Kolmogorov-Smirnov test. Nonparametric comparison of two samples: Mann-Whitney U test

7.4. Hypothesis testing of paired data. Nonparametric Wilcoxon Signed-Rank test

UNIT 8. RELATIONSHIP BETWEEN QUANTITATIVE AND QUALITATIVE VARIABLES: ANALYSIS OF VARIANCE (ANOVA) AND REGRESSION

8.1. One-way ANOVA. Tests *a priori* and *a posteriori*

8.2. Regression: Least squares, significance of the regression and confidence intervals for population parameters. Linearity and utility tests

UNIT 9. RELATIONSHIP BETWEEN TWO RANDOM QUANTITATIVE VARIABLES: CORRELATION

9.1. Correlation Coefficient. Significance of correlation coefficient. Comparison between regression and correlation

UNIT 10. RELATIONSHIP BETWEEN QUALITATIVE VARIABLES: CHI-SQUARE TESTS

10.1. Goodness-of-fit of theoretical distributions frequency distributions

10.2. Homogeneity and independence tests

10.3. McNemar test for paired data

Methodology

Theory lectures:

The lectures will be taught with face-to-face methodology, although the interaction and participation of the students will be made possible and estimulated to the maximum. The classes will be supported by audiovisual media. The material used in class by the teacher, will be available on the Virtual Campus; students are encouraged to print and bring it to class to use as a support when taking notes. The student will also be encouraged to deepen the knowledge acquired in class using the recommended bibliography and simulation software.

Problem classes / Practice seminars:

Given the character and orientation of the subject, the classes of problems, conveniently interspersed with those of theory, will play a key role in its development and in the learning of the subject.

Through the Virtual Campus collections of problems will be delivered, organized according to the topics of the subject, which the student must develop both in class and individually. Most of these problems will be practical cases that, in solving them, allow the student a greater understanding of the knowledge acquired in the theory classes and in personal study.

In the classes of problems, tools such as *Kahoot* will also be used for the consolidation of content and as a diagnosis of the knowledge acquired.

In the practical seminars, conveniently interspersed with the theory classes, the methodology and dynamics of the SPSS software (or other statistical package) will be introduced. The student must use them in practical classes in order to achieve the learning object of the subject.

Practical Classes:

The practical classes are a fundamental point for the correct fulfillment of the objectives of the subject. In them the students will have to solve practical cases, previously selected, by means of statistical software.Learning includes both the introduction and manipulation of data, as well as the use of the main facilities offered by this software for data analysis. The practices will be carried out individually or in pairs. The development of these classes will be linked to the theoretical and problems classes with good temporal correlation.

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	20	0.8	1, 14, 15, 2, 4, 5, 8, 16
Seminars and problems classes	6	0.24	1, 14, 15, 3, 2, 4, 5, 6, 7, 8, 13, 9, 10, 16
Theory lectures	24	0.96	15, 3, 2, 4, 5, 6, 7, 13, 9, 16
Type: Supervised			
Consolidation practices	7	0.28	15, 3, 2, 4, 5, 6, 7, 8, 12, 16
Type: Autonomous			
Personal study	42	1.68	1, 14, 15, 2, 4, 7, 8, 12, 10
Questionnaires of practices	7	0.28	1, 3, 2, 5, 6, 7, 13, 9, 10, 16
Resolution of exercises	24	0.96	15, 3, 2, 5, 6, 7, 13, 9, 10
Tests resolution	10	0.4	

Assessment

Students with CONTINUOUS ASSESSMENT (CA)

The subject will be taught with TE, SEM and PLAB typologies. PLAB type classes are mandatory.

The competences of the subject will be evaluated according to the following criteria:

- 2 multiple-choice exams where errors will be subtracted (with one or more correct answers per question), corresponding to the content of TE and SEM classes, with conceptual questions and problem solving [tests T1 (30%) and T2 (35%)]
- 2 practical exams with computer [tests P1 (10%) and P2 (15%)] and attendance and completion of practices and resolution of practical questionnaires [Qs (10%)]

Theoretical contents:		
1st partial test	T1	30%
2nd partial test	T2	35%
Practical contents:		
1st partial test	P1	10%
2nd partial test	P2	15%
Assistance and completion of practices and resolution of Qs questionnaires	Qs	10%

Students with SINGLE ASSESSMENT (SA)

The subject will be taught with TE, SEM and PLAB typologies. PLAB type classes are mandatory.

To carry out the PLAB classes, you can choose between:

- Do them with the students of continuousevaluation, the same days and hours in which the scheduled practices are carried out.
- Carry them out independently, in computer rooms without regulated hours, or on their personal computer if they have the necessary software.

The Questionnaires of the PLAB will have to be presented before the two tests corresponding to the practical contents (tests P1 and P2, of the 1st and 2nd part of the subject, respectively).

The competences of the subject will be evaluated according to the following criteria:

- 1 multiple-choice exam where errors will be subtracted (with one or more correct answers per question) corresponding to the content of TE and SEM classes, with conceptual questions and problem solving [tests T1 (30%) and T2 (35%)]
- 2 practical exams with computer [tests P1 (10%) and P2 (15%)] and attendance and completion of practices and resolution of practical questionnaires [Qs (10%)]

Theoretical contents:		
Una única prueba con contenido teórico/práctico	T	65%
Practical contents:		
1st partial test	P1	10%
2nd partial test	P2	15%
Completion of practices and resolution of Qs questionnaires	Qs	10%

In the T test, the contents of the entire theory program of the subject corresponding to the material taught in the TE and SEM type classes will be evaluated. It will be done coinciding with the same date set in the calendar for the last continuous assessment theory test (T2), and it will mean 65% of the final grade for the subject.

The P1 and P2 tests will be carried out jointly with the rest of the continuous assessment students.

Ratings:

The minimum global qualification necessary to pass the subject will be 5 points.

- CA students: the minimum score in the theoretical exams to be able to average and pass by continuous assessment will be 3.0 points.
- SA students: the minimum score in the P1 and P2 exams, as well as the average score of the Qs questionnaires, must be 4.0 points.
- In both previous cases, the grade for the course will be a maximum of 4.0 points.

It will be considered that a student will obtain the qualification of "Not evaluable" if the evaluation of all the evaluation activities carried out does not allow him to reach the overall qualification of 5 points in the event that he had obtained the highest grade in all of them.

Recovery/Final Exam:

- There will be a recovery exam, either for those students who have not passed the subject by course, or for those who want to raise their grade (which implies giving up the grade obtained in the previous theoretical exams).
- The recovery/final exam will include all the material, although its result will represent 65% of the final grade since the remaining 35% will continue to depend on the results of the practical part.
- Only students who have been previously evaluated in a set of activities whose weight is equivalent to a minimum of two thirds of the total grade for the subject may take the recovery exam.

Repeating students:

From the second registration, students who have passed the practical contentsin one of the last three courses, may decide between two options:

- Repeat the practical classes and take advantage of the same evaluation scheme as the newly arrived students.
- Be evaluated only for theoretical content. In this case:
  CA students: the percentages of the T1 and T2 tests will be 40% and 60% in the case of partial exams and 100% in the final exam.
  SA students: T test will represent 100% of the grade in any case.

Exam revision:

Following the regulations of the University, the procedure, place, date and time of the review of exams will be announced.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Practical test with computer - 1st partial	10%	2	0.08	1, 14, 15, 3, 5, 6, 13, 12, 11, 9, 10
Practical test with computer - 2nd partial	15%	2	0.08	1, 14, 15, 3, 2, 5, 6, 8, 13, 12, 11, 9, 10
Theoretical and practical questions - 2nd partial	35%	3	0.12	1, 15, 3, 2, 5, 6, 8, 13, 12, 11, 9, 10, 16
Theoretical and practical questions - 1st partial	30%	3	0.12	1, 15, 2, 4, 7, 13, 11, 9, 10, 16

Bibliography

Basic bibliography:

Milton JS. Estadística para biología y ciencias de la salud. 3a. Edición. Madrid: Interamericana. McGraw-Hill, 2007.

Taylor RA, Blair RC. Bioestadística. México: Pearson Education, 2008.

Daniel WW. Bioestadística. Base para el análisis de las ciencias de la salud. 4a Edición. México: Limusa Wiley, 2002.

Sentís J, Pardell H, Cobo E, Canela J. Manual de Bioestadística. 3a. Edición. Barcelona: Masson, 2003.

Sorribas A, Abella F, Gómez X, March J. Metodologia estadística en ciències de la salut: Del disseny de l'estudi a l'anàlisi de resultats. Edicions de la Universitat de Lleida i F.V. Libros. 1997.

Ferrán M, SPSS para Windows. Análisis Estadístico. McGraw-Hill, 2001.

Visauta B. Analisis estadístico con SPSS 14. Estadística básica. 3a Edición. McGraw-Hill, 2007.

Martínez-González MA, Sánchez-Villegas A, Toledo E, Faulin FJ. Bioestadística amigable. 4a. Edición. Elsevier. 2020

Web links:

https://www.ibm.com/docs/SSLVMB_27.0.0/pdf/es/IBM_SPSS_Statistics_Brief_Guide.pdf

http://www.hrc.es/bioest/M_docente.html

https://link-springer-com.are.uab.cat/book/10.1007%2F978-3-319-20600-4

http://davidmlane.com/hyperstat/index.html

https://seeing-theory.brown.edu

http://vassarstats.net

http://Statdistributions.com/

Simulators:

http://demonstrations.wolfram.com/ - http://demonstrations.wolfram.com/topic.html?topic=Statistics&limit=20

http://socr.ucla.edu/SOCR.html

Software

In the practical classes, the statistical program IBM SPSS or an equivalent one will be used.