Universitat Autònoma de Barcelona

2020/2021

Biostatistics and Data Analysis

Code: 101917
ECTS Credits: 6

Degree	Type	Year	Semester
2501230 Biomedical Sciences	FB	1	2

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: Mercedes Campillo Grau
Email: Mercedes.Campillo@uab.cat

Use of Languages

Principal working language: spanish (spa)
Some groups entirely in English: No
Some groups entirely in Catalan: No
Some groups entirely in Spanish: No

Teachers

Leonardo Pardo Carrasco

Prerequisites

There are no official prerequisites, however it is recommended to have previous knowledge of elementary mathematics that includes the concepts of derivation and integration.

Objectives and Contextualisation

Biostatistics and Data Analysis aims to introduce students to the fundamental knowledge and use of basic tools of knowledge according to the scientific method.

The course will address issues relating to research in the fields of biology and medicine using mathematical method and especially from probability theory. This approach will quantify accurately, significant relationships between different phenomena related to human health and disease from the perspective of biomedical research.

To achieve these objectives, the student must work with different conceptual and methodological tools, instruments needed to develop a vision of biomedicine according to 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

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

4.4. Bayes Theorem

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

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

4.7. 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, $\chi 2$, Student's t and Fisher Snedecor F

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. 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 magistral methodology, trying to encourage maximum interaction and student participation. The classes will be supported by audiovisual media. The material used in class by the teacher will be available on the Virtual Campus. It is recommended to print and bring to class this material for use as support when taking notes. Students will be encouraged to deepen the knowledge acquired in class using bibliography and the simulation software recommended.

Problem classes / seminars:

Given the character and orientation of the subject, the classes of problems, appropriately put into place with those of theory, will play a key role in its development and in the learning of the subject.

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

In the problem classes, tools such as *Kahoot* will also be used to consolidate content and to diagnose the knowledge acquired.

In the practice seminars, appropriately interleaved with 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 to achieve the learning of the subject.

Practical Classes:

The practical classes are essential for the proper fulfillment of the objectives of the course. In them, students will solve practical cases previously selected, using statistical software. Learning includes both the introduction and manipulation of data, using the facilities offered by the software for data analysis. The practices will be

conducted individually or in pairs. The development of these classes will be linked to the theoretical and problems classes with good temporal correlation.

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

Competences will be evaluated using the following criteria:

- Multiple-choice tests (with one or more correct answers per question) conceptual questions and troubleshooting [T1 tests (30%) and T2 (35%)]
- Practical tests with computer [P1 tests (10%) and P2 (15%)] and,
- Attendance and reporting practices (10%)

Theoretical tests:		
1st partial test	T1	30%
2nd partial test	T2	35%
Practical tests:		
1st partial test	P1	10%
2nd partial test	P2	15%
Attendance and reporting practices		10%

Scores:

- The overall minimum score required to pass the course by continuous assessment will be 5 points.
- To average and pass the course by continuous assessment, the minimum score in the theory exams will be 3,0 points.

- If students do not pass the subject for continuous assessment, since they do not reach a minimum of 3,0 points in any of the theory exams, the course mark will be 4 points maximum.
- A student will be considered "Non-evaluable" if the marks of all evaluation activities not let to achieve overall rating of 5, in case of having obtained the highest score in all of them.

Recovery Exam (Final):

- There will be a recovery exam, either for those students who have not passed the subject by continuous evaluation, or for those who wish to raise the grade (which implies renouncing the marks obtained in the two theoretical exams by continuous evaluation).
- Only students who have been previously evaluated in a set of activities, the weight of which equals a minimum of two thirds of the total grade of the subject, may be undergo the recovery exam.
- The recovery exam will include all the subject, although its result will represent 65% of the final grade, as the remaining 35% will continue depending on the results of the practical part.

Repeating students:

- From the second enrollment, the students can decide whether or not, repeat the practical classes or only conduct the theoretical exams.
- In the latter case the percentage of the tests will be 40% and 60% (for 1st and 2nd partials, respectively) in case of continuous assessment and 100% for the final exam.

Exams revisions:

Following the University regulations, the procedure, the place, the date and time of the exams revision 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, 2001.

Daniel WW. Bioestadística. Base para el análisis de las ciencias de la salud. 4a Edición. 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. Programación y Análisis estadístico. McGraw-Hill, 1996.

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

Web links:

http://www.bioestadistica.uma.es/libro/

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