

Biostatistics and Data Analysis

Code: 101917
ECTS Credits: 6

| Degree | Type | Year | Semester |
|-----------------------------|------|------|----------|
| 2501230 Biomedical Sciences | FB | 1 | 2 |

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
Ramón Guixa González
Arnau Cordomi Montoya

Prerequisites

No official prerequisites, but it is recommended that the students have previous knowledge of elementary mathematics concepts including 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.

Skills

- 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.
- Develop critical thinking and reasoning and communicate ideas effectively, both in the mother tongue and in other languages.
- Develop independent learning habits and motivation to continue training at postgraduate level.
- Develop independent learning strategies.

- Develop scientific knowledge, critical reasoning and creativity.
- Generate innovative and competitive proposals for research and professional activities.
- Identify and understand the advances and challenges of research.
- Read and critically analyse original and review papers on biomedical issues and assess and choose the appropriate methodological descriptions for biomedical laboratory research work.
- Show respect for the ethical and legal aspects of research and professional activities.
- Work as part of a group with members of other professions, understanding their viewpoint and establishing a constructive collaboration.

Learning outcomes

1. Design, plan and interpret different studies in order to tackle public health problems.
2. Determine the sample size needed to contrast the hypothesis.
3. Develop critical thinking and reasoning and communicate ideas effectively, both in the mother tongue and in other languages.
4. Develop independent learning habits and motivation to continue training at postgraduate level.
5. Develop independent learning strategies.
6. Develop scientific knowledge, critical reasoning and creativity.
7. Distinguish between the different sources of information on health problems.
8. Draw up and contrast hypotheses and identify the errors associated with them.
9. Estimate population parameters from the corresponding sample parameters.
10. Generate innovative and competitive proposals for research and professional activities.
11. Identify and understand the advances and challenges of research.
12. Interpret problems and intervention measures in public health.
13. Show respect for the ethical and legal aspects of research and professional activities.
14. Understand and critique scientific articles on statistics.
15. 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 inferential 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: Frequency 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, χ^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 QUALITATIVE VARIABLES: CHI-SQUARE TESTS

- 8.1. Goodness-of-fit of theoretical distributions frequency distributions
- 8.2. Homogeneity and independence tests
- 8.3. McNemar test for paired data

UNIT 9. RELATIONSHIP BETWEEN QUANTITATIVE AND QUALITATIVE VARIABLES: ANALYSIS OF VARIANCE (ANOVA) AND REGRESSION

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

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

UNIT 10. RELATIONSHIP BETWEEN TWO RANDOM QUANTITATIVE VARIABLES: CORRELATION

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

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 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, 7, 8, 15 |
| Seminars and problems classes | 6 | 0.24 | 13, 14, 6, 3, 2, 1, 7, 8, 9, 11, 12 |
| Theory lectures | 24 | 0.96 | 13, 14, 6, 3, 2, 1, 7, 8, 9, 10, 11, 12, 15 |
| Type: Supervised | | | |
| Consolidation practices | 7 | 0.28 | 6, 3, 1, 7, 8, 12, 15 |

Type: Autonomous

| | | | |
|-----------------------------|----|------|---------------------------|
| Personal study | 42 | 1.68 | 14, 6, 5, 4, 1, 7, 11, 12 |
| Questionnaires of practices | 7 | 0.28 | 6, 5, 4, 1, 8, 12 |
| Resolution of exercises | 24 | 0.96 | 14, 6, 5, 2, 1, 8, 9, 12 |
| Tests resolution | 10 | 0.4 | 6, 5, 4 |

Evaluation

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 in computer rooms [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% |

- 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 points**.
- Failure to pass the course by continuous assessment, because not to have achieved a minimum of 3 points in any of the theoretical exams, imply a global score of 4 points.
- There will be a **final exam** –of recovery for those students who have not passed the subject by continuous assessment, or for those who wish to increase rating (**which involves giving up the mark obtained in theoretical test by continuous evaluation**– at which can be presented to those students who have been previously evaluated in a series of activities whose weight equals to **a minimum of two thirds** of the total grade of the subject. The final 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 rating.
- From the second enrollment, 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.
- A student will be considered "Not assessable" if the value of all evaluation activities not let to achieve overall rating of 5, in case of having obtained the highest score in all of them.
- Following the University regulations, the procedure, the place, date and time of the exams revision will be announced.

Evaluation activities

| Title | Weighting | Hours | ECTS | Learning outcomes |
|-------|-----------|-------|------|-------------------|
|-------|-----------|-------|------|-------------------|

| | | | | |
|---|-----|---|------|----------------------------------|
| Practical test in computer rooms - 1st partial | 10% | 2 | 0.08 | 13, 6, 5, 4, 1, 7, 12 |
| Practical test in computer rooms - 2nd partial | 15% | 2 | 0.08 | 13, 14, 6, 5, 4, 3, 2, 1, 8, 9 |
| Theoretical and practical questions - 2nd partial | 35% | 3 | 0.12 | 14, 6, 3, 2, 1, 8, 9, 10, 11, 15 |
| Theoretical and practical questions - 1st partial | 30% | 3 | 0.12 | 14, 6, 3, 1, 7, 10, 11, 12, 15 |

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.bioestadística.uma.es/libro/>

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

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

Simulators:

<http://web.udl.es/usuarios/q3695988/wenessim/Pagines/index.htm>

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

<http://www.socr.ucla.edu/Applets.dir/>