Universitat de Lleida

DEGREE CURRICULUM

DATA ANALYSIS IN HEALTH RESEARCH 1

Coordination: GOMEZ ARBONES, XAVIER

Academic year 2021-22

Subject's general information

Subject name	DATA ANALYSIS IN HEALTH RESEARCH 1
Code	14064
Semester	1st Q(SEMESTER) CONTINUED EVALUATION

Typology	Degree	Course	Character	Modality
	Master's Degree in Health Research	2	COMPULSORY	Blended learning

Course number of credits (ECTS)	6

Type of activity, credits, and groups	Activity type	PRAULA	TEORIA
	Number of credits	2.4	3.6
	Number of groups	1	1

Coordination	GOMEZ ARBONES, XAVIER
Department	MEDICINE
Teaching load distribution between lectures and independent student work	Classroom lectures: 12 hours. This subject has been planned as a semi-face-to-face subject, although there are days with face-to-face sessions. If finally the face-to-face sessions could not be carried out due to COVID-19 or other circumstances, the sessions they will be made through synchronous videoconference through the CV, or will be made available to students in video format / presentation recorded on the CV. The rest of the hours are for autonomous and tutoriset students work
Important information on data processing	Consult this link for more information.
Language	Catalan, Spanish and some material in English

Teaching staff	E-mail addresses	Credits taught by teacher	Office and hour of attention
GOMEZ ARBONES, XAVIER	xavier.gomez@udl.cat	6	
RUBINAT ARNALDO, ESTHER	esther.rubinat@udl.cat	0	

Subject's extra information

In this subject, as a complement to the subject of research study design, we will deal with the analysis of the data obtained during the research project, and also aspects related to the definition of variables, obtaining registration of the data and the plannification of the statistical procedures that must appear in the project.

The course begins with an introduction to the basis of **probability**, which support the descriptive and analytical statistical procedures of inferential statistics. We will approach the topics of sampling and sample size calculation

Then, we continue on **descriptive statistical concepts**, since they are the first step of any statistical analysis .

For statistical procedures we are going to use the **PSPP** program. GNU PSPP is a program for statistical analysis of sampled data. It is a free as in freedom replacement for the proprietary program SPSS, and appears very similar to it with a few exceptions. The most important of these exceptions are, that there are no "time bombs"; your copy of PSPP will not "expire" or deliberately stop working in the future. Neither are there any artificial limits on the number of cases or variables which you can use. There are no additional packages to purchase in order to get "advanced" functions; all functionality that PSPP currently supports is in the core package. PSPP is a stable and reliable application. It can perform descriptive statistics, T-tests, anova, linear and logistic regression, measures of association, cluster analysis, reliability and factor analysis, non-parametric tests and more. Its backend is designed to perform its analyses as fast as possible, regardless of the size of the input data. You can use PSPP with its graphical interface or the more traditional syntax.

The following topics are addressed to concepts related to **analytical statistics or contrast ypothesis testing**. The most important statistical procedures are presented. Without going into depth, more advanced procedures are also exposed, such as survival analysis, multivariate procedures, ROC curves. To do this, the PSPP program is used, also some SPSS commands are presented (The UdL has a campus license for SPSS and is available in the university's computer rooms)

Finally, we deal with the **interpretation of results** of an statistically or clinically significant point of view; as well as the procedures and tools that can help us establish relevant conclusions .

The duration of the subject does not allow you to achieve maximum statistical expertise. The subject is eminently practical and applied, and if it is intended that by completing it, the student will beable to write the statistical procedures of a project, perform a basic statistical analysis and interpret the results from a statistical and clinical point of view.

Learning objectives

1. Knowledge of the basics of the design, execution, analysis and interpretation of research studies.
2. Be able to design an appropriate database to a research study and fill it with the results. Validate the results and the database.
3. Deal with the concept of probability and biological variability. Recognize in biostatistics and epidemiology a tool for working with samples.

4. Know and differentiate the statistical the descriptive and analytical methods used in biomedicine
5. P and reflect on the concept of power.
6. Interpret the results of statistical tests and be able to evaluate the adequacy of the statistical methods.
7. Kwon the basics of using statistical software.
8. Develop a computerized database and be able to analyze data to obtain conclusions
9. Transform and operate with variables.
10. Apply statistical methods to describe and compare.
11. Basic knowledge about evaluation of diagnostic tests, multivariate analysis and survival analysis

Competences

Transversal skills:

- Ability to use instruments of articles critical assessment of qualitative and quantitative research.
- Knowledge to apply scientific writing language in communicating health outcomes
- Ability to describe and apply the most common techniques for exploring and analyzing data, analyze relationships between variables and test hypothesis in quantitative and qualitative research.

Subject contents

1. Introduction to descriptive statistics.
2. Descriptive statistics. of centralization, dispersion and position.
3. Introduction to probabilidad, confidence intervals and inferential statistics.
4. Contrast of independence between two qualitative variables: Chi-Square.
5. Contrast of normality of a numerical variable.
6. Contrast of a numerical variable observed in two samples: t-test, Mann-Whitney.
7. Contrast of a numerical variable in three or more samples: ANOVA and Kruskal-Wallis.
8. Correlation and linear regression.
9. Result interpretation. Confidece interval.

Methodology

The teaching methodology is based in lectures-seminars (video) and guided activities with the participation of students. The classes will be supported with iconography (slides and transparencies) and computer presentations.

We recommend to actively participate in all the sessions and work in advanced the contents of the lectures.

The learning activity is managed through the Virtual Campus (CV) of the University of Lleida.

The CV provides students with basic and additional material for students. Activities, tests and forums are proposed. Tutorials are done by video conference.

It is possible that, given the evolution of COVID-19, the methodology of the subject will have to be adapted or modified. This subject has already been planned as a semi-face-to-face subject, although there are two days with face-to-face sessions. If finally the face-to-face sessions could not be carried out due to COVID-19 or other circumstances, the sessions they will be done through synchronous videoconferencing on the CV, or will be made available to students in video format / presentation recorded on the CV.

Information on the transmission and recording of personal data of teachers and students of the University of Lleida as a result of teaching in UdL facilities and at a distance in the Catalan or Spanish guide.

Automàtic translation:Information on the transmission and recording of personal data of teachers and students of the University of Lleida as a result of teaching in UdL facilities and at a distance
The University of Lleida informs that, depending on the changes to which it is obliged in accordance with the instructions of the health authorities, the provisions of the UdL or the assurance of the quality of teaching, it may transmit, record and use the image, voice or,

where appropriate, the physical environment chosen by teachers and students, with the aim of teaching in UdL facilities or at a distance. In turn, it encourages the people affected so that, in the case of distance teaching, they choose the spaces that have the least impact on their privacy.

And, in general, it is recommended to opt preferably for interactions in the chat or without activating the camera, when teaching activities are not carried out that, due to their characteristics, require an oral or visual interaction.

The person responsible for the registration and use of this personal data is the University of Lleida –UdL– (contact details of the representative: General Secretariat. Plaça de Víctor Siurana, 1, 25003 Lleida; sg@udl.cat; data contact details of the data protection delegate: dpd@udl.cat).

This personal data will be used exclusively for the purposes inherent in the teaching of the subject. In particular, the recording fulfills the following functions:

• Provide access to online content and, where appropriate, asynchronous training.

• Guarantee access to content for students who, due to technological, personal or health reasons, among others, have not been able to participate.

• Constitute a study material for the preparation of the evaluation.

It is absolutely forbidden to use the data transmitted and the recordings for other purposes, or in areas outside the Virtual Campus, where they will remain archived, in accordance with the intellectual and industrial property policy of all content included on proprietary websites. of the UdL.

If there are any, the records will be kept for the time decided by the teacher, in accordance with strictly academic criteria, and, in all likelihood, must be removed at the end of the current academic year, in the terms and conditions provided for in the regulations on the conservation and disposal of the UdL's administrative documents, and the document evaluation tables approved by the Generalitat de Catalunya (http://www.udl.cat/ca/serveis/arxiu/) .This personal data is essential for teaching in the subject, and the definition of teaching procedures, especially that made at a distance, is a power of the UdL within the framework of its right to university autonomy, as provided for in Article 1.1 and Article 33.1 of Organic Law 6/2001, of 21 December, on Universities. For this reason, the UdL does not need the consent of the people affected to transmit or record their voice, image and, where applicable, the physical environment they have chosen, for this sole purpose, to teach. teaching in the subject.

The UdL will not transfer the data to third parties, except in the cases strictly provided for in the Law.

Affected people can access their data; request rectification, deletion or portability; oppose the treatment and request its limitation, provided that it is compatible with the purposes of teaching, by writing to the address dpd@udl.cat. They can also submit a complaint to the Catalan Data Protection Authority, through the Authority's electronic office (https://seu.apd.cat) or by non-electronic means.

Development plan

Face-to-face sessions:

See CV information, please.

Evaluation

For the evaluation of the subject, the UdL evaluation regulations are followed. Available at http://www.udl.cat/ca/udl/norma/ordenaci-/.

The evidences that make up the assessment of the subject are work and tests related to the contents and competencies of the subject; problem solving or cases; activities related to the contents and competencies of the subject; and written exams.

The language of the evaluation activities are Catalan, Spanish and English. The student can write the answer, if applicable, in any of the official languages of the University

The assessment is continuous and takes place within the teaching period delimited for the subject or subject, in accordance with the academic calendar of the course approved by the Governing Council.

The activities and tests will be carried out preferably through the tools of the UdL CV, therefore it is essential to have a computer and access to the CV during the academic year. To do the activities you must run statistical software either with your own computer or by the computers available at the UdL.

The final exams are scheduled online and the conditions and format of the test are published a few days before (number of questions, penalty for incorrect answers, …).

Students who require or plan to require adaptations in the assessment tests must contact the teacher responsible

for the subject during the first 15 days from the beginning of the course to assess their situation.

The final weight of the evaluation activities is:

- **Attendance** at face-to-face sessions (or videoconference, if applicable): 30%. To pass the course you must attend a minimum of 2/3 of the sessions. Exceptional situations of non-attendance must be discussed with the teacher at the beginning of the course.

- **Tests, activities, forum, individual or group work** on the contents of the sessions: 30%. To pass the subject you must get a grade higher than 50% of the possible grade in each activity.

- Active **participation** in the sessions and in the development of the subject: 5%

- **Final exam**: 35%. To pass the course you must get a grade higher than 50% of the possible grade. This activity could be recovered.

Assessment activities or assessment results may be taken into account, which may serve to modulate grade in specific situations.

The subject is passed if the final grade taking into account all the evaluative evidence is higher than 5/10.

Rating scale:

0.00-4.99: suspended,

5.00-6.99: approved,

7.00-8.99: notable,

9.00-10.00: excellent.

The assessment system is discussed with students enrolled in the first days. No alternative evaluation is previse.

Given the situation due to COVID-19, if applicable, the evaluation system may be subject to modifications, which will be discussed and agreed with the students, as well as with the direction of the master's degree.

Bibliography

Fundamental text books:

1. Abella F, Fajó M, Gómez X, March J, Sorribas A. Metodología estadística en ciencias de la salud. Del diseño del estudio al análisis de los resultados. Edicions de la UdL y F.V. Libros Eines 26, 2001.
2. Argimón Pallás JM, Jiménez Villa J. Métodos de investigación aplicados a la atención primaria. Mosby/Doyma. Barcelona, 1994.
3. Armitage P, Berry G. Estadística para la investigación biomédica. Ediciones Doyma. Barcelona, 1992.
4. Hulley SB, Cummings SR. Diseño de la investigación clínica. Un enfoque epidemiológico. Ediciones Doyma. Barcelona, 1993.
5. Woolson RF. Statistical methods for the analysis of biomedical data. John Wiley & Sons. Chichester, 1987.