

A practical approach to solving statistical data analysis problems – resolving examples and tasks

Faculty

Szkoły Doktorskie

Specialization

Międzynarodowa Szkoła Doktorska

Kind

Stacjonarne

Type

Szkoła Doktorska

Term

2nd semester

Subjects ID

Course supervisor

Subjects name:

A practical approach to solving statistical data analysis problems – resolving examples and tasks

Unit

Department of Haemostasis and Haemostatic Disorder, Chair of Biomedical Sciences

Course profile

academic

Speciality

not applicable

Level of course unit

PhD students

Course aims

The course aims to discuss the various stages of collecting and processing experimental data. The program covers the most basic statistical analyses, as well as selected specialized analyses, and prepares students for practical assessment of the reliability of collected data and analyses' outcomes. After short, block-based introductions to a specific topic, students will have the opportunity to conduct analyses using STATITICA software. Also, other internet-based tools will be used for demonstration. Students receive ready-made data sets for analysis and can also perform analyses using their own data, brought to class on USB drives. For each statistical analysis method discussed in class, practical examples will be performed, demonstrating the STATISTICA software path required for the given analysis. Assessment of the course will consist of independent data analysis under the instructor's guidance, a description of the analysis performed, and a discussion of the results. As part of the assessment, students will report on their achievements, concerns, and questions.

Types of educational activities

Subject	Language course	Selfstudy	Lecture	Excercises	Laboratory	Seminars	practical	e-learning	Professional practice	Other (what?)	ECTS (pts)
	ENG	15	2	18							

ECTS Credits

not applicable

Name of course unit's faculty

prof. dr hab. Cezary Watała

Prerequisites

Obtaining a Master of Science Degree or the equivalent is mandatory. Prerequisites: a PhD student who will attend this course does not need to have knowledge of advanced statistical sciences. Classes are introductory in the basics of statistics necessary to understand the content of introductory lectures, excercises and individual topics are discussed from scratch. Classes are practical/theoretical (10:1) and run on the basis of analyzes performed on computers. Some of the issues are demonstrative.

Learning activities and teaching methods

Lecture: oral presentation and discussion.

Practical classes: oral presentation preparing students for practical classes; problem solving with the use of a computers and dedicated software, discussion of the obtained results.

Course unit content:

A practical approach to solving statistical data analysis problems – resolving examples and tasks

Meeting 1 – Schemes of the choice amongst statistical procedure

- How do we select the right tests for the analyzed data? What do we need to know? What assumptions must be met?
- How do we select dependent, independent, continuous, and discrete variables? What does our choice depend on?
- What determines the data distribution and why is it so important?
- Can we modify the distribution of measurement data without manipulating the data?

Organizing data and preparing it for further analysis

- What to do with data that fails to meet test assumptions? Which assumptions should be considered the most critical for different methods?
- How to deal with data that fails to meet assumptions, i.e., outliers.
- Can the problem of outliers be solved, and how? Methods for transforming measurement data and resampling. The philosophy behind outlier removal.
- Outliers and missing data: consequences of missing data for analysis; possible solutions.

Meeting 2 – Your first step – what is the purpose of your research?

- How to formulate research hypotheses and statistical hypotheses?
- What conditions must a properly formulated statistical hypothesis meet?
- Where did all this come from? A bit of history of metaphysical research
- When does a scientific hypothesis make sense?
- What is statistical hypothesis parity and when is it not met?
- Verifying the consistency of the experimental and theoretical models: general verification of the assumptions of the working hypothesis
- Working hypothesis, research question, purpose of the study: what do we want to prove or what question(s) do we want to answer?
- Sample size estimation: is it necessary to create groups of sufficiently large size?
- When do excessively large groups pose a problem?
- Conformity of the research model to the statistical model for the results

- Strategies for Building Statistical Hypothesis Pairs: Reject-Support and Accept-Support – When to Use Which?
- A fundamental step in experimental design – assessing the necessary sample size and satisfactory statistical power – whether a priori or a posteriori.
- A brief summary: what steps lead us to properly planning a biomedical experiment?
- When the outline is the final touch... how to deal with it and is it worth it?

Meeting 3 – How to demonstrate differences between compared groups?

- What are inferential tests for?
- Simple and more complex comparisons – from two-sample comparison tests to various ANOVA models and post hoc tests
- Blocked experiments and block analysis – a neglected problem of fundamental importance
- Multivariate analyses – when to use block analysis and when to use analysis of covariance?
- Nonparametric equivalents of analysis of variance – when to use them? Are there alternatives?
- When can inferences about differences be misleading? The jackknife technique and the bootstrap technique

Meeting 4 – Relationship Between Variables – correlation or regression – maybe so, maybe so, which is better I don't know ...

- Correlation and Regression – What Is Their Identity, and How Do They Differ?
- Simple and More Complex Correlation and Regression Models
- Simple and Multiple Parametric Regression – Selecting Dependent, Independent, and Dummy Variables
- Simple and Partial Correlation – Redundancy or Tolerance – What's More Important?
- A More Advanced Nonparametric Model – Logistic Regression
- Simple and Multi-Term Frequency Tables – What is the Yates Correction and When to Use It?
- When to Use Frequency Tables and When to Use Exact Tests?

Meeting 5 – What we missed earlier, but might be worth mentioning...

- Randomization models used in biomedical research; practical tools and methods
- Advantages and limitations of multiparameter data analysis used in medical research
- Examples of tempting multiparameter analyses:
 - Linear discriminant analysis
 - Canonical analysis

And finally – some miscellaneous facts ... namely:

- How to present statistical analysis results in various types of medical publications? *
- Graphical presentation methods for data analysis in publications and presentations – how to create an informative scientific graph? What software should you use? What should you pay attention to?
- How to compile tables? What should you include in them? How should you format headers? What should the table legend include?
- A short report on basic errors in biomedical research publications

Knowledge

After accomplishing the course, a student has a knowledge of and understands:

Methodology of research, principles of presenting and disseminating results, including open access

Proficiency in using the Statistica program in performing basic operations on measurement data in this program

Knowledge in selecting the most appropriate and useful methods of statistical data analysis.

Knowing where to seek further help in case of unconventional problems and queries,

(P8S_WG)

Skills

After accomplishing the course, a student has skills allowing:

using a knowledge from various scientific disciplines to creative identification and solving complex problems and accomplishing research tasks,

defining an aim of the study and research hypothesis,

developing and using experimental methods and techniques,

drawing conclusions on the basis of obtained results

performing critical analysis of results

planning and performing individual and joint research projects, including international projects

(P8S UO)

Competencies

After accomplishing the course, a student is ready to:

- **critical assessment of achievements in a given scientific discipline**
- **critical assessment of his/her own contribution in a development of a given scientific discipline**
- **be aware of a significance of a knowledge in solving basic and practical problems (P8S_KK)**

Required learning resources

Stanisz A. Przystępny kurs statystyki, tomy 1, 2 i 3. StatSoft Polska, Kraków 2006-2007, ISBN 9788388724183.

Recommended learning resources

Zar J. Biostatistical Analysis, Pearson Education Limited, Edinburgh Gate, Harlow, Essex CM20 2JE, England and Associated Companies throughout the world, ISBN 10: 1-292-02404-6, ISBN 13: 978-1-292-02404-2, wyd V lub wcześniejsze

Armitage P., Berry G. Statistical Methods in Medical Research, Wiley-Blackwell, SBN 9780632052578

Stanisz A. Modele regresji logistycznej, Statsoft Polska, Kraków 2016, ISBN: 9788388724732.

Thomas Hill, Pawel Lewicki. Statistics: Methods and Applications, Statsoft Inc., Tulsa 2006, ISBN, 1884233597, 9781884233593

Bręborowicz, G.H. Moczko J.A. Nie samą biostatystyką ... Uniwersytet Przyrodniczy w Poznaniu (Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu), 2014/2016, ISBN: 9788377120125

Francuz P., Mackiewicz R. Liczby nie wiedzą skąd pochodzą, Katolicki Uniw. Lubelski (KUL), Lublin 2007, ISBN 978-83-7363-567-8

Assessment methods

After accomplishing the course, students will be asked to prepare an oral presentation including problem solving with the use of a computers and dedicated software, discussion of the obtained results

Rules for making up absences from classes:

Additional information

Consultations for PhD students: Zakład Zaburzeń Krzepnięcia Krwi, room 153, ul. Mazowiecka 6/8, 92-215 Łódź, prof. dr hab. Cezary Watała. It is necessary to establish a date and an hour of consultation via email (cezary.watala@umed.lodz.pl).

Didactic supervisor of the Department:

dr hab. Jacek Golański

jacek.golanski@umed.lodz.pl

tel. 422725731

Signature

I hereby state that the content of the curriculum included in the syllabus below is the result of my individual work completed as part of work contract/cooperation resulting from a civil law contract, and that author rights to this title are not the property of a third party.

Course language

English/Polish