COURSE DESCRIPTION CARD - SYLLABUS

Course name
Statistics and data analysis [S1Inf1>SAD]

Course

Field of study
Computing

Year/Semester
2/4

Area of study (specialization)
–

Profile of study
general academic

Level of study
first-cycle

Course offered in
Polish

Form of study
full-time

Requirements
compulsory

Number of hours

Lecture
32

Laboratory classes
24

Other
0

Tutorials
0

Projects/seminars
0

Number of credit points

4,00

Coordinators

dr hab. inż. Katarzyna Filipiak prof. PP
katarzyna.filipiak@put.poznan.pl

Lecturers

Prerequisites

Students should have knowledge concerning mathemathical analysis, matrix analysis, and probability theory.

Course objective

The aim of the course is to provide students with the basic knowledge of statistical inference methods and the use of the R package. The theoretical knowledge acquired is intended to develop the ability to apply it in practice in solving engineering problems.

Course-related learning outcomes

Knowledge:
Students should have:
1. well founded knowledge on statistics useful for understanding fields of computer science and other engineering disciplines (K1st_W4-W5)
2. knows basic methods of descriptive statistics, data presentations, statistical inference constructing histograms, correlation and regression, parametric and non-parametric tests (z-test, t-Student test, F-test, chi-test) which are particularly useful to solve problems within computer science, in particular analysis

algorithms, computer systems and artificial intelligence (K1st_W7)

Skills:
1.is able to calculate basic statistical measures (K1st_U3)
2. is able to use descriptive statistics measures and hypothesis testing to solve simple experimental tasks of computer science (K2st_U4)
3. can verify statistical hypotheses for computer sciences and related ones (K1st_U4)
4. is able to analyse simple problems of data analysis, choose an appropriate statistical method, present and interpret results of the statistical analysis
Social competencies
1. understands that knowledge and skills related to computer science and data mining quickly becomes non relevant and it is necessary to self-learn and extend his/her knowledge and skills (K1st_K1-2)

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Lectures - written test based on the material presented during the lectures; the student has to collect at least 50% of possible points;
Laboratory classes - two tests comprising probability theory and mathematical statistics; the student has to collect at least 50% of possible points from each test; the first test is scheduled for the eighth class, and the second one for the last class in the semester

Programme content

1. Elements of descriptive statistics
2. Random variables
3. Sampling distributions
4. Statistical inference related to one and many populations
5. Correlation and regression analysis
6. Goodness-of-fit and independence tests: chi-square test
7. Nonparametric tests

Course topics

1. Elements of descriptive statistics
2. Random variables - probability distribution, cumulative distribtuion, expectation, standard deviation, binomial distribution, exponential distribution, normal distribtuion
3. Sampling distributions - Student t-distribution, chi-square distribution
4. Statistical inference: point and interval estimation for single population, hypothesis testing, comparison of two populations, one- and two-ways analysis of variance (ANOVA)
5. Correlation and regression analysis - Pearson correlation coefficient, linear regression, test for significance of regression, curvilinear and multiple regression
6. Goodness-of-fit chi-square tests - for frequency distribution, normality and for independence of variables (contingency tables)
7. Nonparametric tests - Wilcoxon sum rank test and Mann-Whitney test, Wilcoxon signed rank test, Kruskal-Wallis test, Friedman test, Spearman correlation coefficient, Spearman test of independence

Teaching methods

Lectures (multimedia presentations) - presenting the theory connected with a current students' knowledge, presenting a new topic preceded by a reminder of related content known to students from other subjects; solving simple examples with the use of R package.

Laboratory classes - solving examples with the use of R package, discussions in groups, applications of statistical methods to solve real problems in groups an individually.

Bibliography

Basic
1. Krysicki, W., J. Bartos, W. Dyczka, K. Królikowska and M. Wasilewski: Rachunek prawdopodobieństwa i

statystyka matematyczna w zadaniach, wydanie 8. PWN Warszawa, 2012
2. Bobrowski, D. and K. Maćkowiak-Łybacka: Wybrane metody wnioskowania statystycznego. Wyd. PP, Poznań, 2004

Additional
1. Devore, J.L.: Probability and Statistics for Engineering and Sciences, Brooks/Cole, 2012
2. Ross, S.M.: Introductory Statistics, Elsevier, 2010

Breakdown of average student's workload

	Hours	ECTS
Total workload	100	4,00
Classes requiring direct contact with the teacher	58	2,00
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	42	2,00