MINISTRY OF EDUCATION AND SCIENCE OF RUSSIAN FEDERATION

National Research Tomsk State University
Department of Applied Mathematics and Cybernetics

APPROVED:

Dean of the Dept., Professor

A.M. Gortsev

December, the 10th, 2014.

Course syllabus

SOFTWARE APPLICATIONS FOR STATISTICAL AND PROBABILITY ANALYSIS

Master's program title

Statistical and probability analysis of queuing systems

Major study:

01.04.02 – Applied mathematics and computer science

Graduate qualification (degree)

Master

Tomsk
2014

1. Course objective

Course objective is to introduce students to essential modern software applications, supporting statistic methods applied for data analysis in economics, medicine, sociology and other areas. Course aims at teaching students to solve statistical analysis problems using modern software and developing operating skills for essential software modules.

2. Course position in Mater's program

"Software applications in statistical and probability analysis" course introduces students to modern software applications allowing to solve probability analysis tasks and perform statistical data processing, provides knowledge and skills necessary to work with a variety of software applications in the field, and naturally complements the "Applied statistical analysis" course.

3. Student's competencies, obtained as a learning outcome

Together with other courses in basis part of Master's program, the "Software applications in statistical and probability analysis" course provides a framework to obtain the following competencies (GC – general competences; PC – professional competences):

- Competence to use in-depth theoretical and practical knowledge in applied mathematics and informatics (GC-3);
- Competence to acquire and apply new knowledge and skills individually using information technologies, including knowledge and skills in a new area of expertise, not related directly to the scope of activity, competence to extend scientific world outlook (GC-4);
- Competence and commitment to active collaboration in scientific, production and social areas of activity (GC-7);
- Competence to make free use of russian and foreign languages for business communication, competence to active social mobility (GC-8);
- Competence to carry out scientific research and acquire new scientific and applied results (PC-1);
- Competence to develop conceptual and theoretical models of scientific problems (PC-2);
- Competence to perform in-depth problems analysis, to set and validate scientific, design and engineering problems (PC-3);
- Competence to project management, research planning, risk management, project team leading (PC-5);
- Competence to participate in international projects in major specialty (PC-11)

Course learning outcomes for students are:

- Knowledge – essential statistical methods and applications supporting those methods implementation;
- Skills – to be conversant in the variety of software applications, to determine the best fitting application for particular problem, to solve a broad set of problems of applied statistical analysis with a specific application.

4. Topics and abstracts

Topic 1. Current situation. Review of software applications on the market.

Review of software applications, allowing to solve applied statistical analysis problems. General capabilities and differences for Excel, Mathcad, Statistica and R applications. General review of other applications (Statgraphics, SPSS, etc.).

Topic 2. Working with Excel spreadsheets.

Primary data processing. Graphical analysis. Working with matrices. Working with distributions. Correlation and regression analysis in Excel. Time-series analysis.

Topic 3. Statistical analysis in Mathcad.

Primary data processing. Core statistical functions. Correlation and regression analysis. Special econometric methods. GLS. Econometric systems of equations. Modeling of stochastic processes and parameter estimators.

Topic 4. Working with R.

Introduction to R application. Creating a file. Core functionality. Graphics basics. Data visualization. Various distributions generation. Hypothesis testing, parameters estimators and samples numeric characteristics estimators. Correlation and regression analysis in R. Variance analysis. Time-series and predictors.

Topic 5. Core functionality of Statistica.

Creating a file. Primary data processing. Core statistical functions. Graphical analysis. Correlation and regression analysis of numeric data. Rank correlation. Contingency tables. Variance analysis. Time-series and predictors. Cluster analysis. Discriminant analysis.

Course total value is 2 credits (72/32 hours).

Hours distribution over the course

№	Topics	Semester	Week	Total hours	Class hours			Self-studies	Assessment form
					including				
					lectures	seminars	Lab practicals		
1	Software applications review	2		6			2	4	Lab reports during the semester
2	Working with Excel spreadsheets	2		6			2	4	
3	Statistical analysis in Mathcad	2		16			6	10	
4	Working with R	2		20			8	12	
5	Core functionality in Statistica	2		24			14	10	
TOTAL				72			32	40	End-of-term test

5. Educational technologies:

Active forms of class activity are used over the course: seminars, discussions, case studies, students' research groups' results. Master students will meet russian and foreign professors, participate in russian and international conferences throughout the course. One of the essential active learning methods for professional competences is an ongoing seminar with leading researchers and teachers participation, which is a basis for individual masters' curriculums adjustment.

6. Students' self-study educational framework:

Lab practice topics

1. Overall software review.
2. Excel. Primary data processing. Graphical analysis. Correlation and regression analysis. Applying software to solve econometric equations systems.
3. Mathcad. Primary data processing. Core functionality for distributions processing (generation, parameter estimators, etc.).
4. Mathcad. Correlation and regression analysis.
5. Mathcad. Special econometric methods. Econometric equations systems.
6. R. Introduction. Creating a file. Core functionality. Graphics.
7. R. Various distributions generation. Parameters estimators. Hypothesis testing.
8. R. Regression analysis.
9. R. Variance analysis.
10. Statistica. Creating file. Primary data processing. Essential statistical functions. Graphical analysis.
11. Statistica. Regression and correlation analysis of numeric data.
12. Statistica. Non-parametric analysis. Grade correlation.
13. Statistica. Categorized data analysis. Contingency tables.
14. Statistica. Variance analysis.
15. Statistica. Time-series and predictors.
16. Statistica. Classification methods.

.

Sample self-study tasks

Task 1. Find sample numerical characteristics estimates and plot histogram and empiric distribution using functions of learned applications. Estimate numerical characteristics and distribution parameters. Test distribution type hypothesis.

Task 2. Perform regression analysis for two samples: test for dependence, determine dependence type, formulate regression equation, and estimate coefficients. Test resulting model quality. Make a prediction.

Task 3. Find sample correlation coefficients, defining dependence based on several samples. Test hypothesis on significant difference from zero for all correlation coefficients. Select a valid measure of correlation ratio depending on input data nature.

Task 4. Find a sample value of contingency coefficient and test lack of statistical dependence hypothesis based on provided table of two categorized data.

Task 5. Select a trend model and estimate its coefficients for a given time-series. Make a prediction.

Task 6. Solve simultaneous equations.

Task 7. Perform a classification for a given set of objects and different number of classes (i. e. segment objects into homogenous groups based on a particular property).

Note. All the statistical data is provided to students by the teacher.

7. Recommended literature

a) basic

1. V. Borovikov. Topical introduction to Statistica. – Saint-Petersburg.: Piter, 2000. – 269 P. [in russian]
2. V. Borovikov. Statistica. Art of the data analysis on a computer. For professionals.– Saint-Petersburg.: Piter, 2003. – 688 P. [in russian]
3. E. Vukolov. Statistical analysis essentials. Statistical methods and operations research practicum using Statistica and Excel: workbook. – Moscow.: FORUM:INFRA-M, 2004. – 464 P. [in russian]
4. I. Zariadov, Introduction to statistical software R: variable types, data structures, data input and output, graphics: workbook. – Moscow: RUDN, 2010. – 207 P. [in russian]
5. I. Zariadov, Statistical software R: probability theory and mathematical statistics: Workbook. – Moscow: RUDN, 2010. – 141 P. [in russian]
6. O. Salmanov. Mathematical economics with Mathcad and Excel. – Saint-Petersburg.: BHV-Peterburg, 2003. – 464 P. [in russian]
7. E. Chekotovsky. Graphical analysis of statistical data in Microsoft Excel.-Moscow.- "Williams" Publishing, 2002. – 464 P. [in russian]
8. A. Khalafian. Statistica 6. Statistical data analysis. Textbook.– Moscow.: "Binome-press", 2010. – 528 P. [in russian]

b) additional

1. S.Aivazian, V.Mkhitarian; "Applied Statistics in tasks and exercises. University textbook.".- Moscow.:UNITY-DANA, 2001.- 270P [in russian]
2. T. Hill, P. Lewicki. STATISTICS: Methods and Applications. - StatSoft, Tulsa, 2007.- 800 P.
3. B. Dretzke. Statistics with Microsoft Excel. – Pearson Education, 2011.- 288 P.
4. S. Stowell. Using R for Statistics.- Apress, 2014.- 250 P.
5. R. W. Soutas-Little, D. J. Inman. MathCAD Manual for Statistics.- Prentice Hall, 1999.- 166 P.

Educational infrastructure and supplies:

Teachers and students have access to all the necessary infrastructure including computers with the required software (Mathcad, Statistica, R), internet access, database access, book stock of TSU Library (3.8 mil. books).

The program is completed in accordance with the requirements of the FGOS VPO with the advice of PrOOP VPO for majoring specialty 01.04.02 - Applied Mathematics and Computer Science.

Author: T. V. Kabanova, PhD.
Referee: Prof. A. A. Nazarov.

The program was approved at the meeting of the Academic Council of the Department of Applied Mathematics and Cybernetics of Tomsk State University, Date: 10.12.2014, Rec. No. 317.