

# Elementary Statistics and Computing

Instructors: Henryk Banaszak, Michał Wenzel

## A. Object and scope

The aim of the course is to present basic methods and concepts of statistical description and elementary statistical data reporting.

After completion students should be able to analyze by themselves survey data on elementary level, including computing simple multivariate linear models and interpreting their parameters.

## B. Organization

Course is composed of 20 hours of lectures and 20 hours of workshops.

Lectures are dedicated to brief introduction of statistical concepts and description of their formal properties.

Workshops are dedicated to practical application of statistical tools in data analysis and take place in the computer laboratory. In the laboratory participants do series of exercises using *SPSS* package and practice writing statistical data analysis reports.

## C. Requirements for completion

Report containing results of analysis of artificial data set using concepts and techniques introduced during the course. Requirements for the report are enclosed in **Requirements for "Elementary Statistics Report"** document.

## I. Elementary statistics

### A. Lectures (20h)

#### 1. Statistical description

1. Data matrix, case, variable, variables' value
2. Frequency, proportion, variable distribution
3. Central tendency measures and their properties
4. Dispersion measures and their properties
5. Standardized variable
6. Joint bivariate distribution.
7. Covariance
8. Conditional distribution, stochastic independence
9. First type regression of means
10. Variance decomposition, eta-square coefficient
11. Linear regression and correlation coefficient
12. Multiple linear regression

## **B. Workshops (20h)**

### **1. Calculations and graphic: from Excel to Word**

Computing in Excel: formula, arguments, outcome.  
 Computing percentages, cumulative frequencies, consecutive differences.  
 From cumulative to ordinary distribution  
 Graphics in EXCEL - LUCE rule  
 Table and chart from EXCEL to Word

### **2. SPSS objects**

Windows of SPSS: data, syntax, output  
 Customising SPSS outlook  
 Working files of SPSS: data, syntax, output, journal,  
 Creating simple dictionary of simple data set  
 Simple run: making univariate distribution

### **3. Working with SPSS Syntax files**

Menu mode - batch mode  
 SPSS output: commands, errors, warnings  
 Mixed mode: menu - pasting commands - execution  
 Syntax editor operations: help, paste variable names,  
 Data editor operations: toolbar

### **4. Data management**

Working with datasets: splitting files, selecting cases, saving  
 Creating new variable  
 Compute - arithmetic expressions

### **5. Presentation of statistical analysis outcome**

SPSS tables: frequencies, conditional distributions, regression details  
 Pivot and other table operations - cleaning SPSS Output  
 Publishing results of analysis: SPSS Output – Excel – Word

### **6. Single variable description**

Computing frequencies, means and variances  
*Writing comparison of univariate distributions: shape, central tendency, dispersion*

### **7. Bivariate description**

Computing 1st type regression of means, linear regression  
*Writing comparison of bivariate dependencies: 1st type regression, linear regression - shape, direction, strength, linearity.*

### **8. Multivariate description**

Computing multiple linear regression: equation coefficients, Rsq, Rsq increase  
*Writing comparison of multiple regression models: direction, strength, explanatory power*

## C. Readings

Statistical part of the course is based on author's *Elementary Statistics Lecture Notes, Part I and II*. They are accessible, as the remaining relevant course materials, from CSS Intranet.

Knowledge of **Basic readings** is sufficient to follow the lecture and workshops but supplementary readings are strongly recommended.

Computing part of the course is based on the author's *Computing in CSS - Handbook* accessible from CSS Intranet.

### 1. Basic readings

2003 Banaszak, Henryk, *Elementary Statistics Lecture Notes, Part I and II*. CSS Warsaw.

2004 Banaszak, Henryk, *Computing in CSS - Handbook*. CSS Warsaw.

### 2. Supplementary readings

1975a Kim Jae-On, Kohout Frank, J. : Multiple regression analysis: subprogram REGRESSION, in: Nie Norman, H. et. al., *SPSS. Statistical Package for the Social Scientists. Second Edition*, pp. 320-342, McGraw-Hill

1975b Kim Jae-On, Kohout Frank, J. Analysis of variance and covariance: subprogram ANOVA and ONE-WAY, in: Nie Norman H. et. al., *SPSS. Statistical Package for the Social Scientists. Second Edition*, pp. 398-410, McGraw-Hill.

1997 Norusiss, Marija J., *SPSS 7.5 Guide to Data Analysis*, Prentice-Hall, Inc., New Jersey

1993 Vogt, W. Paul, *Dictionary of STATISTICS and METHODOLOGY a nontechnical guide for the social sciences*, Sage Publications.