Elements of mathematical statistics by the use of programming language R

Wasyl Kowalczuk, Ph.D., Dr. Habil.

Department of Theory of Continuous Media

The aim of this cycle of lectures and laboratory exercises is to present basic concepts in analysis, interpretation, and presentation of scientific data with the help of an open source programming language R, which certainly is one of the most suitable software environments for statistical computing and graphical representation of the results of data analysis. During this course the theoretical overview of classical probability distributions, descriptive and inference statistics together with correlation and regression analysis will be given to enhance understanding of the formal numerical methods used in statistical analysis of the scientific data.

Main topics:

1. Introduction to programming in R. From data to graphics.
2. Descriptive statistics: measures of central tendency (mean, median, mode), measures of variability/dispersion (standard deviation/variance, minimum/maximum values, quantiles, quartiles, interquartile range), measures of shape (kurtosis, skewness), scatter plots, histograms, box-and-whiskers plots.
3. Elements of probability theory: discrete/continuous probability distributions, cumulative distribution function, probability density function, law of large numbers, central limit theorem.
4. Statistical inference: hypotheses testing, test statistics, type I and type II errors, significance level, power of the test, parametric/non-parametric tests.
5. Estimation statistics: point and interval estimations, confidence intervals, population and sample effect sizes, sampling methods, sample size determination, precision of the estimate.
6. Correlaton analysis: Pearson correlation coefficient, Spearman's and Kendall tau rank correlation coefficients, correlation matrices, multiple correlation coefficient.
7. Regression analysis: linear/nonlinear regression models, assumptions, regression diagnostics, interpolation/extrapolation, power and sample size calculations, applications.

The total number of lecture hours: 30, laboratory exercises: 30 hours, self-teaching: 45 hours, direct tutoring and consultations: 15 hours.

ECTS Points: 4.