INTRODUCTION

The articles we would like to present are the results of the 21st Conference on Multivariate Statistical Analysis – MSA 2002. The conference took place in Łódź, in the period 4–6 November 2002. It was organised by the University of Łódź (Chair of Statistical Methods) and Polish Statistical Association. The range of subjects of interests of the Conference papers was wide, from the theory of probability, through statistical inference to different applications of statistical methods in various disciplines: finance, insurance and medicines.

The papers are divided into three thematic sections:

1. Probability.
2. Statistical Inference.
3. Applications.

The first paper devoted to the use of stochastic gradient boosting in nonparametric tree-based regression was proposed by Eugeniusz Gatnar (*Gradient Boosting in Regression*). The author studied three aggregation methods used in classification: bootstrap aggregation (bagging), adaptive resample and combine (boosting) and adaptive bagging (hybrid bagging-boosting procedure).

Tadeusz Gerstenkorn (*A Compound of an Inflated Pascal Distribution with the Poisson One*) focused his attention on a compound of an inflated Pascal distribution with the Poisson one. Probability function of the compound distribution Pascal-Poisson, factorial, crude and incomplete moments as well recurrence relations of this distribution are presented.

The next paper by Krzysztof Jajuga (*Tail Dependence in Bivariate Distributions*) was concerned with the problem of tail dependence for bivariate data. The particular emphasis was put on the conditional correlation coefficients and tail dependence coefficients. It was shown how the latter could be analyzed through copula analysis.

Grażyna Trzpiot (*Effectiveness of Stochastic Dominance in Financial Analysis*) focused on portfolio analysis that can be regarded as a problem of choosing the best investment project from all possible investments. This choice depends on, the unique for each investor, utility function and the distribution of the return of the investment project. The results of analysis the properties of the optimal efficient set according SD criteria for asymmetric distribution were presented.

The paper by Jerzy Korzeniewski (*Analysis of Point Processes Observed with Noise with Applicational Example*) presented an example of the application of point processes observed with noise are aerial photographs of forests with the aim of estimating the actual number of trees on a given area. A new algorithm to estimate directly the number of actual trees was proposed, where the only assumption on which the new measure depends is the natural assumption about forest density being locally constant.

Sebastian Sitarz (*Stochastic Orders in Discrete Dynamic Programming*) considered a problem of dynamic optimization with values of criteria function in the set of the random variables. Dynamic model with finite number of stages, states and decision variables was described. Such a dynamic process is evaluated regarding values of the random variables. The random variables have to fulfil some conditions, if they are to be applied to dynamic optimization. These conditions are described in presented paper and there is given a review of stochastic orders, which can be used in the model.

Tomasz Jurkiewicz and Krzysztof Najman (*Proposition of Applying K-Means Classification Method and the SOM Type Neural Network to Improve the Efficiency of Small Domains Estimation in a Representative Study of Small and Medium-Sized Enterprises*) presented a proposition of two-stage estimation process. In the first stage, using the SOM-type neural networks and using the k-means classification method the similarity of components belonging to the small domain with the components belonging to the remaining part of the sample is determined. The second step consists in using the information only from those domains, which are similar to the studied small domain with the help of appropriately construed weights. Authors presented the results of the above procedure in the analysis of the building industry on the basis of a representative study of small and medium-sized enterprises. They have also undertaken an attempt to estimate the errors of the synthetic estimation method modified in such a way.

In the paper by Tomasz Żądło (*On Synthetic Ratio Estimator Based on Superpopulation Approach*), properties of a predictor of the form of synthetic ratio estimator of domain total, known from randomisation approach, were considered. The proof of its ξ-unbiasedness for simple regression superpopulation model in strata was shown. For the model BLU predictor was also presented. Equations of prediction variances of both predictors were derived. For considered predictors the problem of model misspecification was considered and equations of prediction mean square errors were derived.

Adam Depta (*The Use of Blume and Vasicek Methods in the Estimation of Beta Coefficient in the Single-Index Model*) presented alternative methods

of valuation of coefficients beta. The estimation of future coefficients beta can be received by delimitation the coefficients beta from past data and use these coefficients as the estimation of future coefficients beta.

The next section entitled *Statistical Inference* starts with the contribution of Czesław Domański (*Some Remarks on Statistical Inference for Complex Samples*). The author presented the problems in estimation and verifications of hypothesis of consistency of distributions for complex samples, where observations in these samples are stochastically dependent and have different distribution.

Krystyna Pruska (*Tests for Ratio of Two Means in Case of Small Areas*) disscused testing procedures for verification of hypothesis which says that there is no difference between the ratio of small area mean and population mean for analysed variable and auxiliary variable. The properties of one considered procedure were investigated with the use of simulation methods.

The next paper by Janusz Wywiał (*On Estimation of Dominant of Multidimensional Random Variable*) considered the problem of estimation of the mode of a continuous distribution function of multidimensional random variable. The biased estimators of values of modal regressions were proposed. Moreover, the well-known "jackknife" procedure was proposed to evaluate the mean square errors of the estimators.

Aleksandra Baszczyńska (*Some Remarks on the Choice of the Kernel Function in Density Estimation*) focused on kernel method in density estimation, with particular emphasis on influence of the choice of the kernel function $K(u)$ on the quantity of smooothing. Monte Carlo study was presented, where seven kernel functions (Gaussian, Uniform, Triangle, Epanechnikov, Quartic, Triweight, Cosinus) are used in density estimation.

Wojciech Gamrot (*On Application of Logistic Regression to Mean Value Estimation in Two-Phase Sampling for Nonresponse*) investigated alternative estimators for two-phase sampling scheme using estimates of response probabilities obtained on the basis of logistic regression model. The results of Monte Carlo simulation study comparing the properties of these estimators were also presented. In the simulations, the data from the Polish 1996 Agricultural Census were used.

The next paper by Alina Jędrzejczak (*Properties of the Cox Consistency Test in the Case of Income Distribution Analysis*) presented main properties of the Cox statistic which is based on likelihood ratio. The presented results were obtained by means of the Monte Carlo experiment. The theoretical distributions most often used in income distribution analysis as the gamma, lognormal, Dagum and Singh-Maddala were taken into consideration.

Dorota Pekasiewicz (*Application of Simulation Methods to Estimation of Variance of Nonparametric Sequential Estimator of Mean*) proposed an applying simulation methods to estimate the variance of a nonparametric

estimator of mean. An application of bootstrap methods to estimate the variance of a synthetic estimator of the mean in sequential estimation was also presented.

The last paper in this section by Agnieszka Rossa (*Unbiased Estimation of Survival Probabilities for Censored Data with Known Censoring Times*) investigated a class of unbiased estimators of survival probability $P(T_i > t)$ under random and independent censorship model is considered, where the potential survival times T_i are possibly unobserved, but the censoring times Z_i and $\min(T_i, Z_i)$ are known and the sample size is random.

In the group of articles dealing with applications Bronisław Ceranka and Małgorzata Graczyk (*Construction of Optimum Balance Weighing Designs Based on Balanced Block Designs*) discussed problem of estimation of the individual unknown measurements (weights) of p objects when we have at our disposal n measurement operations (weighings).

Małgorzata Misztal (*On the Application of Classification and Regression Trees in Medical Diagnosis*) considered a decision tree as a graphical presentation of the recursive partitioning the learning set into homogenous subsets considering dependent variable y. The aim of this paper is to present some applications of regression and classification trees in medical diagnosis for solving decision – making problems.

The paper by Anna Szymańska (*Application of Selected Statistical Methods in Assessing Homogeneity of Insurance Portfolio*) focused on the assessment of selected methods of testing portfolio homogeneity illustrated with an example of motor insurance.

Monika Zielińska (*Dynamic Schemes of Hedging – Delta Hedging and Delta-Gamma Hedging on Currency Market*) presented two dynamic schemes of hedging: delta hedging and delta-gamma hedging with examples from the Polish currency market. The hedging techniques which reduce the loss without excluding the profits from currency movements are given preference. Their application requires access to reliable forecasts of future currency movements, as well as certain readiness to bear that kind of risk.

The paper by Alicja Ganczarek (*APT Model For Electricity Prices on the Day Ahead Market of the Polish Power Exchange*) considered model of the dependence of the electricity price on macroeconomic factors such as changes in the dollar price, the deutsche mark price, the rate of inflation, the rate of unemployment, price changes in the mining industry, the production of the manufacturing sector, the output of the mining industry and weather conditions. The aim of this article was the empirical verification of the price model on the Day Ahead Market (DAM) of the Polish Power Exchange in 2001 based on the principal components method.

Tomasz Kozdraj (*Remarks on Bayesian Networks and Their Applications*) focused on Bayesian networks that are directed acyclic graphs that represent

dependencies between variables in a probabilistic model. This paper explores the nature of implications for Bayesian networks beginning with an overview and comparison of inferential statistics and Bayes' Theorem. It presents the possibilities of applications of Bayesian networks in a field of economic problems and also focuses on the problem of learning.

The next paper by Jarosław Michalak (*Using Control Charts to Detect Small Process Shifts*) dealt with the selection of proper SPC charts is essential to effective statistical process control implementation and use. This paper shows that, the Cumulative-sum control charts (CUSUM) and Exponentially Weighted Moving Average control charts (EWMA) are appropriate to detect these shifts.

Marek Szajt (*Identification of Patent-Activity Level with the Usage of Discriminant Analysis*) tried to separate particular country groups in Europe on the basis of patent activity. The division has been made with the usage of statistical methods – mainly discriminant function. The analysis presented in the thesis allows characterizing particular participants and drawing one' s attention to the differences in innovative policy conducted in different countries.

The paper by Aneta Włodarczyk, Tomasz Szmigiel (*Theoretical Aspects of Using Markov Models in Research of Exchange Rate Volatility*) was concerned with modeling of short-run exchange rate fluctuations using Markov models.

Jan Żółtowski (*Application of Probit Models and Selected Discrimination Analysis Methods for Credit Decision Evaluation*) disscused the problem of evaluation to which of the two groups the person applying for a credit should be assigned to: a) those who possess the creditworthiness; b) those who do not possess the creditworthiness. It analyses the possibility of applying the probit models and the discrimination analysis methods using the quadratic and linear discrimination function. An evaluation of the correctness of the classification based on the real data from a commercial bank is conducted.

Aleksandra Baszczyńska
Czesław Domański