XLV Conference on
Mathematical Statistics
in Będlewo

2 – 6 December, 2019

Contents

Conference programme

Monday, 2nd of December	
8:00	Breakfast
9:15–9:25	Opening of the conference
9:25–10:10	**Piotr Zwiernik** *Totally positive distributions: Markov structures, graphical models, and convex optimization*, part I (p. 13)
10:10–10:30	Coffee break
10:30–10:50	**Jan Mielniczuk** *Distributions of general information-theoretic criteria and their applications* (p. 40)
10:50–11:10	**Małgorzata Łazęcka (PhD Stud),** Jan Mielniczuk *Selekcja zmiennych oparta na przybliżeniach warunkowej informacji wzajemnej* (p. 36)
11:10–11:25	**Maryia Shpak (PhD Stud)** *Structure learning for Bayesian Networks via penalized likelihood method* (p. 47)
11:25–11:50	Coffee break
11:50–12:10	**Joanna Wasiura-Maślany (PhD Stud)** *Prawie pewna wersja centralnego twierdzenia granicznego dla sum zmiennych losowych z losowymi indeksami* (p. 54)
12:10–12:30	**Paweł Kurasiński (PhD Stud)** *Dokładne prawa wielkich liczb i ich zastosowania* (p. 35)
12:30–12:50	**Luiza Pańczyk (PhD Stud),** Mariusz Bieniek *On the choice of the optimal single order statistics in quantile estimation* (p. 43)
13:00	Lunch
15:00–15:20	**Zbigniew Szkutnik** *Regularizations induced by filters and a discrepancy principle for Poisson inverse problems* (p. 50)
15:20–15:40	**Monika Mokrzycka (PhD Stud)** *On some covariance structure estimators under doubly multivariate model* (p. 41)
15:40–16:00	**Grzegorz Wyłupek** *Data-driven tests for trend under right-censored model* (p. 55)
16:00–16:30	Coffee break
16:30–16:50	**Magdalena Szymkowiak** *Generalized aging intensity functions with applications to lifetime analysis* (p. 51)
16:50–17:10	**Agnieszka Goroncy** *Górne oszacowania uogólnionych statystyk pozycyjnych z rozkładów o rosnącej uogólnionej intensywności awarii* (p. 23)
17:10–17:30	**Krzysztof Jasiński** *Asymptotyczne zachowanie k-tych rekordów z rozkładów dyskretnych* (p. 27)
18:00	Dinner

Tuesday, 3th of December	
8:00	Breakfast
9:00–9:45	**Piotr Zwiernik** *Totally positive distributions: Markov structures, graphical models, and convex optimization*, part II (p. 13)
9:45–10:00	Coffee break
10:00–10:20	**Tadeusz Inglot** *Intermediate efficiency of tests for uniformity under heavy-tailed alternatives* (p. 25)
10:20–10:40	**Bogdan Ćmiel,** Teresa Ledwina *Validation of Association* (p. 19)
10:40–10:55	**Malwina Janiszewska (PhD Stud),** Anna Szczepańska-Álvarez *Testing of the association between two groups of features* (p. 26)
10:55–11:30	Coffee break
11:30–11:50	**Błażej Miasojedow** *Non-asymptotic Analysis of Biased Stochastic Approximation Schemes* (p. 37)
11:50–12:10	**Patrick Tardivel** *Geometry for uniqueness and model selection for LASSO and SLOPE estimators* (p. 52)
12:10–12:30	**Wei Jiang (PhD Stud)** *Adaptive Bayesian SLOPE - high-dimensional model selection with missing values* (p. 28)
12:30–12:50	**Michał Kos (PhD Stud),** Małgorzata Bogdan *False Discovery Rate Control for the Linear and Logistic Regression via SLOPE Method* (p. 33)
13:00	Lunch
14:30–15:15	**Piotr Zwiernik** *Totally positive distributions: Markov structures, graphical models, and convex optimization*, part III (p. 13)
15:15–15:30	Coffee break
15:30–16:00	**Małgorzata Bogdan** *VARCLUST: mathematical bases, R package and application* (p. 16)
16:00–16:20	**Krzysztof Rudaś (PhD Stud),** Szymon Jaroszewicz *Properties of uplift estimators for small samples* (p. 46)
16:20–16:40	**Mateusz John (PhD Stud)** *Testing Toeplitz covariance structure under multivariate model* (p. 29)
16:40–17:00	**Adam Mieldzioc (PhD Stud)** *Porównanie estymatorów macierzy kowariancji o strukturze wstęgowej macierzy Toeplitza* (p. 39)
17:00–17:10	Coffee break
17:10–18:40	Discusion Panel: Modern Statistics and new challenges – Małgorzata Bogdan moderator
18:40	Dinner
19:20	Committee meeting

Wednesday, 4th of December	
8:00	Breakfast
9:00–14:00	Excursion
14:00–15:00	Lunch
15:00–15:45	**Grzegorz Rempała** *Introduction to Mathematical Theory of Chemical Reaction Networks with Applications in Biosciences*, part I (p. 11)
16:00–16:15	Coffee break
16:15–16:30	**Marek Kociński** *On the Estimation on the Financial Market with the Liquidity Shortage* (p. 32)
16:30–16:50	**Piotr Sulewski** *Kolmogorov-Smirnov-like goodness of fit test with two-component test statistics* (p. 48)
16:50–17:10	Coffee break
17:10–17:30	Barbara Kowalczyk**, Wojciech Niemiro,** Robert Wieczorkowski *Badania ankietowe dotyczące drażliwych kwestii: nowa odmiana "Item Count Technique"* (p. 42)
17:30–17:50	**Katarzyna Brzozowska-Rup, Mariola Chrzanowska** *Structural Equation Models - selected aspects* (p. 17)
17:50–18:10	Sylwia Filas-Przybył**, Tomasz Klimanek** *Employment-related commuting to provincial capital cities and their functional areas in Poland* (p. 31)
19:00	Conference Dinner

	Thursday, 5th of December
8:00	Breakfast
9:00–9:45	**Grzegorz Rempała** *Introduction to Mathematical Theory of Chemical Reaction Networks with Applications in Biosciences*, part II (p. 11)
9:45–10:00	Coffee break
10:00–10:20	**Piotr Pokarowski** *P-values for model selection and comparison* (p. 44)
10:20–10:40	**Wojciech Rejchel** *Variable selection in high-dimensional binary regression* (p. 45)
10:40–10:55	**Anna Szczepańska-Álvarez,** Adolfo Álvarez, Bogna Zawieja *Simulation study about the properties of MLE of separable covariance matrix* (p. 49)
10:55–11:20	Coffee break
11:20–11:40	Tomasz Górecki**, Mirosław Krzyśko,** Waldemar Wołyński *Measuring and testing mutual dependence for functional data* (p. 34)
11:40–12:00	Maciej Karczewski, **Andrzej Michalski** *A data driven kernel estimator of the density function* (p. 38)
13:00	Lunch
15:00–15:45	**Grzegorz Rempała** *Introduction to Mathematical Theory of Chemical Reaction Networks with Applications in Biosciences*, part III (p. 11)
15:45–16:00	Coffee break
16:00–16:30	**A. Jokiel-Rokita,** R. Magiera, P. Skoliński *Estimation for trend-renewal processes* (p. 30)
16:30–16:50	**Katarzyna Filipiak** *Structured covariance matrices: identification and testing* (p. 21)
16:50–17:05	**Jacek Badnarz** *Measuring the use intensity of global migration network nodes – a simple proposal* (p. 15)
17:10–17:30	**Tomasz Cąkała** *Poisson Tree Particle Filter* (p. 18)
18:00	Dinner

Friday, 6th of December	
8:00	Breakfast
9:00–9:20	Anna Czapkiewicz, **Antoni Leon Dawidowicz**
	On some Limit Theorem for Markov Chain (p. 20)
9:20–9:40	**Wojciech Zieliński**
	Confidence interval for odds ratio (p. 56)
9:40–10:00	Coffee break
10:00–10:20	**Przemysław Grzegorzewski**
	Two-sample goodness-of-fit test for fuzzy data (p. 24)
10:20–10:40	**Paweł Teisseyre**
	Learning classifier chains under feature extraction budget (p. 53)
10:40–11:00	**Konrad Furmańczyk**
	Estimation of classification error for misspecified binary regression model (p. 22)
11:00–11:15	Closing of the conference
11:45	Lunch

Invited Speakers

Grzegorz Rempała studied mathematics at the University of Warsaw from 1987 to 1991, and then he worked at the Computer Science Institute of the Polish Academy of Sciences from 1991 to 1992. In 1992 he moved to the US where in 1996 he completed his PhD thesis at Bowling Green State University. His advisor was Prof. Arjun K. Gupta. In 1998 he nostrified his degree at the University of Warsaw in the Faculty of Mathematics, Informatics and Mechanics. In 2007 he received his habilitation from Warsaw University of Technology. From 1996 until 2008 Rempała was a professor (full professor from 2007) in the Department of Mathematics at the University of Louisville. In 2008 he joint the Department of Biostatistics at the Medical College of Georgia and in 2013 he moved to The Ohio State University (OSU) where he is currently a professor in the Division of Biostatistics and in the Department of Mathematics. In 2016 he was named the interim director of the Mathematical Biosciences Institute at OSU, and served in this position until 2018. Rempała is known for his work on random matrices, in particular on a random permanent function. He has also established some results in nonparametric statistics related to central limit theorems for products of random variables. More recently, he has worked on mathematical models of chemical reactions, diversity of molecular populations, and disease spread across contact networks.
Source: based on: https://en.wikipedia.org/wiki/Grzegorz_Rempala

Introduction to Mathematical Theory of Chemical Reaction Networks with Applications in Biosciences

Grzegorz Rempała

Ohio State University, USA

Chemical Reaction Networks (CRNs) are popular mathematical models often used to describe processes in systems biology, epidemiology, ecology, population dynamics, communication networks, and, of course, (bio)chemistry. Classical CRN is described by a collection of identical units (e.g. molecules) classified into finite number of types (chemical species) and interacting with each another through some pre-determined set of rules (reactions). Reactions are typically combinations of death and birth events where units (molecules) of one type consume or produce unites of other types. CRNs may be used to model macro-level phenomena, like an onset and spread of a disease in a finite population, or to describe processes at molecular level, like a transcription and translation of an RNA molecule. The mathematical theory of CRN is quite rich and touches upon multiple areas of modern Mathematical Sciences including graphical models, stochastic jump processes, percolation theory, and both approximate and exact statistical inference. The series of 3 lectures will cover some aspects of CRNs theory including:
(a) classical Markovian and deterministic formulation for mass-action dynamics
(b) model reduction methods based on quasi steady state approximations and network deficiency theory
(c) basic statistical inference for reaction rates.
The theory will be illustrated with examples from molecular biology and mathematical epidemiology. The material is largely based on my summer school lectures delivered at the Politecnico di Torino in June.

Piotr Zwiernik obtained MSc in Economics, Warsaw School of Economics in 2003 and in 2006 MSc in Mathematics, Warsaw University. In 2007-2010 Piotr Zwiernik was a PhD student in Statistics at the University of Warwick, the title obtained in 2011. In 2011 he was for a postdoctorate at the Mittag-Leffler Institute in Stockholm, Sweden and at IPAM in Los Angeles, USA. In 2012 in the Department of Mathematics and Computer Science at TU Eindhoven, The Netherlands and in 2013-2014 in the Department of Statistics at University of California, Berkeley, USA. In 2015 in the Department of Mathematics at University of Genoa, Italy. Currently, from 2016 Piotr Zwiernik is Assistant Professor in the Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain. His research interests include graphical models with hidden data, algebraic statistics, applications of algebraic geometry and combinatorics in statistics, time series analysis, phylogenetic analysis, symbolic computations.
Source: based on: http://econ.upf.edu/ piotr

Totally positive distributions: Markov structures, graphical models, and convex optimization

Piotr Zwiernik

Universitat Pompeu Fabra

Probability distributions that are multivariate totally positive of order 2 (MTP2) appeared in the theory of positive dependence and in statistical physics through the celebrated FKG inequality. The MTP2 property is stable under marginalization, conditioning and it appears naturally in various probabilistic graphical models with hidden variables. Models of exponential families with the MTP2 property admit a unique maximum likelihood estimator. In the Gaussian case, the MLE exists also in high-dimensional settings, when $p >> n$, and it leads to sparse solutions. The main aim of this lecture is to give an idea of what the MTP2 condition is as well as to show how total positivity becomes useful in statistical modelling. I will discuss in more detail two cases: the Gaussian distribution and the binary Ising model.

Literatura

[1] S. Fallat, S. Lauritzen, K. Sadeghi, C. Uhler, N. Wermuth, and P. Zwiernik (2017), *Total positivity in Markov structures,* Annals of Statistics 45(3), 1152–1184

[2] S. Lauritzen, C. Uhler, and P. Zwiernik (2019), *Maximum likelihood estimation in Gaussian models under total positivity,* Annals of Statistics, Vol. 47, No. 4, 1835–1863

[3] S. Lauritzen, C. Uhler, and P. Zwiernik (2019), *Total positivity in structured binary distributions,* arXiv:1905.00516

[4] R. Agrawal, U. Roy, and C. Uhler (2019), *Covariance Matrix Estimation under Total Positivity for Portfolio Selection,* arXiv:1909.04222

Abstracts

Measuring the use intensity of global migration network nodes – a simple proposal.

Jacek Bednarz

Instytut Ekonomii i Finansów, Katolicki
Uniwersytet Lubelski Jana Pawła II

The aim of this work is to explore the possible types of results that simple methods for the use intensity measurement of global migration network can deliver. Modifications of two statistical i.e. taxonomic methods, that help to characterize the relative intensity of the global migration network, are proposed. The motivation is to understand the bipolar nature of the migration network node behaviour. In this regard, the major finding is the possibility of divergent behaviour between two opposite node types, that satisfy the weak link hypothesis formulated by Granovetter (1973).

Keywords: migration; network, node implosion; network vitality; weak link.

VARCLUST: mathematical bases, R package and application

Małgorzata Bogdan

Uniwersytet Wrocławski

We present a *VarClust* algorithm devoted to the clustering of multivariate data under the assumption that variables is a given group can be obtained as linear combinations of a relatively small number of hidden latent variables (factors), corrupted by the random noise. With the ultimate goal of selecting a model by maximization of the posterior probability, we build a novel K-medioids algorithm made of two steps based on the Laplace approximation. In the first step, given a partition of the data, we estimate the dimensions of different clusters using the PEnalized SEmi-integrated Likelihood Criterion (PESEL) and represent the respective subspaces by providing the appropriate number of first principal components for each cluster. Then, we perform the partition step, where the similarity between a given variable and a linear subspace is measured by the Bayesian Information Criterion in the multiple regression model relating this variable to the set of the subspace's PCs.

We first state some theoretical results including PESEL consistency and convergence properties of *VarClust* when the dimensions of the cluster are fixed. Then, extensive simulations are provided testing the convergence of VarClust in a general setting and the ability of the algorithm to identify a hidden low-dimensional structure. In particular, numerical comparisons with other algorithms show that VarClust may outperform some popular machine learning tools for sparse subspace clustering. Finally, we report the results of real data analysis including TCGA breast cancer data and meteorogical data, which show that the algorithm can lead to meaningful clustering in the real data problems. The proposed method is implemented in the publicly available R package varclust.

This is a joint work with Piotr Sobczyk, Stanislaw Wilczynski and Mateusz Staniak from University of Wroclaw, Piotr Graczyk and Fabien Panloup from University of Angers (France), Julie Josse from Ecole Polytechnique in Paris and Valerie Seegers from the University Hospital in Angers.

Structural Equation Models - selected aspects

Katarzyna Brzozowska-Rup,
Mariola Chrzanowska

Kielce Uniwersity of Technology, Statistical Office in Kielce,
Warsaw University of Life Sciences, Statistical Office in Kielce

There are a number of important attributes in economics, social, behavioral, and health sciences which cannot be observed directly. Examples of such attributes include shadow economy, happiness, depression, anxiety, social competence, latent construct from multiple biomarkers, etc. The necessity to include non-observable (latent) variables, measurement with errors, and the complex structure of the relationship among a set of variables gives grounds for using the structural equation models (SEM). SEM does not designate a single statistical technique but is in fact a flexible set of techniques, applicable to both experimental and nonexperimental data. The versatility of the method, which is undoubtedly its advantage, may, on the other hand, lead to serious difficulties and thereby to questioning of the reasonableness of using the model. The aim of the article is to review the selected elements of the SEM estimation process and indicate a new approach.

Keywords: structural equation modeling (SEM), Maximum Likelihood Method, Likert scale, Weighted Ordinary Least Squares, Bayesian method.

Literatura

[1] K. A. Bollen (1989), *Structural Equations with Latent Variables,* New York: John Willey & Sons

[2] D. Kaplan (2000), *Structural Equation Modeling: Foundations and Extensions,* Thousand Oaks, CA: Sage Publications

[3] S. Chow, M.R. Ho, E. L. Hamaker, C.V. Dolan (2010), *Equivalence and Differences Between Structural Equation Modeling and State-Space Modeling Techniques,* Structural Equation Modeling: A Multidisciplinary Journal, 17:2

[4] P. Tarka (2018), *An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences,* Qual Quant 52:313-354, doi:10.1007/s11135-017-0469-8

[5] K.K. Holst, E.Budtz-Jørgensen (2019), *A two-stage estimation procedure for nonlinear structural equation models,* Biostatistics,
doi:10.1093/biostatistic/kxy082

Poisson Tree Particle Filter

Tomasz Cąkała

University of Warsaw

Since the seminal paper of Andrieu, Doucet, Holenstein (2010), particle filters and particle MCMC methods have emerged as one of the main tools in bayesian inference. We introduce a new version of particle filter in which the number of 'children' of a particle at a given time has a Poisson distribution. As a result, the number of particles is random and varies with time. An advantage of this scheme is that descendants of different particles can evolve independently, which greatly improves parallel implementations. Moreover, particle filter with Poisson resampling is readily adapted to the case when a process of interest is a continuous time, piecewise deterministic semi-Markov process. We show that the basic techniques of particle MCMC, namely particle independent Metropolis-Hastings, particle Gibbs Sampler and its version with ancestor sampling, work under our Poisson resampling scheme. Our version of particle Gibbs Sampler is uniformly ergodic under the same assumptions as its standard counterpart. We present simulation results which indicate that our algorithms can compete with the existing methods.

The presentation is based on joint work with Wojciech Niemiro and Blazej Miasojedow.

Literatura

[1] C. Andrieu, A. Doucet, R. Holenstein, *Particle Markov chain Monte Carlo methods.,* Journal of the Royal Statistical Society B, 2010

[2] F. Lindsten, M.I. Jordan, T.B. Schön, *Particle Gibbs with Ancestor Sampling,* Journal of Machine Learning Research, 2014

[3] F. Lindsten, R. Douc, E. Moulines, *Uniform ergodicity of the particle Gibbs sampler,* Scand. J. Stat, 2015

[4] T. Cakala, B. Miasojedow, W. Niemiro:, *Particle MCMC with Poisson resampling: parallelization and continuous time models.,* arXiv:1707.01660v2

Validation of Association

Bogdan Ćmiel and Teresa Ledwina

AGH University of Science and Technology,
Polish Academy of Sciences

We will consider how the quantile dependence function can be used to construct tests for independence and to provide an easily interpretable diagnostic plot of existing departures from the null model. For a pair of random variables X and Y with bivariate cumulative distribution function H and with continuous margins F and G, respectively, the quantile dependence function is defined as follows

$$q(u,v) := \frac{C(u,v) - uv}{\sqrt{uv(1-u)(1-v)}}, \quad (u,v) \in (0,1)^2,$$

where $C(u,v) = H(F^{-1}(u), G^{-1}(v))$. It gives an insight into how the dependence structures changes in different parts of the joint distribution. We define new estimators of the dependence function, discuss some of their properties, and apply them to construct new tests of independence. Numerical evidence is given to the test's benefits against three recognized independence tests introduced in the previous years.

Literatura

[1] B. Ćmiel and T. Ledwina (2019), *Validation of Association,* arXiv:1904.06519

[2] T. Ledwina and G. Wyłupek (2014), *Validation of positive quadrant dependence,* Insurance: Mathematics and Economics, 56, 38-47.

On some Limit Theorem for Markov Chain

Anna Czapkiewicz

Faculty of Management,
AGH University of Science and Technology

Antoni Leon Dawidowicz

Faculty of Mathematics and Computer
Science, Jagiellonian University

The goal of this paper is to describe conditions which guarantee a central limit theorem for random variables, which distributions are controled by hidden Markov chains. We proved that when a Markov chain is ergodic and random variables fullfiled Lindeberg's condition then the Central Limit Theorem is true.

Literatura

[1] A. Czapkiewicz, A.L.Dawidowicz(2019), *On some Limit Theorem for Markov Chain.*, Submitted to Annals of Applied Probability

[2] R. J.Serfling(1968) , *Contributions to Central Limit Theory for Dependent Variables.*, Annals of Mathematical Statistics, 39,4,1158–1175

Structured covariance matrices: identification and testing

Katarzyna Filipiak

Institute of Mathematics, Poznań University of Technology

For doubly multivariate data we present the method of identification of the separable covariance structures using the entropy loss function as well as we give the likelihood ratio test procedure for testing the structure. We show the properties of both methods of choosing covariance (identification through the discrepancy function and testing). Finally, we apply both of the methods to real data example.

Literatura

[1] K. Filipiak, D. Klein, A. Roy (2017), *A comparison of likelihood ratio tests and Rao's score test for three separable covariance matrix structures,* Biometrical Journal 59, 192--215

[2] K. Filipiak, D. Klein, A. Markiewicz, M. Mokrzycka (2019), *Approximation with a Kronecker product structure with one component as compound symmetry or autoregression via entropy loss function,* Sumbitted

Estimation of classification error for misspecified binary regression model

Konrad Furmańczyk

Institute of Information Technology, WULS-SGGW

The main objective of the talk is to provide a bound for excess risk for classification for misspecified binary regression model. First, we consider case when logistic model is fitted to binary model and second we consider a linear fit. In numerical study we compare the logistic model and linear classification model.

Literatura

[1] M. Kubkowski (2019), *Misspecification of binary regression model: properties and inferential procedures.,* Ph.D. Thesis. Warsaw University of Technology. Warsaw 2019.

[2] P. Ruud (1983), *Sufficient conditions for the consistency of maximum likelihood estimation despite misspecification of distribution in multinomial discrete choice models,* Econometrica, 51:225-228

[3] T. Zhang (2004), *Statistical behavior and consistency of classification methods based on convex risk minimization,* Ann. Stat., 32 (1): 56-134

Górne oszacowania uogólnionych statystyk pozycyjnych z rozkładów o rosnącej uogólnionej intensywności awarii

Agnieszka Goroncy

Uniwersytet Mikołaja Kopernika w Toruniu

W referacie zostaną przedstawione optymalne górne nieujemne oszacowania dla uogólnionych statystyk pozycyjnych, opartych na próbach prostych $X_1, X_2, \ldots$ z rozkładu odystrybuancie F, centrowanych względem średniej z populacji. Dodatkowo nakładamy ograniczenie na F, która należy do rodziny rozkładów o rosnącej uogólnionej intensywności awarii, w szczególności do rodziny o rosnącej gęstości (ID) lub rosnącej intensywności awarii (IFR). Oszacowania zostały wyznaczone metodą rzutowania na odpowiedni stożek wypukły, wyrażone w jednostkach skali odpowiadających odchyleniu standardowemu rozkładu bazowego. Podane zostaną również warunki osiągania równości dla wyznaczonych oszacowań.

Literatura

[1] M. Bieniek (2008), *On families of distributions for which optimal bounds on expectations of GOS can be derived,* Comm. Stat. Theor. Meth. 37, 1997–2009

[2] A. Goroncy (2019), *Optimal upper bounds on expected kth record values from IGFR distributions,* Statistics 53:5, 1012–1036

[3] A. Goroncy (2019), *Upper bounds on expected generalized order statistics based on IGFR distributions,* w przygotowaniu

Two-sample goodness-of-fit test for fuzzy data

Przemysław Grzegorzewski

Faculty of Mathematics and Information Science,
Warsaw University of Technology
and Systems Research Institute, Polish Academy of Sciences

When the output of an experiment consists of a random sample of imprecise data that could be satisfactorily described by fuzzy numbers we need a model which allow to grasp both aspects of uncertainty that appear in such data: randomness, associated with data generation mechanism and fuzziness, connected with data nature, i.e. their imprecision. To cope with this problem Puri and Ralescu (1986) introduced the notion of a fuzzy random variable.

However, in analyzing fuzzy data from a statistical perspective one immediately come upon some key obstacles, like the nonlinearity associated with the fuzzy number arithmetic, the lack of suitable probability distribution models or no Central Limit Theorems for random mechanisms producing fuzzy data which could be directly applied in statistical inference. To overcome these drawbacks one usually applies an appropriate metric between fuzzy data and a bootstrapped central limit theorem for general space-valued random elements (see, e.g., Blanco-Fernández et al., 2014).

In our presentation we consider a two-sample goodness-of-fit problem for fuzzy data. We suggest its solution based on another methodology, i.e. combining some L^2-type metric and a permutation test.

Literatura

[1] A. Blanco-Fernández, M. R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M. A. Gil, G. González-Rodríguez, M. T. López, M. A. Lubiano, M. Montenegro, A. B. Ramos-Guajardo, S. de la Rosa de Sáa, B. Sinova (2014), *A distance-based statistic analysis of fuzzy number-valued data,* Int. J. Approx. Reason. 55, 1487–1501.

[2] M. L. Puri, D. A. Ralescu (1986), *Fuzzy random variables,* J. Math. Anal. Appl. 114, 409–422.

Intermediate efficiency of tests for uniformity under heavy-tailed alternatives

Tadeusz Inglot

Faculty of Pure and Applied Mathematics,
Wrocław University of Science and Technology

We show that for local alternatives which are not square integrable the intermediate (or Kallenberg) efficiency of the Neyman-Pearson test for uniformity with respect to the classical Kolmogorov-Smirnov test is equal to infinity. Contrary to this, for local square integrable alternatives the intermediate efficiency is finite and can be explicitly calculated.

Literatura

[1] B. Ćmiel, T. Inglot and T. Ledwina, (2018), *Intermediate efficiency of some weighted goodness-of-fit statistics,* submitted

[2] T. Inglot (2019), *Intermediate efficiency of tests under heavy-tailed alternatives,* Probab. Math. Statist., accepted

[3] T. Inglot, T. Ledwina and B. Ćmiel, (2018), *Intermediate efficiency in some nonparametric testing problems with an application to some weighted statistics,* ESAIM Probab. Statist., accepted

Testing of the association between two groups of features.

Malwina Janiszewska, Anna Szczepańska-Álvarez

Poznań University of Life Sciences, Poland

Nowadays, in many experiments the various types of characteristics are analyzed. The features can be divided according to some properties of the sets. For example, in medicine, the relations between physical, biological and chemical features are studied. In agriculture, biochemical and biophysical traits are often observed. In our presentation, we examine the association between two different groups of characteristics measured on the same objects. One set of variables is observed in the time points and second one does not change in the time. We estimate the dispersion matrix based on maximum likelihood estimation and test separability and independence between two types of features. Moreover, we present simulations of distribution of test statistic in likelihood ratio test, when the sample size is small. The results we illustrate in the real data.

Literatura

[1] A Szczepańska-Álvarez, Ch. Hao, Y. Liang, D. von Rosen (2016), *Estimation equations for multivariate linear models with Kronecker structured covariance matrices,* Communications in Statistics-Theory and Methods. Vol 46, no 16, 7902-7915

[2] A Roy and R. Khattree (2005), *On implementation of a test for Kronecker product covariance structure for multivariate repeated measures data,* Statistical Methodology 2, 297-306

[3] G. Marsaglia and G. P. H. Styan (1974), *Equalities and Inequalities for Ranks of Matrices,* Linear and Multilinear Algebra. Vol 2, 269-292

[4] M. Krzyśko (2009), *Podstawy wielowymiarowego wnioskowanie statystycznego,* Wydawnictwo naukowe UAM

[5] T. Kollo and D. von Rosen (2005), *Advanced Multivariate Statistics with Matrices,* Springer, Dordrecht

Asymptotyczne zachowanie k-tych rekordów z rozkładów dyskretnych

Krzysztof Jasiński

Wydział Matematyki i Informatyki,
Uniwersytet Mikołaja Kopernika w Toruniu

Niech $W_0^{(k)}, W_1^{(k)}, \ldots$ oraz $R_0^{(k)}, R_1^{(k)}, \ldots$ bedą odpowiednio słabymi k-tymi rekordami oraz k-tymi rekordami pochodzącymi od obserwacji z rozkladu dyskretnego. Naszym celem będzie zbadanie asymptotycznego zachowania się $W_{n+m}^{(k)}/W_n^{(k)}$ oraz $R_{n+m}^{(k)}/R_n^{(k)}$, $n \to \infty$, w oparciu o teorie funkcji regularnie zmieniających się.

Literatura

[1] K. Jasiński (2018), *Asymptotic properties of the ratio of weak kth records,* Commun. Statist. Theor. Meth. DOI: 10.1080/03610926.2018.1535074

[2] K. Jasiński (2019), *Limiting behavior of the ratio of kth records,* Stat. Probab. Lett. 150, 29–34

Adaptive Bayesian SLOPE – high-dimensional model selection with missing values

Wei Jiang

CMAP, Ecole Polytechnique, Paris

We consider the problem of variable selection in high-dimensional settings with missing observations among the covariates. To address this relatively understudied problem, we propose a new synergistic procedure – adaptive Bayesian SLOPE – which effectively combines the SLOPE method (sorted l_1 regularization) together with the Spike-and-Slab LASSO method. We position our approach within a Bayesian framework which allows for simultaneous variable selection and parameter estimation, despite the missing values. As with the Spike-and-Slab LASSO, the coefficients are regarded as arising from a hierarchical model consisting of two groups: (1) the spike for the inactive and (2) the slab for the active. However, instead of assigning independent spike priors for each covariate, here we deploy a joint "SLOPE" spike prior which takes into account the ordering of coefficient magnitudes in order to control for false discoveries. Through extensive simulations, we demonstrate satisfactory performance in terms of power, FDR and estimation bias under a wide range of scenarios. Finally, we analyze a real dataset consisting of patients from Paris hospitals who underwent severe trauma, where we show excellent performance in predicting platelet levels. Our methodology has been implemented in C++ and wrapped into an R package ABSLOPE for public use.

Literatura

[1] W. Jiang , M. Bogdan, J. Josse , B. Miasojedow , V. Rockova, TraumaBase Group (2019), *Adaptive Bayesian SLOPE – high-dimensional model selection with missing values,* arXiv:1909.06631

Testing Toeplitz covariance structure under multivariate model

Mateusz John

Institute of Mathematics
Poznan University of Technology, Poznan, Poland

The goal of this talk is to show the procedure for testing banded Toeplitz structure of covariance matrix under the multivariate model. Proposed tests are based on the likelihood ratio and Rao score test statistics, in which the maximum likelihood estimator of a penta-diagonal Toeplitz matrix have been replaced by asymptotic estimator (see [1]) or shrinkage estimator (cf. [2]). The properties of the test will be verified.

Literatura

[1] K. Filipiak, A. Markiewicz, A. Mieldzioc, A. Sawikowska (2018), *On projection of a positive defnite matrix on a cone of nonnegative defnite Toeplitz matrices,* Electron. J. Linear Algebra, 33, 74-82

[2] M. John and A. Mieldzioc (2019), *The comparison of the estimators of banded Toeplitz covariance structure under the highdimensional multivariate model,* Comm. Statist. Simulation Comput.
DOI: 10.1080/03610918.2019.1614622

Estimation for trend-renewal processes

A. Jokiel-Rokita, R. Magiera, P. Skoliński

Wrocław University of Science and Technology

The trend renewal process (TRP) model was introduced by Lindqvist [4]. The TRP is defined as a time transformed renewal process and incorporates both time trend and renewal-type behavior. The most commonly used models for recurrent events such as nonhomogeneous Poisson processes and renewal processes are special cases of TRP's. Inhomogeneous gamma processes, introduced by Berman [1], are also the TRP's.

In the presentation, we make some review of estimation methods for semiparametric and parametric TRP models. We present in detail some new results for two representatives of inhomogeneous gamma processes, obtained in [2] and [3].

Literatura

[1] M. Berman (1981), *Inhomogeneous and modulated gamma processes,* Biometrika 68, 143–152

[2] A. Jokiel-Rokita, R. Magiera (2018), *Estimation and prediction for the modulated power law process,* Mathematical and Statistical Methods for Actuarial Sciences and Finance, MAF 2018 / Marco Corazza [i in.] Eds. Cham: Springer International Publishing AG, 443–447

[3] A. Jokiel-Rokita, R. Magiera, P. Skoliński (2018), *Parameters estimation for an inhomogeneous gamma process with log-linear intensity,* Raporty Wydziału Matematyki Politechniki Wrocławskiej. Ser. PRE nr 24, 24 s.

[4] B. H. Lindqvist (1993), *The trend-renewal process, a useful model for repairable systems,* Malmö, Sweden. Society in Reliability Engineers, Scandinavian Chapter, Annual Conference

Employment-related commuting to provincial capital cities and their functional areas in Poland

Sylwia Filas-Przybył
Tomasz Klimanek

Statistical Office in Poznań

The purpose of the presentation is to report the results of the study "Employment-related commuting in 2016", for selected provincial capital cities and their functional areas (FUAs). The presentation describes measures for communes that make up the functional area of a given provincial capital city. These measures include: the ratio of incoming to outgoing commuters (indicator of the "attractiveness" of the local labour market), and the share of outgoing commuters in the total number of employed persons (a measure of the strength of the push factors that drive migration).

The approach presented in the paper involves also converting commuter data into the graph format, where communes of residence and employment are treated as vertices and commuting flows are regarded as edge weights. Values of structural measures computed for the network of commuter traffic are then interpreted to enable conclusions about the socio-economic situation. We also present how decisions regarding the expansion of a given FUA, could be informed by network analysis. The proposal is based on two selected simple structural measures (density, clustering coefficient) describing the graph in relation to commuting ties.

Literatura

[1] G.P. Clemente and R. Grassi (2018), *Directed clustering in weighted networks: A new erspective,* Chaos, Solitons & Fractals, 107,26–38

[2] G. Fagiolo (2007), *Clustering in complex directed networks,* Physical Review E, 76(2) 026106

[3] K.Kruszka (2010), *Dojazdy do pracy w Polsce. Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem [in Polish] [Commuting to work. Territorial identification of population flows related to working],* Ośrodek Statystyki Miast Urzędu Statystycznego w Poznaniu

[4] D.J. Watts and J. Strogatz (1998), *Collective dynamics of 'small-world' networks,* Nature. 393 (6684): 440–442

On the Estimation on the Financial Market with the Liquidity Shortage

Marek Kociński

Institute of Economics and Finance
Warsaw University of Life Sciences SGGW

The level of liquidity is an important characteristic of the financial market. The liquidity shortage implies the transaction costs which depend on the trading volume and the price dynamics parameters of the financial asset: drift and volatility. The problem of estimation of drift and volatility of the financial asset price with the use of transactions data will be considered.

False Discovery Rate Control for the Linear and Logistic Regression via SLOPE Method

Michał Kos, Małgorzata Bogdan

Uniwersytet Wrocławski

Sorted L-One Penalized Estimator (SLOPE) is a solution to a following convex optimization problem:

$$\hat{b} = \arg\min_{b \in \mathbb{R}^p} \left(-l(b) + \sum_{i=1}^{p} \lambda_i |b|_{(i)} \right) \ ,$$

where: $l(b)$ is a loglikelihood function of linear or logistic regression, the $\{\lambda_i\}_{i=1}^{p}$ is a positive, non-increasing sequence of tuning parameters and the $|b|_{(i)}$ is the i-th largest element of $(|b_1|, ..., |b_p|)'$.

SLOPE chooses columns of the design matrix $X_{n \times p}$ that are associated with non-zero elements of $\hat{b}$ and identifies them as important predictors. In linear regression, when the design matrix $X_{n \times p}$ is orthogonal, SLOPE with the sequence of tuning parameters selected according to the thresholds of the Benjamini - Hochberg (BH) procedure for multiple testing controls the False Discovery Rate (FDR).

During the session we will present new results illustrating that SLOPE asymptotically controls FDR for the linear and logistic regression, when entries of the design matrix are iid variables from the normal distribution. We will discuss both low dimensional set-up, where p is fixed and n goes to infinity, and the high dimensional set-up, where p may diverge to infinity much quicker than n. We will illustrate our asymptotic results with computer simulations. Apart from the Gaussian design matrix we will also consider the practical case of the design matrix containing genotypes of independent genetic markers.

Literatura

[1] M. Bogdan, E. van den Berg, C. Sabatti, W. Su, E. Candes (2015), *SLOPE - Adaptive Variable Selection via Convex Optimization,* Annals of Applied Statistics Vol. 9, No. 3, 1103-1140

[2] Michal Kos, Malgorzata Bogdan (2019), *On the asymptotic properties of SLOPE,* arXiv:1908.08791

Measuring and testing mutual dependence for functional data

Tomasz Górecki, Mirosław Krzyśko, Waldemar Wołyński

Department of Mathematics and Computer Science
Adam Mickiewicz University, Poznań, Poland

We propose new measures of mutual dependence for multivariate functional data. Each measure is zero if and only if the vectors of functional features are mutually independent. The proposed measures base on the functional rV coefficient and distance correlation coefficient. The first one is appropriate for linear mutual dependence and the second one for non-linear mutual dependence between the vectors of functional features. Based on the proposed coefficients we can test mutual dependence. The implementation of corresponding tests is demonstrated by both simulation results and real data examples.

Literatura

[1] Górecki T., Krzyśko M., Wołyński W., *(2017) Correlation analysis for multivariate functional data. In: Data Science, Studies in Classification, Data Analysis, and Knowledge Organization, Springer International Publishing 243–258,*

[2] Górecki T., Krzyśko M., Wołyński W., *Independence test and canonical correlation analysis based on the alignment between kernel matrices for multivariate functional data. Artifical Inteligence Review DOI: 10.1007/s10462-018-9666-7,*

[3] Jin Z., Matteson D.S., *(2018) Generalizing distance covariance to measure and test multivariate mutual dependence via complete and incomplete V-statistics. Journal of Multivariate Analysis 168:304-322,*

Dokładne prawa wielkich liczb i ich zastosowania

Paweł Kurasiński

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

W prezentacji będziemy rozważać dokładne prawa wielkich liczb, czyli zbieżność prawie pewną do pewnej niezerowej i skończonej stałej, ważonych sum niezależnych zmiennych losowych o nieskończonej wartości oczekiwanej. W referacie zostaną przedstawione dokładne prawa wielkich liczb dla ilorazów zmiennych losowych, a w szczególności ilorazów statystyk porządkowych. Zwrócimy również uwagę na konieczność badania dokładnych praw wielkich liczb dla zmiennych losowych o różnych rozkładach.

Literatura

[1] P. Matuła, P. Kurasiński, A. Adler (2019), *Exact strong laws of large numbers for ratios of the smallest order statistics,* Statist. Probab. Lett. 152, 69–73

[2] P. Matuła, A. Adler, P. Kurasiński (2018), *On exact strong laws of large numbers for ratios of random variables and their applications,* Communications in Statistics – Theory and Methods

Selekcja zmiennych oparta na przybliżeniach warunkowej informacji wzajemnej

Małgorzata Łazęcka, Jan Mielniczuk

Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska

We will present information-based methods of selecting active predictors for the response. We show that two popular information-based selection criteria which are both derived as approximations of Conditional Mutual Information CMI (which is the expected value of the mutual information (MI) of two random variables given the value of a third; MI measures the mutual dependence between the two variables) may exhibit different behaviour resulting in different orders in which predictors are chosen in variable selection process in the Generative Tree Model (GMT). Distributions of predictors in this model are mixed gaussians and the responce is a binary variable.

We will show explicit formulae for CMI and its two approximations in the GMT. We will establish expressions for entropy of a multivariate gaussian mixture and its mutual information with mixing distribution.

Literatura

[1] G. Brown, A. Pocock, M.J. Zhao, M. Luján (2012), *Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection,* J. Mach. Learn. Res. 13(1), 27–66

[2] S. Gao, G. Ver Steeg, A. Galstyan (2016), *Variational Information Maximization for Feature Selection,* NIPS, 487–495

Non-asymptotic Analysis of Biased Stochastic Approximation Schemes

Błażej Miasojedow

University of Warsaw

Stochastic approximation (SA) is a key method used in statistical learning. Recently, its non- asymptotic convergence analysis has been a fundamental issue considered in many papers. However, most of these analyses are made under restrictive assumptions such as unbiased gradient es- timates and convex objective function, which significantly limit their applications to sophisticated tasks such as online and reinforcement learning. These restrictions are all essentially relaxed in this work. In particular, we consider two general SA schemes to minimize a non-convex objective function. We consider update procedure whose mean field is not necessarily of gradient-type, covering in particular approximate second-order method and allow the one-step update to be a biased estimator of the target mean-field. We illustrate these settings with the online EM algorithm and the policy-gradient method for average reward maximization in reinforcement learning.

Literatura

[1] B. Karimi, B. Miasojedow, E. Moulines, H.-T. Wai, *Non-asymptotic Analysis of Biased Stochastic Approximation Scheme*, Proceedings of the Thirty-Second Conference on Learning Theory, 1944–1974, 2019,

,

A data driven kernel estimator of the density function

Maciej Karczewski, Andrzej Michalski

Department of Mathematics, Wroclaw
University of Environmental and Life Sciences

In the article we propose the concept of local effectiveness of kernel density estimator based on the distance from Marczewski-Steinhaus (cf Karczewski and Michalski, 2018a, 2018b). Literature suggests several different approaches to kernel density estimation (e.g. Berlinet and Devroye, 1994). We focused on three main approaches: Silverman's rule of thumb, cross-validation methods and the plug-in methods (Silverman, 1986, Givens and Hoeting, 2005, Wand and Jones, 1995). We demonstrate that none of considered estimators are optimal on each of selected intervals. In this paper, we present a data driven estimator based on a linear convex combination of the most effective kernel density estimators. Thus we create a new estimator combining the best features of all previously assessed estimators. All numerical calculations were done for an experimental data recording groundwater level on a melioration facility (see Michalski, 2016).

References

[1] A. Berlinet, L. Devroye, *A comparison of kernel density estimates,* Publications de l'Institut de Statistique de l'Université de Paris, vol. XXXVIII –Fascicule 3, 3–59, 1994.

[2] M. Karczewski, A. Michalski, *The study and comparison of one-dimensional kernel estimators – a new approach. Part 1. Theory and methods,* doi.org/ 10.1051/itm-conf/20182300017, 2018a.

[3] M. Karczewski, A. Michalski, *The study and comparison of one-dimensional kernel estimators – a new approach. Part 2. A hydrology case study,* doi.org/ 10.1051/itmconf/ 20182300018, 2018b.

[4] G. H. Givens, J. A. Hoeting, *Computational Statistics,* New York: Wiley & Sons, 2005.

[5] E. Marczewski, H. Steinhaus, *On a certain distance of sets and the corresponding distance of functions,* Coll. Mathematicum 6, 319–327, 1958.

[6] A. Michalski, *The use of kernel estimators to determine the distribution of groundwater level,* Meteorology Hydrology and Water Management,Vol.4, 1, 41–46, 2016.

[7] B. W. Silverman, *Density Estimation for Statistics and Data Analysis,* London: Chapman & Hall, 1986.

[8] M. P. Wand, M. C. Jones, *Kernel Smoothing,* London: Chapman & Hall, 1995.

Porównanie estymatorów macierzy kowariancji o strukturze wstęgowej macierzy Toeplitza

Adam Mieldzioc

Uniwersytet Przyrodniczy w Poznaniu

Estymacja macierzy kowariancji odgrywa kluczową rolę w statystyce matematycznej. Znalezienie dobrze uwarunkowanego estymatora macierzy kowariancji o zadanej strukturze i dobrych własnościach statystycznych może być zadaniem trudnym i czasochłonnym obliczeniowo.

W pracy zostaną porównane estymatory macierzy kowariancji o strukturze wstęgowej macierzy Toeplitza uzyskane za pomocą metody największej wiarogodności z ograniczeniem na współczynnik uwarunkowania macierzy (Won, Lim, Kim, Rajaratnam, 2013) i metody „shrinkage" (Ledoit i Wolf, 2004; John i Mieldzioc, 2019) w modelu, w którym badane jest wiele cech lub jedna cecha w wielu punktach czasowych.

Literatura

[1] M. John and A. Mieldzioc (2019), *The comparison of the estimators of banded Toeplitz covariance structure under the high-dimensional multivariate model,* Communications in Statistics - Simulation and Computation

[2] O. Ledoit and M. Wolf (2004), *A well-conditioned estimator for large-dimensional covariance matrices,* Journal of Multivariate Analysis 88(2), 365–411

[3] JH. Won, J. Lim, SJ. Kim, B. Rajaratnam (2013), *Condition-Number-Regularized Covariance Estimation,* Journal of the Royal Statistical Society. Series B (Statistical Methodology) 75(3), 427–450

Distributions of general information-theoretic criteria and their applications

Jan Mielniczuk

Instytut Podstaw Informatyki PAN i Wydział
Matematyki i Nauk Informacyjnych PW

We study distributions of general information-theoretic criteria introduced in [1] and apply the proved results to conditional independence testing and feature selection. We show that when the criterion is properly normalized and centred, the distribution is either normal or that of quadratic form of normal variables. For the second case we study conditions under which it occurs and show that they crucially depend on values of criterion's parameters. In particular we study the behaviour of two popular criteria, JMI and CIFE and show that their behaviour may differ. Obtained results are applied to construct rejection region for JMI based test of the hypothesis H_0 that candidate variable X and binary target Y are conditionally independent given multivariate X_S, where X_S is a vector of already chosen predictors. The limiting distribution on H_0 is approximated either by a scaled chi square distribution or normal distribution in a data adaptive way motivated by obtained results.

The presentation is based on joint research with Mariusz Kubkowski i Małgorzata Łazęcka.

Literatura

[1] G. Brown, A. Pocock, M-J. Zhao and M. Luján (2012), *Conditional likelihood maximisation: a unifying framework for information theoretic feature selection* , Journal of Machine Learning Research. 13, 27-66

On some covariance structure estimators under doubly multivariate model

Monika Mokrzycka

Institute of Plant Genetics, Polish Academy of Sciences

The maximum likelihood estimation (MLE) of separable covariance structure with one component as compound symmetry or autoregression of order one has been widely studied in the literature. In this talk we propose two different approaches to estimation of mentioned covariance structures based on minimization of the Frobenius norm [1] and the entropy loss function [2, 3]. Since most of the proposed estimators are not given in explicit form, the properties such as biasedness and MSE are compared with the use of simulation studies.

The presented results were carried out in cooperation with K. Filipiak, D. Klein and A. Markiewicz.

Literatura

[1] K. Filipiak and D. Klein (2018), *Approximation with a Kronecker product structure with one component as compound symmetry or autoregression,* Linear Algebra Appl. 559, 11–33

[2] K. Filipiak, D. Klein, M. Mokrzycka (2018), *Estimators comparison of separable covariance structure with one component as compound symmetry matrix,* Electron. J. Linear Algebra 33, 83–98

[3] K. Filipiak, D. Klein, A. Markiewicz, M. Mokrzycka (2019), *Approximation with a Kronecker product structure with one component as compound symmetry or autoregression via entropy loss function,* Submitted

Badania ankietowe dotyczące drażliwych kwestii: nowa odmiana "Item Count Technique"

Barbara Kowalczyk, Wojciech Niemiro, Robert Wieczorkowski

Szkoła Główna Handlowa, Uniwersytet Warszawski, Główny urząd Statystyczny

Jeśli w badaniu ankietowym pojawiają się pytania „drażliwe" (dotyczące np. nieakceptowanych społecznie zachowań), wielu respondentów może odmówić odpowiedzi lub udzielić odpowiedzi nieprawdziwej. Statystycy opracowali techniki konstruowania ankiet w ten sposób, aby zwiększyć poczucie anonimowości respondentów i w rezultacie otrzymywać bardziej wiarygodne wyniki badań. Jeden z ciekawszych sposobów nosi nazwę „Item Count Technique". Przestawimy nową odmianę tej techniki. Respondentów dzieli się na dwie grupy. Zadaje się każdemu respondentowi pytanie „obojętne" oraz pytanie drażliwe. Respondent podaje ankieterowi tylko sumę lub różnicę odpowiedzi na oba pytania. Do analizy wyników tak zaplanowanego badania używa się metody momentów lub metody największej wiarygodności w połączeniu z algorytmem EM.

Literatura

[1] Tian G-L, Tang M-L, Wu Q, Liu Y. (2017), *Poisson and negative binomial item count techniques for surveys with sensitive question,* Stat Methods Med Res 26, 931-947

[2] Trappman M, Krumpal I, Kirchner A and Jann B. (2014), *Item Sum: A New Technique for Asking Quantitative Sensitive Questions,* J Surv Stat Methodol 2, 58-77

On the choice of the optimal single order statistics in quantile estimation

Luiza Pańczyk, Mariusz Bieniek

Instytut Matematyki UMCS w Lublinie

Let $X_{1:n} \leqslant X_{2:n} \leqslant \ldots \leqslant X_{n:n}$ denote the order statistics of the sample of size n. We consider the problem of the estimation of population quantile $F^{-1}(p)$ of order $p \in (0,1)$ by the use of appropriately chosen single order statistic $X_{j:n}$. It is intuitively clear that j should be as close to np as possible, so the usual choice is either $[np]$ or $[np] + 1$. For large n, the difference is irrelevant, but for small sizes it is not so obvious which choice gives 'better' results. We present rather simple conditions on n and p, implying which choice is better. Our criterion of optimality is based on determining sharp bounds on the bias $EX_{j:n} - F^{-1}(p)$. The proofs of our results heavily depend on well known properties of binomial distribution, such as its median or mode.

P-values for model selection and comparison

Piotr Pokarowski

Univeristy of Warsaw

P-wartości mogą być używane do zgodnego wyboru czy porównywania modeli w następujący sposób:

- każdemu modelowi przyporządkowujemy p-wartość ilorazu wiarygodności tego modelu w stosunku do referencyjnego modelu zerowego, a następnie wybieramy model o najmniejszej *p*-wartości (największej istotności).

- Takie podejście zaprezentowaliśmy 10 lat temu z Janem Mielniczukiem. Jest ono związane z bayesowskim wyborem modelu choćby dlatego, że odwrotność *p*-wartości pomnożonej przez rozmiar próby n, czyli $F = \frac{1}{np-\text{val}}$, zwany dalej czynnikiem Fishera, jest przybliżeniem „obiektywnego" czynnika Bayesa. W referacie podam przykład zastosowania F oraz opowiem o częstościowej interpretacji *p*-wartości.

1. Ostatnio, po analizie danych powstałych w ramach projektu oceny reprodukowalności wyników w różnych dziedzinach nauki, grupa wpływowych statystyków zaproponowała zmianę wprowadzonego przez Fishera standardowego progu istotności 0.05 na 0.005. Autorzy przekonują, że pozwoli to kilkukrotnie zwiększyć reprodukowalność. W referacie argumentuję, że taki sam wzrost reprodukowalności otrzymamy, gdy zamiast progu 0.005 użyjemy reguły

$$F > 2 \iff p - \text{val} < \frac{0.5}{n} \tag{*}$$

Jak wiadomo Fisher uzasadniał próg 0.05 na podstawie eksperymentów, w których $n \approx 10$, natomiast w projekcie reprodukowalności mamy zazwyczaj $n \approx 100$. Reguła (*) uogólnia zatem oba wymienione progi, a ponadto dodaje uzasadnienie bayesowskie: odrzucaj nulla, gdy alternatywa ma 2 razy większą szansę.

2. W klasyfikacji szukamy kierunku rozdzielającego dwie wielowymiarowe populacje X_0 i X_1 maksymalizując pewną odległość między nimi. Przykładem takiego postępowania jest maksymalizacja AUC. W testowaniu hipotez jest podobnie, ale mamy asymetrię w dostępie do danych, bo X_0 to populacja teoretyczna, a X_1 - empiryczna. Okazuje się, że gdy będziemy estymować AUC między teorią, a eksperymentem, a w roli kierunków dyskryminacji obsadzimy ilorazy wiarygodności, to najlepszy kierunek dyskryminacji daje wspomniane na wstępie kryterium minimalnej *p*-wartości. W ten sposób otrzymujemy częstościową interpretację *p*-wartości.

Variable selection in
high-dimensional binary regression

Wojciech Rejchel

Nicolaus Copernicus University in Toruń

We consider the binary regression model that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are i.i.d. random vectors such that $X_i \in \mathbb{R}^p$ and $Y_i \in \{0, 1\}$. In the talk we assume that

$$P(Y_i = 1 | X_i) = g(\beta' X_i), \quad i = 1, \ldots, n \qquad (*)$$

for some true parameter β and *unknown* function g. The goal is to find the set of relevant predictors $T = \{1 \leqslant j \leqslant p : \beta_j \neq 0\}$, even if the number of predictors p is much larger than the sample size n.

In the talk we propose a simple and computationally fast procedure solving this problem that is based on applying the standard Lasso for linear regression to the binary regression model (*). Statistical properties of this algorithm in variable selection are investigated.

Properties of uplift estimators for small samples

Krzysztof Rudaś[1,2], Szymon Jaroszewicz[2]

Politechnika Warszawska[1],
Instytut Podstaw Informatyki Polskiej Akademii Nauk[2]

Uplift modeling is an approach, which allows for predicting effect of an action (e.g. new marketing campaign or medical treatment). To achieve this we divide our population into two subgroups: treatment, which is subjected to the action and control on which no action is taken. Then we estimate difference between effects in treatment and control group. In [1] we introduced and analysed two basic methods of estimation: double and uplift regression. We also proposed new method, called corrected uplift regression, joining advantages of two previous approaches. We also calculated asymptotic distributions of these estimators.

In my presentation I will show basic properties of these three approaches. They can be used to explain behaviour of estimators for small samples, where difference of mean squared error is most significant.

Literatura

[1] K.Rudaś, S. Jaroszewicz (2018), *Linear regression for uplift modeling,* Data Mining and Knowledge Discovery, 32(5), 1275–1305

[2] K.Rudaś, S. Jaroszewicz (2019), *Shrinkage Estimators for Uplift Regression,* Proc. of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD'19)

Structure learning for Bayesian Networks via penalized likelihood method

Maryia Shpak

Instytut Matematyki UMCS / MIMUW

We consider a Gaussian Bayesian Network model and our goal is to learn the sparse structure of the network based on observed data. In essence, Bayesian Network is a graph probability model where the dependence structure is given by a Directed Acyclic Graph (DAG) and

$$f(X_1, \ldots, X_n) = \prod_i f(X_i | \text{Parents}(X_i)),$$

where $f(X_1, \ldots, X_n)$ is the joint density over all the variables in the network and $f(X_i | \text{Parents}(X_i))$ are conditional densities of variables X_i given their parents (meaning nodes in the graph from which there exist arrows towards the node X_i).

Our method consists of two stages. In the first stage we use partitioning algorithm proposed by [1]. In the second stage we build linear regression models, which agree with partitioning. Finally, selection of the final DAG is done by thresholded LASSO method.

Literatura

[1] J. Kuipers, G. Moffa (2017), *Partition MCMC for Inference on Acyclic Digraphs*, Journal of the American Statistical Association, vol. 112, pp. 282–299

Kolmogorov-Smirnov-like goodness-of-fit test with two-component test statistics

Piotr Sulewski

Pomeranian University in Słupsk

Any goodness-of-fit test (GoFT) has its unique measure of discrepancy between theoretical and sample distributions commonly called the test statistics. In pre-Monte Carlo era inventors of GoFTs had their hands tied. Formulas of test statistics had to be simple. This simplicity was an indispensable condition for test statistics distributions to be derived in an analytical way. It was the only way they had. In the Monte Carlo era inventors of GoFTs have their hands freed owing to Lilliefors. In that case the author of this paper gets to a far-reaching modification of the KS GoFT. The aim is, of course, to extract more information about the discrepancy mentioned above and to increase the power of the test. The proposed test statistics has two components. The first component further denoted Δ is, of course, the greatest absolute value of a difference between theoretical and empirical cumulative distribution functions. The second component denoted r is the order statistics at which the greatest difference is located. Both components, Δ and r, are random variables. The conditional distributions of the random variable Δ_r given r=r*,r*=1,2,...,n, where n is a sample size, were determined with the Monte Carlo method. Then this distribution served as bases for determining conditional critical values Δ_r^{crit}. The decision rule remains unchanged. The null hypothesis is accepted when $\Delta_{r*} < \Delta_{r*}^{crit}$.

Literatura

[1] H.W.Lilliefors (1967), *On the Kolmogorov-Smirnov test for normality with mean and variance unknown,* Journal of the American statistical Association 62(318), 399–402

[2] N. Smirnov (1948), *Table for estimating the goodness of fit of empirical distribution,* Annals of Mathematical Statistics 19, 279–281

Simulation study about the properties of MLE of separable covariance matrix.

Anna Szczepańska-Álvarez[1],
Adolfo Álvarez[2], Bogna Zawieja[1]

[1] Poznań University of Life Sciences, Poland

[2] Collegium Da Vinci, Poland

We analyse the properties of MLE of covariance matrix: $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$, where $\boldsymbol{\Sigma}$ is the unstructured matrix and $\boldsymbol{\Psi}$ is the partitioned matrix. In the considered case the explicit form of estimators does not exist and $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\Psi}}$ are determined iteratively. We present the convergence and bias of algorithms in cases where $25\%, 50\%$ and 75% of the structure of $\boldsymbol{\Psi}$ is known.

Literatura

[1] A Szczepańska-Álvarez, Ch. Hao, Y. Liang, D. von Rosen (2017), *Estimation equations for multivariate linear models with Kronecker structured covariance matrices*, Communications in Statistics-Theory and Methods. Vol 46, no 16, 7902-7915

Regularizations induced by filters and a discrepancy principle for Poisson inverse problems

Zbigniew Szkutnik

Wydział Matematyki Stosowanej, AGH, Kraków

The problem of inverse estimation of the intensity function of a Poisson process will be considered. Given observed a Poisson process (i.e. a random point measure) on an abstract space with the intensity function $g = \mathcal{K}f$, where $\mathcal{K}$ is a compact operator acting between some separable Hilbert spaces, the goal is to estimate the function of interest f. This is a general form of an ill-posed Poisson inverse problem. Ill-posedness requires some sort of regularization and a method of data-driven choice of some regularization parameter. A general form of the Morozov discrepancy principle, suitable for such problems, will be discussed, along with a family of regularizations induced by filters, that includes, e.g., well-known Tikhonov, truncated SVD and Landweber methods. Unlike other methods described in literature, we shall work in the infinite dimensional setting with no need of discretization. Some asymptotic properties of the resulting estimators, as well as applications to some specific stereological problems, formulated as Poisson inverse problems, will be presented.

Generalized aging intensity functions with applications to lifetime analysis

Magdalena Szymkowiak

Institute of Automation and Robotics,
Poznan University of Technology

We define and study a family of generalized aging intensity functions to be used in reliability analysis of items and compound structures. The defined functions allow us to measure and compare aging tendencies of lifetime random variables in various time scalings. They describe quantitatively the differences of random lifetime variabilities ascertained by the star ordering and IFRA (increasing failure rate average) condition. Moreover, we present characterizations of lifetime distributions by means of the generalized aging intensity functions. Finally, we propose an idea of goodness-of-fit tests for compound parametric hypotheses applying nonparametric estimates of the support dependent generalized aging intensities determined for generated and real data.

Literatura

[1] K. Danielak and T. Rychlik (2003), *Sharp bounds for expectations of spacings from DDA and DFRA families,* Statistics and Probability Letters, 65, 303–316

[2] A.W. Marshall and I. Olkin (2007), *Life Distributions: Structure of Nonparametric, Semiparametric and Parametric Models,* Springer Series in Statistics, New York

[3] M. Szymkowiak (2019), *Measures of aging tendency,* Journal of Applied Probability, 56(2), 358–383

Geometry for uniqueness and model selection for LASSO and SLOPE estimators

Patrick Tardivel

Instytut Matematyczny, Uniwersytet Wrocławski

LASSO and SLOPE estimators are defined as minimizers of the penalized residual sum of squares where the penalty is respectively the l_1 norm and the SLOPE norm (a generalization of the l_1 norm). Because in both cases the objective function is not strictly convex the uniqueness of the minimizer is not obvious. For both LASSO and SLOPE estimators we give a necessary and sufficient condition under which the uniqueness of the minimizer occurs. In addition, we show how a geometric condition involving the hypercube (resp. sign permutahedron) gives insights about the accessible models for LASSO estimator (resp. SLOPE estimator).

Learning classifier chains under feature extraction budget

Paweł Teisseyre

Institute of Computer Science, Polish Academy of Sciences

We study a problem of learning classifier chains in multi-label classification under a constraint on the number of features. Such constraints are important in domains where including redundant features is associated with costs and risk of negative effects. For example in medical diagnosis, features correspond to diagnostic tests which may be very expensive or may cause negative effects. Standard classifier chains tend to select too many features, when feature selection method is embedded in base learner, which is due to the fact that selection is performed separately for each of the models in the chain. We present the novel method which use matrix regularization to select features shared across the models. The experiments carried out on a large clinical database show that the proposed method achieves higher accuracy than related methods when the number of features is limited.

Literatura

[1] P. Teisseyre, D. Zufferey, M. Slomka (2019), *Cost-sensitive classifier chains: selecting low-cost features in multi-label classification,* Pattern Recognition, Volume 86, 290–319,

Prawie pewna wersja centralnego twierdzenia granicznego dla sum zmiennych losowych z losowymi indeksami

Joanna Wasiura-Maślany

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

W ostatnich latach pojawia się wiele prac poświęconych tak zwanym twierdzeniom granicznym ze zbieżnością prawie na pewną. W twierdzeniach tego typu, nazywanych również prawie pewną wersją centralnych twierdzeń granicznych, rozważane są dwa typy zbieżności ciągów zmiennych losowych: zbieżność według rozkładów i zbieżność z prawdopodobieństwem jeden. Tego typu zbieżności odgrywają ważną rolę w rachunku prawdopodobieństwa i statystyce matematycznej. Prawie pewne wersje twierdzeń granicznych łączą te dwa typy zbieżności. Z tego też powodu, prawie pewne wersje twierdzeń granicznych można rozpatrywać z jednej strony jako twierdzenia dotyczące słabej zbieżności ciągu miar, a z drugiej strony jako mocne prawa wielkich liczb.

Głównym celem mojego referatu jest przedstawienie twierdzeń granicznych ze zbieżnością prawie na pewną dla sum niezależnych zmiennych losowych z losowymi indeksami. Główne wyniki uogólniają i rozszerzają twierdzenia dotyczące tej tematyki, które zostały opublikowane w ostatnich latach przez innych autorów.

Data-driven tests for trend under right-censored model

Grzegorz Wyłupek

Institute of Mathematics, University of Wrocław

Let X_{li}^0, $i = 1, \ldots, n_l$, be the independent, identically distributed, non-negative, un-observable survival times from the population with continuous cumulative distribution function F_l and related survival function S_l. Let U_{li}, $i = 1, \ldots, n_l$, be the corresponding independent, identically distributed, non-negative, unobservable censoring times from the population with continuous cumulative distribution function G_l and related survival function C_l. We consider the k-sample problem, therefore, $l = 1, \ldots, k$, and, we assume that all the random variables are independent. Since the data are right-censored, we only observe

$$X_{li} = \min\{X_{li}^0, U_{li}\} \quad \text{and} \quad \Delta_{li} = \mathbb{1}(X_{li}^0 \leqslant U_{li}), \; i = 1, \ldots, n_l, \; l = 1, \ldots, k.$$

The censoring status $\Delta_{li} = 1$ says that X_{li} is the original survival time X_{li}^0, otherwise $\Delta_{li} = 0$, resulting in $X_{li} = U_{li}$, which means that X_{li}^0 has been censored.

We are interested in detection of the upward trend, in the strong sense, between the consecutive survival functions, i.e. the relation,

$$S_1(t) \leqslant \cdots \leqslant S_k(t), \quad \text{for all} \quad t \in [0, +\infty),$$

with at least one strict inequality for at least one t. We denote it simply as

$$S_1 \leqslant \cdots \leqslant S_k, \; S_i \lneqq S_j, \; 1 \leqslant i < j \leqslant k.$$

Actually, we are going to test the hypothesis

$$\mathcal{H} : \text{the upward trend } S_1 \leqslant \cdots \leqslant S_k, \; S_i \lneqq S_j, \; 1 \leqslant i < j \leqslant k, \text{ does } \underline{\text{not}} \text{ hold,}$$

against the alternative

$$\mathcal{A}_+ : \text{the upward trend } S_1 \leqslant \cdots \leqslant S_k, \; S_i \lneqq S_j, \; 1 \leqslant i < j \leqslant k, \text{ holds,}$$

in the presence of infinite-dimensional nuisance parameters $G_1, \ldots, G_k$.

In the talk, we introduce a new test for the problem $(\mathcal{H}, \mathcal{A}_+)$, as well as discuss its asymptotic properties.

Confidence interval for odds ratio

Wojciech Zieliński

Department of Econometrics and Statistics, SGGW
Nowoursynowska 159, PL-02-787 Warszawa
e-mail: wojciech_zielinski@ssggw.pl
http://wojtek.zielinski.statystyka.info

In medical research two treatments are compared on the basis on binary data. The very important indicator is odds ratio. In applications widely used is asymptotic confidence interval. Unfortunately this confidence interval has some disadvantages.
In the paper a new confidence interval for odds ratio is constructed.

List of Participants

1. **dr Ewa Bakinowska**
   Politechnika Poznańska
   `ewa.bakinowska@put.poznan.pl`

2. **dr Jacek Bednarz**
   Instytut Ekonomii i Finansów Katolicki Uniwersytet Lubelski Jana Pawła II
   `bednarz@kul.pl`

3. **dr hab. Mariusz Bieniek**
   Instytut Matematyki, Uniwersytet Marii Curie-Skłodowskiej w Lublinie
   `mariusz.bieniek@umcs.lublin.pl`

4. **dr hab. Małgorzata Bogdan**
   Uniwersytet Wrocławski
   `malgorzata.bogdan@uwr.edu.pl`

5. **dr Katarzyna Brzozowska-Rup**
   Politechnika Świętokrzyska
   `brzozows@poczta.fm`

6. **dr hab. Monika Budzyńska**
   Instytut Matematyki, Uniwersytet Marii Curie-Skłodowskiej w Lublinie
   `monika.budzynska@umcs.lublin.pl`

7. **mgr Tomasz Cąkała**
   Uniwersytet Warszawski
   `tc360950@students.mimuw.edu.pl`

8. **dr Mariola Chrzanowska**
   Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
   `mariola_chrzanowska@sggw.pl`

9. **dr Bogdan Ćmiel**
   Wydział Matematyki Stosowanej, Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
   `cmielbog@gmail.com`

10. **dr hab. Anna Czapkiewicz**
    AGH Akademia Górniczo-Hutnicza
    `gzrembie@cyf-kr.edu.pl`

11. **dr hab. Antoni Leon Dawidowicz**
    Uniwersytet Jagielloński
    `Antoni.Leon.Dawidowicz@im.uj.edu.pl`

12. **mgr Alicja Dembczak-Kołodziejczyk**
    Uniwersytet Opolski
    `alicja.dembczak@uni.opole.pl`

13. **dr Orest Dorosh**
    Narodowe Centrum Badań Jądrowych
    `sub3kv@gmail.com`

14. **dr Diana Dziewa-Dawidczyk**
    WZIiM SGGW
    `diana_dziewa_dawidczyk@sggw.pl`

15. **mgr Sylwia Filas-Przybył**
    Urząd Statystyczny w Poznaniu
    `s.filas@stat.gov.pl`

16. **dr hab. Katarzyna Filipiak**
    Instytut Matematyki, Politechnika Poznańska
    `katarzyna.filipiak@put.poznan.pl`

17. **dr hab. Konrad Furmańczyk**
    Instytut Informatyki Technicznej SGGW
    `konfur@wp.pl`

18. **dr Barbara Gołubowska**
    Instytut Podstawowych Problemów Techniki PAN
    `bgolub@ippt.pan.pl`

19. **dr Agnieszka Goroncy**
    Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w Toruniu
    `gemini@mat.umk.pl`

20. **prof. Przemysław Grzegorzewski**
    Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska
    `pgrzeg@mini.pw.edu.pl`

21. **prof. Tadeusz Inglot**
    Wydział Matematyki, Politechnika Wrocławska
    `Tadeusz.Inglot@pwr.edu.pl`

22. **mgr Malwina Janiszewska**
    Uniwersytet Przyrodniczy w Poznaniu, Katedra Metod Matematycznych
    i Statystycznych
    `malwina.janiszewska@up.poznan.pl`

23. **dr Krzysztof Jasiński**
Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w Toruniu
`krzys@mat.umk.pl`

24. **Wei Jiang (PhD Student)**
CMAP, Ecole Polytechnique, Paris

25. **mgr Mateusz John**
Instytut Matematyki, Politechnika Poznańska
`mateusz.john@put.poznan.pl`

26. **dr hab. Alicja Jokiel-Rokita**
Wydział Matematyki, Politechnika Wrocławska
`alicja.jokiel-rokita@pwr.edu.pl`

27. **dr Joanna Karłowska-Pik**
Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w Toruniu
`joanka@mat.umk.pl`

28. **dr Tomasz Klimanek**
Urząd Statystyczny w Poznaniu
`t.klimanek@stat.gov.pl`

29. **dr Marek Kociński**
Institute of Economics and Finance Warsaw University of Life Sciences
SGGW
`marek_kocinski@sggw.pl`

30. **mgr Michał Kos**
Uniwersytet Wrocławski
`michal.kos@uwr.edu.pl`

31. **dr hab. Wasyl Kowalczuk**
Instytut Podstawowych Problemów Techniki PAN
`vkoval@ippt.pan.pl`

32. **prof. Mirosław Krzyśko**
Wydział Matematyki i Informatyki, Uniwersytet im. Adama Mickiewicza
w Poznaniu
`mkrzysko@amu.edu.pl`

33. **mgr Paweł Kurasiński**
Instytut Matematyki UMCS w Lublinie
`pawelkurasinski1@gmail.com`

34. **prof. Teresa Ledwina**
Instytut Matematyczny PAN
`ledwina@impan.pl`

35. **mgr Małgorzata Łazęcka**
Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska
m.lazecka@mini.pw.edu.pl

36. **prof. Ryszard Magiera**
Politechnika Wrocławska
ryszard.magiera@pwr.edu.pl

37. **dr Mariola Marciniak**
UKW w Bydgoszczy
znowurobota@gmail.com

38. **dr hab. Błażej Miasojedow**
Instytut Matematyki Stosowanej i Mechaniki, Uniwersytet Warszawski
bmiasojedow@gmail.com

39. **dr hab. Andrzej Michalski**
Katedra Matematyki, Uniwersytet Przyrodniczy we Wrocławiu
apm.mich@gmail.com

40. **mgr Adam Mieldzioc**
Uniwersytet Przyrodniczy w Poznaniu
adam.mieldzioc@mail.up.poznan.pl

41. **prof. Jan Mielniczuk**
Instytut Podstaw Informatyki PAN i Wydział Matematyki i Nauk Informacyjnych PW
miel@ipipan.waw.pl

42. **mgr Grzegorz Mika**
Akademia Górniczo- Hutnicza im. S. Staszica w Krakowie
grzegorz.mika@agh.edu.pl

43. **mgr Monika Mokrzycka**
Instytut Genetyki Roślin PAN
mmok@igr.poznan.pl

44. **Remigiusz Musiałkiewicz**
WZIiM SGGW
musialkiewiczremigiusz@gmail.com

45. **prof. Wojciech Niemiro**
Wydział Matematyki Stosowanej i Mechaniki, UW i Wydział Matematyki i Informatyki, UMK w Toruniu
wniemiro@gmail.com

46. **dr hab. Arkadiusz Orłowski**
Instytut Informatyki Technicznej SGGW
arkadiusz_orlowski@sggw.pl

47.   **Kacper Paczutkowski**
WZIiM SGGW
`kacper.pki@outlook.com`

48. **mgr Luiza Pańczyk**
Instytut Matematyki, Uniwersytet Marii Curie-Skłodowskiej w Lublinie
`luiza.panczyk@o2.pl`

49. **dr hab. Piotr Pawlas**
Uniwersytet Marii Curie-Skłodowskiej w Lublinie
`piotrpawlas@interia.pl`

50. **dr hab. Piotr Pokarowski**
Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski
`pokar@mimuw.edu.pl`

51. **dr Wojciech Rejchel**
Wydział Matematyki i Informatyki, Uniwersytet Mikołaja Kopernika w Toruniu
`iggypop@mat.umk.pl`

52. **prof. Grzegorz A. Rempała**
The Ohio State University
`rempala.3@osu.edu`

53. **mgr Krzysztof Rudaś**
Politechnika Warszawska
`krzysztof.rudas@gmail.com`

54. **mgr Maryia Shpak**
Instytut Matematyki, Uniwersytet Marii Curie-Skłodowskiej w Lublinie
`shpak.maria@gmail.com`

55. **dr Piotr Sulewski**
Instytut Matematyki, Akademia Pomorska w Słupsku
`piotr.sulewski@apsl.edu.pl`

56. **dr Anna Szczepańska-Álvarez**
Uniwersytet Przyrodniczy w Poznaniu Katedra Metod Matematycznych
i Statystycznych
`sanna6@wp.pl`

57. **prof. Zbigniew Szkutnik**
Wydział Matematyki Stosowanej, Akademia Górniczo-Hutnicza im. Stanisława
Staszica w Krakowie
`szkutnik@agh.edu.pl`

58. **dr hab. Magdalena Szymkowiak**
Instytut Automatyki i Robotyki, Politechnika Poznańska
`magdalena.szymkowiak@put.poznan.pl`

59. **dr Patrick Tardivel**
Instytut Matematyczny Uniwersytet Wrocławski
`tardivel@math.uni.wroc.pl`

60. **dr Paweł Teisseyre**
Instytut Podstaw Informatyki PAN
`teisseyrep@ipipan.waw.pl`

61. **mgr Joanna Wasiura-Maślany**
Uniwersytet Marii Curie-Skłodowskiej w Lublinie i Katolicki Uniwersytet
Lubelski
`jwaaa@wp.pl`

62. **dr Grzegorz Wyłupek**
Instytut Matematyczny, Uniwersytet Wrocławski
`wylupek@math.uni.wroc.pl`

63. **dr Olga Zajkowska**
SGGW
`o.zajkowska@gmail.com`

64. **prof. Wojciech Zieliński**
Katedra Ekonometrii i Statystyki SGGW
`wojciech_zielinski@sggw.pl`

65. **mgr Agnieszka Zięba**
Politechnika Warszawska
`Agnieszka.Zieba@mini.pw.edu.pl`

66. **dr Piotr Zwiernik**
Universitat Pompeu Fabra
`piotr.zwiernik@upf.edu`