PROBASTAT 2024

The Eighth International Conference on Mathematical Statistics

ABSTRACTS

May 20 - 24, 2024, Smolenice Castle, Slovak Republic

PROBASTAT 2024
May 20 – 24, 2024, Smolenice Castle, Slovakia

ORGANIZERS
Institute of Measurement Science of the Slovak Academy of Sciences
Faculty of Mathematics, Physics and Informatics, Comenius University Bratislava
Mathematical Institute of the Slovak Academy of Sciences

GOLD SPONSORS
Slovenská sporiteľňa, a.s.
Zurich Insurance Company Ltd
EuroPainClinics

SILVER SPONSOR
KPMG Slovensko spol. s r. o.

MEDIA PARTNERS
International Conference Alerts
COMS – Conference Listings

EDITORS
Samuel Rosa & Andrej Dvurečenskij & Viktor Witkovský

PUBLISHER
Institute of Measurement Science
Slovak Academy of Sciences, Bratislava, Slovak Republic
Bratislava, May 2024

Contents

Abstracts

Some extensions of the Schröter distribution family

Friday I. Agu* Ján Mačutek†

In traditional insurance computations, aggregate claim amounts are often determined using the convolution approach. However, this method faces challenges, particularly when dealing with a large number of claims and claim severity. To address this issue, the Schröter distribution family has emerged as a valuable alternative. This family has a fixed positive probability at 0 and provides a more accurate and comprehensive representation of both the number of claims and claim severity, particularly when the number of claims is discrete. Despite its usefulness, the Schröter distribution family may not fully capture the characteristics of claim amounts truncated at 1, a common practice in insurance where the insurance company may be interested in the number of events that have led to a claim. If a claim has been reported, the least observed claim amount is truncated at 1. In risk theory, truncated distributions play a critical role in modeling claim severity and the time intervals between successive claims. These distributions are essential tools for insurers and actuaries, enabling them to gain insights into the frequency and severity of potential losses, thereby facilitating policy pricing and risk management. In this study, we investigated the truncated Schröter distributions using a combination of truncated probability distribution and convolution approaches. Specifically, we focus on exploring the convolutions of truncated and non-truncated versions of distributions from the Schröter family, aiming to enhance our understanding of truncated claim amounts in insurance.

Acknowledgments

This study was supported by the Slovak Scientific Grant Agency VEGA, under grants **VEGA 2/0120/24 (Ján Mačutek/Friday I. Agu)** and **APP0515 (Friday I. Agu)**.

References

[1] Schröter, K. J. (1990). On a family of counting distributions and recursions for related compound distributions. *Scandinavian Actuarial Journal* **1990**(2-3), 161–175.

[2] Agu, F. I., Mačutek, J., and Szűcs, G. (2023). A simple estimation of parameters for discrete distributions from the Schröter family. *Statistika: Statistics and Economy Journal* **103**(2), 246–251.

[3] Kreer, M., Kizilersü, A., Thomas, A. W., and dos Reis, A. D. E. (2015). Goodness-of-fit tests and applications for left-truncated Weibull distributions to non-life insurance. *European Actuarial Journal* **5**, 139–163.

*Mathematical Institute of the Slovak Academy of Sciences, Štefánikova 898/49, 814 73 Staré Mesto, Bratislava, agu1@uniba.sk

†Mathematical Institute of the Slovak Academy of Sciences, Štefánikova 898/49, 814 73 Staré Mesto, Bratislava, jmacutek@yahoo.com

Data driven modelling of vertical profiles of atmospheric radiation

Jaromír Antoch*

In the Czech Hydrometeorological Institute there exists a unique set of meteorological measurements consisting of the values of vertical atmospheric levels of beta and gamma radiation. An important task required by meteorologists is to assess an upper confidence bound for the beta and gamma counts that could be used for warning purposes.

The primary goal of our talk is to improve understanding of the distribution of environmental radiation based on the measurements of the vertical radioactivity profiles by a sonde system. To that purpose we use a nonlinear quantile regression, which is gradually evolving into a comprehensive approach to the statistical analysis of linear and nonlinear response models for conditional quantile functions.

It is well known that the calculation of the nonlinear regression quantiles is from the computational point of view a very complicated problem. Therefore, we will show under which assumptions the quantiles can be calculated in a much simpler way directly from the estimated intensity of the observed radioactivity process without the need of numerical iterations.

The last part of our talk will concentrate on a combined approach to the quantiles when we have available a parametric model, and when the quantiles are estimated from the corresponding residuals either nonparametrically or parametrically.

Keywords: Richards' growth curve, nonlinear regression, nonhomogeneous Poisson process, vertical atmospheric radiation profiles, RADAC data.

Acknowledgments

Lecture is based on joint work with D. Hlubinka.

*Charles University in Prague, Department of Probability and Mathematical Statistics, Sokolovská 83, CZ – 186 75 Praha 8 – Karlín, Czech Republic, antoch@karlin.mff.cuni.cz

Characteristic function and moment generating function of multivariate folded normal distribution

Matej Benko* Zuzana Hübnerová† Viktor Witkovský‡

Our contribution contains a derivation of the characteristic function of the multivariate folded normal distribution. It is obtained by computation of the moment-generating function using expression in terms of normal probability and cumulative density function. The folded normal distribution arises when only the magnitudes of the elements, but not the signs, of a normally distributed random vector are of interest.

The work corrects the errors in derivation from [1] pointed out in [2] by showing an inconsistency in the covariance for the two dimensional case. We show that our results are correct from this point of view. We also present simulations to visualize that our characteristic function of sum of absolute values of elements on multivariate normal vector tends to the empirical one.

Folded distribution has important applications in various fields and analytically expressed characteristic function allows us to get many properties of it.

Acknowledgments

This work was supported by grant No. FSI-S-23-8161. Participation of the first author at Probastat was supported by the Czech Statistical Society.

References

[1] Chakraborty, A. K. and Chatterjee, M. (2013). On multivariate folded normal distribution. *Sankhya B* **75**(1) 1–15. https://doi.org/10.1007/s13571-013-0064-5

[2] Murthy, G. S. R. and Faraway, J. J. (2015). A note on multivariate folded normal distribution. *Sankhya B* **77** 108–113. https://doi.org/10.1007/s13571-014-0092-9

*Institute of Mathematics, Faculty of Mechanical Engineering, Brno University of Technology, Czech Republic, Matej.Benko@vutbr.cz

†Institute of Mathematics, Faculty of Mechanical Engineering, Brno University of Technology, Czech Republic, hubnerova@fme.vutbr.cz

‡Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia, witkovsky@savba.sk

Probability distributions and calculations for Hake's ratio statistics in measuring effect size

Jozef Hanč* Dominik Borovský† Martina Hančová‡

Today, ratio statistics and distributions play a very important role not only in linear regression theory and metrology but also in diverse application areas. Particular examples [1, 2] are nuclear physics (mass-to-energy ratios), operations research and engineering (safety factors in design, signal-to-noise ratios, radars' distribution), econometrics (economic indicators), biostatistics (enzyme activity, lifespan of red blood cells, medical study ratios), genetics (Mendelian inheritance ratios), and the food and pharmaceutical industries (digestibility measures, component ratios of foods or drugs), as well as meteorology (target-to-control precipitation ratios).

In our study, using open data science tools [3], we investigate the statistical properties and exact probability calculations of the Hake normalized gain, widely used in the physics education community as a measure of effect size and educational effectiveness in introductory physics courses [4]. Our computational procedure involves the application of the Mellin integral transform [5, 6] and its numerical integration using the relatively unfamiliar but highly efficient double exponential quadrature [7, 3].

Knowing that Hake's statistic is a ratio of normal variables, we compared our results with the well-known analytic formulas for its distribution [5] as well as with recent approaches using Clenshaw-Curtis or Fejér quadratures [6]. Our investigation offers not only valuable insights and a better understanding of the Hake ratio for educational researchers but also our proposed numerical approach, in its generality, speed, and precision, shows potential for fast data analysis based on exact probability distributions of products and quotients of random variables in fields like multidimensional statistics or measurement uncertainty analysis in metrology.

Acknowledgments

This work was supported by the Slovak Research and Development Agency under the Contract No. APVV-21-0369, No. APVV-21-0216 and by the Slovak Scientific Grant Agency VEGA under grant VEGA 1/0585/24.

*Institute of Physics, P. J. Šafárik University, Slovakia, jozef.hanc@upjs.sk

†Institute of Physics, P. J. Šafárik University, Slovakia, dominik.borovsky@student.upjs.sk

‡Institute of Mathematics, P. J. Šafárik University, Slovakia, martina.hancova@upjs.sk

References

[1] Nadarajah, S., and Kotz, S. (2006). A note on the ratio of normal and Laplace random variables. *Stat. Methods Appl.* **15**(2) 151–158.

[2] Celano, G., et al. (2014). Statistical Performance of a Control Chart for Individual Observations Monitoring the Ratio of Two Normal Variables. *Qual. Reliab. Eng. Int.* **30**(8) 1361–1377.

[3] Hančová, M., Gajdoš, A., and Hanč, J. (2022). A practical, effective calculation of gamma difference distributions with open data science tools. *J. Stat. Comput. Simul.* **92**(11) 2205–2232.

[4] Nissen, J. M., et al. (2018). Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Phys. Rev. Phys. Educ. Res.* **14**(1) 010115.

[5] Springer, M. D. (1979). *The Algebra of Random Variables.* New York: Wiley.

[6] Jaroszewicz, S., and Korzeń, M. (2012). Arithmetic Operations on Independent Random Variables: A Numerical Approach. *SIAM J. Sci. Comput.* **34**(3) A1241–A1265.

[7] Mori, M. (2005). Discovery of the Double Exponential Transformation and Its Developments. *Res. Inst. Math. Sci.* **41**(4) 897–935.

On the convergence rate for the longest at most T-contaminated runs of heads

István Fazekas* Borbála Fazekas† László Fórián‡

In this paper, we shall consider the well-known coin tossing experiment. Erdős and Rényi [3] obtained results concerning the length of the longest pure head run. A head run containing T tails is called a T-interrupted (in other words T-contaminated) head run. Erdős and Révész [4] presented almost sure limit results for the length of the longest T-interrupted head run. Földes [5] obtained asymptotic results for the distribution of the length of the longest T-interrupted head run using Sevastyanov's Poisson limit theorem. Arratia, Gordon and Waterman [1] applied Poisson approximation to obtain accompanying distributions for the length of the longest T-interrupted head run.

However, numerical studies show, that for $T > 0$, the rates of convergences are slow both in the result of Földes [5] and the result by Arratia, Gordon and Waterman [1]. In this paper, we offer new accompanying distributions for the length of the longest at most T-interrupted head run. To obtain the convergence rate, we apply the Poisson approximation given in [1]. For values of $T = 1, 2$, a proof can be given by using a powerful lemma of Csáki, Földes and Komlós [2]. We apply that lemma for a certain problem of trinary experiment, too. We also present simulation results showing the performance of our theorems.

References

[1] Arratia, R., Gordon, L. and Waterman, M. S. (1990). The Erdős-Rényi law in distribution, for coin tossing and sequence matching. *Ann. Statist.* **18**, no. 2, 539–570.

[2] Csáki, E., Földes, A. and Komlós, J. (1987). Limit theorems for Erdős-Rényi type problems. *Studia Sci. Math. Hungar.* **22**, 321–332.

[3] Erdős, P. and Rényi, A. (1970). On a new law of large numbers. *J. Analyse Math.* **23**, 103-111.

[4] Erdős, P. and Révész, P. (1977). On the length of the longest head-run. *Topics in information theory (Second Colloq., Keszthely, 1975)*, pp. 219–228. Colloq. Math. Soc. János Bolyai, Vol. 16, North-Holland, Amsterdam.

[5] Földes, A. (1979) The limit distribution of the length of the longest head-run. *Period. Math. Hungar.* **10**, no. 4, 301–310.

*University of Debrecen, Hungary, fazekas.istvan@inf.unideb.hu
†University of Debrecen, Hungary, borbala.fazekas@science.unideb.hu
‡University of Debrecen, Hungary, forian.laszlo@inf.unideb.hu

On estimation of block covariance matrix with linearly structured blocks with the use of quadratic subspace properties

Katarzyna Filipiak* Augustyn Markiewicz†
Malwina Mrowińska‡

The aim of this talk is to propose estimation methods of linearly structured partitioned covariance matrix. In contrast to well known structures, as e.g. block compound symmetry, we allow the diagonal blocks of covariance matrix to be of different dimensions. Such a situation is very common in real experiments, especially when heatmaps are used to discover the structure of covariance matrix.

We focus on two cases: when the covariance structure belongs to a quadratic subspace and when it is not the case. In the first situation we show how to estimate covariance structure with the use of projections of sample covariance matrix on respective subspaces. In the second case, since projection does not preserve definiteness, we improve obtained estimates using shrinking method with respective target matrix belonging to some quadratic space.

Acknowledgments

This work was supported by the Poznan University of Technology under grants no. 0213/SBAD/0118 (K. Filipiak) and 0213/SBAD/0119 (M. Mrowińska).

*Poznan University of Technology, Poznań, Poland, katarzyna.filipiak@put.poznan.pl
†Poznan University of Life Sciences, Poznań, Poland, augustyn.markiewicz@up.poznan.pl
‡Poznan University of Technology, Poznań, Poland, malwina.mrowinska@put.poznan.pl

The rocky start of the probability calculus

Christian Genest*

The probability calculus is rooted in games of chance, and arose in the mid 1600s in response to a question raised by an inveterate gambler, or at least so the story goes. In this talk, I will give an overview of the real motivations, major contributions, and significant interactions between some of the key players in the development of this theory in the sesquicentennial period before the publication of Laplace's Analytic Theory of Probability in 1812.

*McGill University, Montréal (Québec) Canada, christian.genest@mcgill.ca

Dynamic graphics for high-density EEG data using R

Zdeňka Geršlová* Martin Lamoš† Stanislav Katina‡

The high-density electroencephalography (HD-EEG) is one of the non-invasive techniques to display human brain activity. During the entire EEG data analysis process, a number of graphical methods are commonly used, including, for example, curves (time domain) or topographical maps (space domain). However, there are still limits in exploring the complex time-space character of the data through only static charts.

The presentation introduces some possibilities of making dynamic and interactive graphics in EEG filed using statistical software R. The interactive elements can be applied at every step of the data analysis from preprocessing and exploratory analysis to model results. An animation enables to follow the change in brain activity (on the head model or on its projection) both in time and space.

The developed methods are illustrated on the examples of HD-EEG recordings during a simple visual motor task in Parkinson's disease patients and a healthy controls. Subjects were instructed to press the response button in case of target visual stimulus presentation. The activity of the brain motor circuits are then evaluated and compared between subject groups.

Acknowledgments

This work was supported by Masaryk University under Grant MUNI/A/1457/2023 and Czech Health Research Council AZV NU21-04-00445.

*Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic, gerslovaz@math.muni.cz

†Brain and Mind Research Program, Central European Institute of Technology, Masaryk University, Brno, Czech Republic, martin.lamos@ceitec.muni.cz

‡Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic, katina@math.muni.cz

Variance in gene regulatory networks: the impact of system volume on network stability

Laura Hajzoková[1,2*] Pavol Bokes[2†]

Gene expression is heterogeneous among genetically identical cells and dynamically fluctuates within them. Mutual interactions of genes form gene regulatory networks that have a naturally stochastic character. In this context, we observe the stability of the protein cascades, depending on the network architecture. Furthermore, unlike deterministic modeling of gene regulatory networks, stochastic modeling allows the system volume parameter to enter the system and influence the network's variability. For large systems, in terms of the number of molecules, the system exhibits almost deterministic behavior. However, in many real-life situations, chemical reactions happen at low concentrations, where the deterministic approach usually fails. By varying the system volume parameter, we explore the stability of a synthetic gene regulatory cascade for different total numbers of molecules and through stochastic simulations, we endeavor to pinpoint the transition regions of stable states.

Acknowledgments

The research was supported by the Czech Ministry of Education, Youth and Sports and the Slovak Research and Development Agency through the Inter-Excellence II program, project LUASK22008/SK-CZ-RD-21-0109.

References

[1] Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions *The Journal of Physical Chemistry* **81** 2340-2361

[2] Hooshangi, S., Thiberge, S. and Weiss, R. (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade *Proceedings of the National Academy of Sciences of the USA* **102** 3581-3586

[3] Thattai, M. and van Oudenaarden, A. (2001) Intrinsic noise in gene regulatory networks *Proceedings of the National Academy of Sciences of the USA* **98** 8614-8619

[4] Zheng, X. and Tao, Y. (2008) Additivity of noise propagation in a protein cascade *J. Chem. Phys.* **128** 165104

[1]Institute of Measurement Science, Slovak Academy of Sciences

[2]Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava

*laura.hajzokova@savba.sk

†pavol.bokes@fmph.uniba.sk

Consistent distribution-free and affine-invariant tests for the validity of independent component models

Marc Hallin*

We propose a family of tests of the validity of the assumptions underlying independent component analysis methods. The tests are formulated as L2–type procedures based on characteristic functions and involve weights; a proper choice of these weights and the estimation method for the mixing matrix yields consistent and affine-invariant tests. Due to the complexity of the asymptotic null distribution of the resulting test statistics, implementation is based on permutational and resampling strategies. This leads to distribution-free procedures regardless of whether these procedures are performed on the estimated independent components themselves or the componentwise ranks of their components. A Monte Carlo study involving various estimation methods for the mixing matrix, various weights, and a competing test based on distance covariance is conducted under the null hypothesis as well as under alternatives. A real-data application demonstrates the practical utility and effectiveness of the method.

Acknowledgments

Based on joint work with Simos G. Meintanis (National and Kapodistrian University of Athens, Greece) and Klaus Nordhausen (University of Jyväskylä, Finland). Supported by the Czech Science Foundation grants GAČR22036365 and GA24-10078S (Marc Hallin).

*Université Libre de Bruxelles, Belgium, mhallin@ulb.ac.be

Systematic design methodology in optimizing simulation & computational algorithms

Martina Hančová* Jozef Hanč† Dominik Borovský‡

At present, computer simulations have become a cornerstone in advancing and managing nearly every field of human activity. They are not only used for investigating various phenomena but, in many cases, have replaced traditional experiments, providing results that would otherwise be unattainable. For example, the results of deep learning models behind AI chatbots like ChatGPT (Zhao et al 2023) are so complex that only benchmarking simulation studies can determine their quality and behavior. However, even though we have solid mathematical theories behind classical computational algorithms, such as exact computations of probability using some integral transforms (Hancova et al 2022), it is necessary to validate them through credible simulations. But what justifies our trust in these simulations, and to what extent can we generalize the results from them (Beisbart et al 2019), e.g., when making life-and-death decisions based on stochastic models for a kidney transplantation system (Cechlarova et al 2021)?

In this contribution, we describe, in detail and through specific examples, one of the current approaches to addressing this issue, which is not yet widespread among statisticians, data scientists, and engineers. This approach uses the concept of the systematic design, introduced over a decade ago (Lorscheid et al 2012). The Systematic Design of Experiments (DOE) can be utilized for screening system variables (identifying the most crucial variables), conducting detailed examinations of their impacts, optimizing systems by discovering combinations of independent variables that lead to optimal responses and enhancing system robustness where appropriate settings of input levels limit the impact of noise and uncontrollable inputs. Finally, systematic design is also effective in evaluating the correctness and efficiency of the mathematical methods and their corresponding computational algorithms.

Acknowledgments

This work was supported by the Slovak Research and Development Agency under the Contract No. APVV-21-0369, No. APVV-21-0216 and by the Slovak Scientific Grant Agency VEGA under grant VEGA 1/0585/24.

*Institute of Mathematics, P. J. Šafárik University, Slovakia, jozef.hanc@upjs.sk

†Institute of Physics, P. J. Šafárik University, Slovakia, martina.hancova@upjs.sk

‡Institute of Physics, P. J. Šafárik University, Slovakia, dominik.borovsky@student.upjs.sk

The challenge of optimal exact design of experiment

Radoslav Harman*

The field of optimal experimental design traditionally focuses on the so-called "approximate" experimental designs, which specify a finite set of experimental conditions and the proportions of trials performed at each of these conditions. The major advantage of the approximate design of experiments is that it can utilize theoretical and numerical tools from convex optimization.

In practical applications, however, we require "exact" experimental designs, which determine a finite sequence of experimental conditions for the trials. An exact design can be derived from an approximate design through heuristics known as "rounding" methods. Unfortunately, the outcome of these rounding methods is often a suboptimal exact design, even when applied to an optimal approximate design.

In this talk, we will first define and review the integer optimization problem of optimal exact design for statistical models with uncorrelated observations, focusing on its theoretical and computational complexity. Next, we will survey various approaches to the numerical computation of optimal exact designs. In more detail, we will discuss methods based on mixed-integer conic mathematical programming formulations. Finally, we will demonstrate these methods on several challenging problems of exact optimal designs under non-standard experimental constraints.

*Comenius University Bratislava, Slovakia, radoslav.harman@fmph.uniba.sk

One sample location test based on the center-outward signs and ranks

Daniel Hlubinka* Šárka Hudecová†

We consider a location test for a single multivariate sample based on the center-outward ranks and signs. Two possible approaches are proposed: The first one is based on a random division of the data into two samples, while the second one uses a symmetrized sample. The asymptotic distributions of the proposed tests are provided. For univariate data, two variants of the symmetrized test statistic are shown to be equivalent to the standard sign and Wilcoxon test respectively. The small sample behavior of the proposed techniques is illustrated by a simulation study that also provides a power comparison for various transportation grids.

Acknowledgments

This work was supported by the Czech Grant Agency GAČR, under grant 23-05737 S.

References

[1] Hlubinka, D. and Hudecová, Š. (2024). One sample location test based on the center-outward signs and ranks. *To appear*.

*Univerzita Karlova, Matematicko-fyzikální fakulta, daniel.hlubinka@matfyz.cuni.cz
†Univerzita Karlova, Matematicko-fyzikální fakulta

Bootstrapping not independent and not identically distributed data

Martin Hrba* Matúš Maciak† Barbora Peštová‡
Michal Pešta§

Classical normal asymptotics could bring serious pitfalls in statistical inference, because some parameters appearing in the limit distributions are unknown and, moreover, complicated to estimated (from a theoretical as well as computational point of view). Due to this, plenty of stochastic approaches for constructing confidence intervals and testing hypotheses cannot be directly applied. Bootstrap seems to be a plausible alternative. A methodological framework for bootstrapping not independent and not identically distributed data is presented together with theoretical justification of the proposed procedures. Among others, bootstrap laws of large numbers and central limit theorems are provided. The developed methods are utilized in insurance and psychometry.

*Charles University, Prague, Hrba@karlin.mff.cuni.cz
†Charles University, Prague, Matus.Maciak@mff.cuni.cz
‡Czech Academy of Sciences, Prague, Pestova@cs.cas.cz
§Charles University, Prague, Michal.Pesta@mff.cuni.cz

Specification tests for count time series

Šárka Hudecová* Marie Hušková† Simos G. Meintanis‡

Time series of counts enjoy numerous applications in various diverse fields as epidemiology, finance, and engineering. Corresponding inferential procedures have been intensively studied in the literature in recent years and various modelling approaches have been proposed. A popular class of models consists of observation-driven models, where the model dynamic is specified by the conditional distribution of the current observation given the past, referred to as INGARCH models. In this contribution, we deal with time series of counts, possibly accompanied by multivariate external covariates, and propose a distributional goodness–of–fit test for a specified class of INGARCH type models. The test is based on a combination of the conditional mean characterization of Bierens and the conditional probability generating function.

*Department of Probability and Mathematical Statistics, Charles University, Prague, hudecova@karlin.mff.cuni.cz

†Department of Probability and Mathematical Statistics, Charles University, Prague, huskova@karlin.mff.cuni.cz

‡University of Athens, Greece, simosmei@econ.uoa.gr

Detection of changes in panel data

Marie Hušková* Charl Pretorius†

Panel regression models with cross-sectional dimension N are considered. The aim is to test, based on T observations, whether the intercept in the model remains unchanged throughout the observation period. The test procedure involves the use of a CUSUM-type statistic derived via a quasi-likelihood argument. Limit behavior under the null distribution of the test under strong mixing and stationarity conditions on the errors and regressors are presented. Both independent panels as well as the case of mild cross-sectional dependence are considered. It is also proposed a self-normalized version of the test which is convenient from a practical perspective in that the estimation of long-run variances is avoided entirely. The theoretical results are supported by a simulation study that indicates that the test works well in the case of small to moderate sample sizes. An illustrative application of the procedure to US mutual fund data demonstrates the relevance of the proposed procedure in financial settings.

Acknowledgments

The research of Marie Huškova was supported by the Czech Science Foundation project GAČR No. 21-13323S.

*Charles University, Prague, Czech Republic, huskova@karlin.mff.cuni.cz
†North-West University, South Africa, charl.pretorius@nwu.ac.za

Estimating parameters of truncated and skew normal distribution with observations below limit of quantitation

Markéta Janošová* Stanislav Katina† Jozef Hanes‡

Clinical research relies heavily on laboratory measurements, however every laboratory equipment has limits to what it can accurately measure. Generally, for every laboratory apparatus three types of limits should be defined – limit of blank, limit of detection and limit of quantitation (LoQ). If a biological sample has such an amount of analyte that it falls below LoQ, this presents us with a problem of estimating parameters of the distribution and making an inference about the whole population, not just those within the measurable range.

Several methods can be used to address the issue, but their performance needs to be considered before selecting, which method to use. While it may be tempting to use methods that are easy to implement, selecting an incorrect method may lead to biased estimates and change the research outcomes.

In this contribution we look at four different methods – ignoring censored observations, replacing censored observations, using a truncated version of target distribution, and using target distribution with censored observations. To compare these methods we designed a simulation study, where generated samples were trimmed from the left at selected quantiles, mimicking situations with certain percentages of the population falling below the limit. Parameters' estimates were then compared to the original values. Simulation study was run separately on skew-normal distribution and truncated normal distribution.

Based on the results of the simulation study for truncated normal distribution, using truncated normal distribution with censored observations provides the best estimates of the parameters. For left-skewed skew-normal distribution, replacement by 0.5-times the limit provides the best estimates of the parameters. For right-skewed skew-normal distribution, using skew-normal distribution with censored observations provides the best estimates of the parameters.

Acknowledgments

This work was supported by specific research of Masaryk University as support for student projects (MUNI/A/1132/2022).

*Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic, janosova@math.muni.cz

†Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic, katina@math.muni.cz

‡Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia, jozef.hanes@savba.sk

Robust functional regression with discretely sampled predictors

Ioannis Kalogridis* Stanislav Nagy†

The functional linear model is an important extension of the classical regression model allowing for scalar responses to be modeled as functions of stochastic processes. Yet, despite the usefulness and popularity of the functional linear model in recent years, most treatments, theoretical and practical alike, suffer either from (i) lack of resistance towards the many types of anomalies one may encounter with functional data or (ii) biases resulting from the use of discretely sampled functional data instead of completely observed data. To address these deficiencies, this paper introduces and studies the first class of robust functional regression estimators for partially observed functional data. The proposed broad class of estimators is based on thin-plate splines with a novel computationally efficient quadratic penalty, is easily implementable and enjoys good theoretical properties under weak assumptions. We show that, in the incomplete data setting, both the sample size and discretization error of the processes determine the asymptotic rate of convergence of functional regression estimators and the latter cannot be ignored. These theoretical properties remain valid even with multi-dimensional random fields acting as predictors and random smoothing parameters. The effectiveness of the proposed class of estimators in practice is demonstrated by means of a simulation study and a real-data example.

Acknowledgments

This work was supported by the Research Foundation Flanders (project 1221122N) and the Czech Science Foundation (project n. 23- 05737S).

*KU Leuven, Belgium, ioannis.kalogridis@kuleuven.be
†Charles University, Czech Republic, nagy@karlin.mff.cuni.cz

Assessment of measurement properties of a questionnaire in R

Barbora Kessel* Arciolanda Macama† Francisca dos Reis‡
Jeanette Melin§ Maria Mache Malomalo¶ Gunilla Priebe‖

Drawing on the large amount of literature on item response theory and on Rasch measurement models, a unified workflow for assessment of measurement properties of a questionnaire has been suggested recently [1]. The recommended steps were implemented in a package in R [2], building on several existing packages for fitting Rasch models. In our contribution, we will exemplify the recommended assessment steps using data on 11 affirmations about gender norms collected among young Angolan women within the SADIMA project [3]. A special focus will be on the suggested approaches for examination of differential item functioning. In particular, the examination of individual items after detecting global differences by model-based recursive partitioning [4] will be discussed.

References

[1] Johansson, M. et al (2023). Valid and reliable? Basic and expanded recommendations for psychometric reporting and quality assessment. *Preprint.* https://doi.org/10.31219/osf.io/3htzc.

[2] Johansson, M. (2024, february 20). RISEkbmRasch vignette. https://pgmj.github.io/raschrvignette/RaschRvign.html

[3] Priebe, G. et al (2024). Social Capital among Young Angolan Women: A Cross Sectional Study in Three Provinces. *Under review.*

[4] Komboz, B., Strobl, C. and Zeilis, A. (2018). Tree-based global model tests for polytomous Rasch models. *Educational and Psychological Measurement* **78**(1) 128–166.

*University of Gothenburg, Sweden, barbora.kessel@gu.se
†The Catholic University of Angola, Angola
‡The Catholic University of Angola, Angola
§Södertörns University, Sweden and Swedish Defence University, Sweden
¶Mwana Pwo, Angola
‖University of Gothenburg, Sweden

Testing identification in mediation & dynamic treatment models

Martin Huber* Kevin Kloiber† Lukáš Lafférs‡

We propose a test for the identification of causal effects in mediation and dynamic treatment models that is based on two sets of observed variables, namely covariates to be controlled for and suspected instruments, building on the test by [1] for single treatment models. We consider models with a sequential assignment of a treatment and a mediator to assess the direct treatment effect (net of the mediator), the indirect treatment effect (via the mediator), or the joint effect of both treatment and mediator. We establish testable conditions for identifying such effects in observational data. These conditions jointly imply (1) the exogeneity of the treatment and the mediator conditional on covariates and (2) the validity of distinct instruments for the treatment and the mediator, meaning that the instruments do not directly affect the outcome (other than through the treatment or mediator) and are unconfounded given the covariates. Our framework extends to post-treatment sample selection or attrition problems when replacing the mediator by a selection indicator for observing the outcome, enabling joint testing of the selectivity of treatment and attrition. We propose a machine learning-based test to control for covariates in a data-driven manner and analyze its finite sample performance in a simulation study. Additionally, we apply our method to Slovak labor market data and find that our testable implications are not rejected for a sequence of training programs typically considered in dynamic treatment evaluations.

JEL Classification: C12, C21, C26

Keywords: mediation, dynamic treatment effects, causality, conditional independence, sequential exogeneity, instrument, covariates, hypothesis test.

Acknowledgments

L. Lafférs acknowledges support provided by the Slovak Research and Development Agency under contracts VEGA 1/0398/23 and APVV-21-0360.

References

[1] Huber, M. and Kueck, J. (2023). esting the identification of causal effects in observational data *arXiv preprint arXiv:2203.15890.*

*University of Fribourg, Department of Economics, martin.huber@unifr.ch
†University of Munich, Department of Economics, kevin.kloiber@econ.lmu.de
‡Matej Bel University, Department of Mathematics, lukas.laffers@umb.sk

Functional approaches for claims reserving using standard chain ladder data

Matúš Maciak* Ivan Mizera† Michal Pešta‡

Typical and commonly used techniques for claims reserving are based on different parametric approaches using some underlying model formulation and aggregated data–so-called run-off triangles. We propose some non-parametric, model-free alternatives that handle the underlying loss development triangles as functional profiles. The estimation is performed in a fully data-driven manner and the overall claim reserve prediction is given in terms of the whole distribution obtained via a permutation bootstrap.

Theoretical justifications are provided and some practical implementation issues are addressed for all three proposed non-parametric methods. In addition, a finite sample evaluation in terms of a full-scale comparison with standard (parametric) reserving techniques is carried on several hundreds of real run-off triangles against known real loss outcomes.

References

[1] Maciak, M. and Mizera, I. and Pešta, M. (2022). Functional Profile Techniques for Claims Reserving. *ASTIN Bulletin* **52**(2) 449–482.

*Charles University, Prague, Czech Republic, maciak@karlin.mff.cuni.cz
†University of Alberta, Edmonton, Canada, imizera@ualberta.ca
‡Charles University, Prague, Czech Republic, pesta@karlin.mff.cuni.cz

Design of experiments for function-on-function linear models

Caterina May* Theodoros Ladas† Kalliopi Mylona‡
Davide Pigoli§

In this work we study optimal experimental designs for precise estimation of the functional coefficient of a linear regression model where both the response and the factors are continuous functions of the time. After obtaining the variance-covariance matrix of an estimator which minimizes the integrated sum of square of errors, we extend the A-optimality and D-optimality criterion to this context. The optimal designs for the dynamic factors are then computed through a suitable algorithm.

Acknowledgments

This work was supported by the major research grant entitled *Multi-objective optimal design of experiments* funded by the EPSRC (Engineering and Physical Sciences Research Council, UK).

*King's College London, UK and Università degli Studi del Piemonte Orientale, Italy, caterina.may@kcl.ac.uk

†King's College London, theodoros.ladas@kcl.ac.uk

‡King's College London, kalliopi.mylona@kcl.ac.uk

§King's College London, davide.pigoli@kcl.ac.uk

Density estimation via maximum likelihood: shape constraints and beyond

Ivan Mizera*

Recent progress on estimation algorithms in density estimation using maximum likelihood, and algorithmic approaches to those, is surveyed and complemented. Apart from possible penalization approaches, the focus is on various shape-constrained situations: either in the unimodal situations, but also under the assumption of the monotonicity of the estimated density (various interpretations of the celebrated Grenander estimator).

*University of Alberta and Charles University, imizera@ualberta.ca

Block covariance matrix estimation with structured off-diagonal blocks

Malwina Mrowińska* Monika Mokrzycka†

This talk deals with the estimation of a block covariance matrix with off-diagonal blocks corresponds to the part of autoregression of order one structure, AR(1). Commonly used maximum likelihood estimation is challenging and time-consuming, thus we propose also another approach based on the least squares method. Some estimates are not always well-conditioned and may not even be definite. Thus, the improvement based on a shrinkage method and an additional algebraic approach is applied. The considered structure can be also expressed as a sum of two matrices: block diagonal matrix and AR(1) matrix. New approach based on estimation of whole AR(1) structure is presented and several estimation method are proposed. All considered estimates are compared with respect to some statistical properties and time needed to determine them.

*Institute of Mathematics, Poznań University of Technology, Poland, malwina.mrowinska@put.poznan.pl

†Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland, mmok@igr.poznan.pl

A criterion and incremental design construction for simultaneous kriging predictions

Helmut Waldl* Werner G. Müller† Paula Camelia Trandafir‡

In this presentation, we further investigate the problem of selecting a set of design points for universal kriging, which is a widely used technique for spatial data analysis. Our goal is to select the design points in order to make simultaneous predictions of the random variable of interest at a finite number of unsampled locations with maximum precision. Specifically, we consider as response a correlated random field given by a linear model with an unknown parameter vector and a spatial error correlation structure. We propose a new design criterion that aims at simultaneously minimizing the variation of the prediction errors at various points. We also present various efficient techniques for incrementally building designs for that criterion scaling well for high dimensions. Thus the method is particularly suitable for big data applications in areas of spatial data analysis such as mining, hydrogeology, natural resource monitoring, and environmental sciences or equivalently for any computer simulation experiments. We have demonstrated the effectiveness of the proposed designs through two illustrative examples: one by simulation and another based on real data from Upper Austria.

Acknowledgments

This work was partly supported by project INDEX(INcremental Design of EXperiments) I3903-N32 of the Austrian Science Fund(FWF). The third author received support for this research from a fellowship provided by the Spanish Ministry of Universities (PRX22/00578).

References

[1] Waldl, H., Müller, W.G., and Trandafir, P. C. (2024). A criterion and incremental design construction for simultaneous kriging predictions. *Spatial Statistics* **59** 100798. https://www.sciencedirect.com/science/article/pii/S2211675323000738

*Department of Applied Statistics, Johannes-Kepler-University Linz, helmut.waldl@jku.at

†Department of Applied Statistics, Johannes-Kepler-University Linz, Werner.Mueller@jku.at

‡Department of Statistics, Informatics and Mathematics, Public University of Navarre, camelia@unavarra.es

How to robustify an experimental design: introducing a new methodology

Alba Muñoz del Río* Víctor Casero-Alonso†
Mariano Amo-Salas‡

The use of nonlinear models implies assuming uncertainty in nominal values of the parameters. In this situation, practitioners seek a robust experimental design. Several approaches have been developed in the literature to obtain robust experimental designs such as adaptive or Bayesian designs, among others. However, the application of these methods is not straightforward or requires an advanced knowledge of statistics. This paper proposes a new methodology for robustifying an optimal experimental design with respect to uncertainty in the nominal values of the parameters. Based on the maximin idea, the method adds support points to the optimal design, to obtain designs that are robust. This methodology comes particularly useful in the field of predictive microbiology where it has been applied to the Baranyi model, one of the most widely used mathematical models to describe the behavior of microorganisms in food products.

Acknowledgments

This work was supported by Ministerio de Ciencia e Innovación [grant number PID2020-113443RB-C21], by Junta de Comunidades de Castilla-La Mancha [grant number SBPLY/21/180501/000126] and by Universidad de Castilla-La Mancha, Plan Propio co-funded by FEDER [grant number 2022-GRIN-34288].

References

[1] Grijspeerdt, K. & Vanrolleghem, P. (1999). Estimating the parameters of the Baranyi model for bacterial growth. *Food Microbiology*. **16**, 593–605.

[2] Baranyi, J. (1997). Simple is good as long as it is enough. *Food Microbiology*. **14**, 189–192.

[3] López-Fidalgo, J. (2023). Optimal Experimental Design: A Concise Introduction for Researchers. Springer Nature.

[4] Calle-Arroyo, C., Amo-Salas, M., López-Fidalgo, J., Rodríguez-Aragón, L. & Wong, W. (2023). A methodology to D-augment experimental designs. *Chemometrics And Intelligent Laboratory Systems*. **237** pp. 104822.

*Universidad de Castilla-La Mancha, Spain, alba.munoz@uclm.es
†Universidad de Castilla-La Mancha, Spain, VictorManuel.Casero@uclm.es
‡Universidad de Castilla-La Mancha, Spain, mariano.amo@uclm.es

On boxplots for functional data

Stanislav Nagy* Tomáš Mrkvička† Antonio Elías‡

The boxplot for functional data is a useful visualisation tool for datasets where each observation is a function. Such boxplots are based on the notion of a functional depth, ranking the observations from the typical to the more extreme ones. While, in principle, any functional depth can be used to construct a boxplot, the most common depths used in practice are integrated depths, e.g., the popular modified band depth.

In this contribution we highlight that integrated depths do not result in well-defined boxplots. Instead, we argue that infimal depths are the only functional depths that provide a valid construction of a functional boxplot.

Acknowledgments

This work was supported by the Czech Science Foundation (project n. 24-10822S).

*Charles University, Czech Republic, nagy@karlin.mff.cuni.cz
†University of South Bohemia, Czech Republic, mrkvicka.toma@gmail.com
‡University of Malaga, Spain, aelias@uma.es

Prediction model for hydrogel formation

Ricardo Negrete* Irene García-Camacha Sergio Pozuelo
Carlos Martin Ester Vazquez

Nowadays, there is a growing interest in hydrogels due to its wide applicability in many fields. Controlling drug release within the body and improving water retention in the soil, benefiting plant growth, are some of their numerous applications. Prediction models play a crucial role in experimentation for optimizing material usage and, consequently, for hydrogel formation. In this study, we propose a model based on ordinal regression with a logistic function to optimally estimate the quantities of monomers integrating the mixture, as well as the time for hydrogel formation. In order to effectively perform the optimization, an adaptation of the Particle Swarm Optimization (PSO) algorithm was developed to deal with the special nature of mixture experiments. Results illustrate an optimal hydrogel synthesis. We also provide a free interactive app using Shiny, which is open to the scientific community. It allows experimenters to precisely determine the mixture subject to their particular experimental conditions.

Acknowledgments

This work was supported by the Advances Materials program (MATERIALES-PCOMPLEM) funded by the Castilla-La Mancha Regional Government, the Ministry of Science and Innovation, and the European Union.

*Castilla-La Mancha University, Ciudad Real, Ricardo.Negrete@uclm.es

Two-sample testing in reinforcement learning

Martin Waltz[1*] Ostap Okhrin[1,2†]

Value-based reinforcement-learning algorithms have shown strong performances in games, robotics, and other real-world applications. The most popular sample-based method is Q-Learning. It subsequently performs updates by adjusting the current Q-estimate towards the observed reward and the maximum of the Q-estimates of the next state. The procedure introduces maximization bias with approaches like Double Q-Learning. We frame the bias problem statistically and consider it an instance of estimating the maximum expected value (MEV) of a set of random variables. We propose the T-Estimator (TE) based on two-sample testing for the mean, that flexibly interpolates between over- and underestimation by adjusting the significance level of the underlying hypothesis tests. A generalization, termed K-Estimator (KE), obeys the same bias and variance bounds as the TE while relying on a nearly arbitrary kernel function. We introduce modifications of Q-Learning and the Bootstrapped Deep Q-Network (BDQN) using the TE and the KE. Furthermore, we propose an adaptive variant of the TE-based BDQN that dynamically adjusts the significance level to minimize the absolute estimation bias. All proposed estimators and algorithms are thoroughly tested and validated on diverse tasks and environments, illustrating the bias control and performance potential of the TE and KE.

References

[1] Waltz, M. and Okhrin, O. (2023). Addressing Maximization Bias in Reinforcement Learning with Two-Sample Testing, arXiv: `https://arxiv.org/pdf/2201.08078`

[1]Technische Universität Dresden, Dresden, Germany

[2]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany

[*]`martin.waltz@tu-dresden.de`

[†]`ostap.okhrin@tu-dresden.de`

Hurdle semi-continuous time series

Šárka Hudecová* Michal Pešta†

Time series containing non-negligible portion of possibly dependent zeros, whereas the remaining observations are positive, are considered. They are regarded as GARCH processes consisting of non-negative values. Our first aim lies in estimation of the omnibus model parameters taking into account the semi-continuous distribution. The hurdle distribution together with dependent zeros cause that the classical GARCH estimation techniques fail. Two different quasi-likelihood approaches are employed. Both estimators are proved to be strongly consistent and asymptotically normal. The second goal consists in the proposed predictions with bootstrap add-ons. The considered class of models can be reformulated as multiplicative error models. The empirical properties are illustrated in a simulation study, which demonstrates computational efficiency of the employed methods. The developed techniques are presented through an actuarial problem concerning insurance claims.

*Charles University, Prague, Czech Republic, hudecova@karlin.mff.cuni.cz
†Charles University, Prague, Czech Republic, pesta@karlin.mff.cuni.cz

Testing structural breaks in large dynamic models

Zuzana Prášková*

A linear dynamic panel data model with cross-sectional dependence is considered. A procedure to detect changes in coefficients of lagged variables is proposed and asymptotic distribution of the test statistic is studied in case that both the number of panels and number of observations are infinitely large. Bootstrap variants of the test statistic are discussed that take into account both the temporal and the cross-sectional dependences.

Acknowledgments

This work was partially supported by the Czech Science Foundation, under grant GAČR 21-13323S.

*Charles University, Czech Republic, praskova@karlin.mff.cuni.cz

Symmetry in functional spaces via characteristic functionals

Hedvika Ranošová* Daniel Hlubinka†

Methods in functional data analysis are often based on generalizations taken from multivariate statistics, however, only some multivariate techniques are available in infinite-dimensional vector spaces. In this poster, we aim to explore the problem of functional symmetry. Multivariate (sign) symmetry can be tested via the characteristic function and its empirical counterpart. In functional settings, we can similarly rely on the characteristic functional defined by the functional space's dual. Thus, we can construct a Cramér-von Mises type statistical test based on characteristic functionals. The characteristic functional approach is compared to a method that utilizes functional principal components and tests of multivariate symmetry.

Acknowledgments

This work was supported by German Research Foundation (DFG) grants KI 1443/6-1 and WE 5988/5-1.

References

[1] Hlávka, Z., Hlubinka, D., and Koňasová, K. (2022). Functional ANOVA based on empirical characteristic functionals. *Journal of Multivariate Analysis* **189** 193–212.

[2] Henze N. and Jiménez-Gamero, M. G. (2021). A test for Gaussianity in Hilbert spaces via the empirical characteristic functional. *Scandinavian Journal of Statistics* **48.2** 406–428.

*Otto-von-Guericke Universität Magdeburg, hedvika.ranosova@ovgu.de
†Charles University Prague, hlubinka@karlin.mff.cuni.cz

Non-classically generalized complex numbers with applications to probability distributions and their characteristic functions

Wolf-Dieter Richter*

Classical generalizations of complex numbers often begin by determining a non-commutative but distributive product for every two or more basic elements of an algebraic structure. A geometric interpretation of this product within the whole space or algebraic structure is generally not a motivating starting point, but has emerged as an interesting aspect in some individual cases. For example, the scalar and the vectorial parts of a certain quaternion product in a four-dimensional space reflect a Euclidean scalar product and a vector product in the three-dimensional Euclidean space, respectively.

In contrast, the non-classical generalizations of complex numbers that are presented in this talk begin in the two-dimensional case with the consideration of generalized circles that are level sets of a norm, antinorm or semi-antinorm $||.||$, and movements along such level lines as well as transitions between these lines. The latter are described based on the situation with complex numbers by varying an angle variable or a generalized radius variable and thus motivate the definition of $||.||-$related generalized polar coordinates and of a vector-valued product of generalized complex numbers. In this way, the relationship between the imaginary unit of usual complex numbers and the Euclidean unit circle, which is conveyed by Euler's formula, as well as the latter formula itself are far-reaching generalized. The resulting algebraic structure is commutative but not distributive, in general.

This approach extends to higher dimensions and applies to probability laws and their characteristic functions. We discuss invariant probability densities and associated Lie groups, heavy and light tailed density generators of Kotz-type and Pearson-type VII, respectively, and static as well as dynamically changing contours of probability density level sets. Particular emphasis is placed on the vector representation of complex numbers, where we transform Gauss' interpretation of complex numbers as points on the plane into the status of an axiom, and the resulting vector representation of characteristic functions. We present some elementary properties of the latter representation such as a polar representation and a vector power expansion as well as further exemplary effects.

References

[1] Richter, W.-D. (2023). On the vector representation of characteristic functions, *Stats* **2023**, 6, 1072–1081.

[2] Richter, W.-D. (2024). Short remark on $(p_1, p_2, p_3)-$complex numbers, *Symmetry* **2024**, 16, 9.

*University of Rostock, Rostock Germany, wolf-dieter.richter@uni-rostock.de

Identification of the personal sales succeed factors in a wealthy automotive market environment with exploratory factor analysis

Elena Říhová* David Říha†

Understanding the intricate dynamics of personal sales success factors in wealthy market environments is paramount for businesses aiming to succeed in competitive industries such as the automotive industry. This study based on exploratory factor analysis to identify key determinants of success in personal sales within a wealthy automotive market. By analyzing data collected from sales interactions and customer feedback, underlying factors influencing sales performance and customer satisfaction are uncovered.

Through a rigorous examination of sales strategies, techniques, and communication methods, the nuanced elements driving success in personal sales are detected. These findings shed light on the multifaceted nature of effective sales practices, offering valuable insights for practitioners and researchers alike.

References

[1] Costello, A. B. and Osborne, J. W. G. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* **10-7**, 1–9.

[2] Fabrigar, L. R. and Wegener, D. T. and MacCallum, R. C. and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* **4-3**, 272–299.

[3] Gorsuch, R. L. (2015). Factor analysis. New York: Psychology Press and Routledge Classic Editions.

[4] Huck, S. W. (2012). Reading statistics and research. Boston, MA: Pearson.

*Škoda Auto University, Mladá Boleslav, elena.rihova@vse.cz
†Prague University of Economics and Business, Prague, david.riha@vse.cz

Mixed-integer linear programming for computing optimal designs

Samuel Rosa* Radoslav Harman†

The problem of constructing an optimal design for a given experiment can be viewed as a discrete optimization problem. As such, it can be solved by general methods of integer programming, which may, however, be too slow for more difficult design problems. One approach to speeding up the solution is to use specialized solvers, provided that a more structured formulation of the problem is known. For instance, for many design criteria a mixed-integer second-order cone programming formulation of the optimal exact design problem was proposed in [1]. We show (see [2]) that for some criteria this mathematical programming specialization can be taken even further. In particular, we provide a mixed-integer linear programming (MILP) formulation for optimal replication-free designs with respect to a wide class of criteria, including A-, I-, G- and MV-optimality. We also show that the MILP formulation can be extended to exact designs with replications and demonstrate some unique advantages of the MILP approach.

Acknowledgments

This work was supported by the Slovak Scientific Grant Agency VEGA, under grant VEGA 1/0362/22.

References

[1] Sagnol, G. and Harman, R. (2015), Computing exact D-optimal designs by mixed integer second-order cone programming, *The Annals of Statistics* **43**, 2198–2224.

[2] Harman, R. and Rosa, S. (2023), Mixed-integer linear programming for computing optimal experimental designs. arXiv: `https://arxiv.org/pdf/2305.17562`

*Comenius University Bratislava, Slovakia, `samuel.rosa@fmph.uniba.sk`
†Comenius University Bratislava, Slovakia, `radoslav.harman@fmph.uniba.sk`

A tensor decomposition approach to EEG eye blink removal

Zuzana Rošťáková* Roman Rosipal†

The human electroencephalogram (EEG), measured under the eyes open condition, is usually corrupted by ocular artefacts due to eye blinks and movement. The efficient removal of ocular artefacts is an essential step in the preprocessing of EEGs. Many ocular correction approaches use linear regression, such as the Gratton-Coles & Donchin method (GCD) [1], or blind source separation, such as independent component analysis (ICA) [2]. It is important to note, however, that both approaches have their limitations, for example the requirement of separate electrooculogram (EOG) signals in GCD or distortions of the EEG signals in ICA due to overcorrection.

This study proposes a tensor-based alternative method of ocular correction called Signal Spectrum Tensor Decomposition and Eye Blink Removal (SPECTER). Preliminary results on two real EEG datasets demonstrated that SPECTER outperformed existing tensor-based approaches for removing eye blinks and was comparable to ICA and GCD. Furthermore, SPECTER produced accurate results when traditional methods of eye blink correction lead to distortions in the EEG signals. Finally we can conclude that SPECTER is also capable of removing a variety of other artifacts and specific EEG rhythms.

Acknowledgments

This work was supported by the Slovak Scientific Grant Agency VEGA under grants VEGA 2/0023/22 and VEGA 2/0057/22, by the APVV-21-0105 grant and by the CHIST-ERA 20-BCI-004 grant.

References

[1] Gratton, G. and Coles, M. G. H. and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology* **55** 468–484.

[2] Wallstrom, G. L. and Kass, R. E. and Miller, A. and Cohn, J. F. and Fox, N. A. (2004). Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. *International Journal of Psychophysiology* **53** 105–119.

*Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia, zuzana.rostakova@savba.sk

†Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia, roman.rosipal@savba.sk

A note on the asymptotic confidence intervals for differences of location parameters

František Rublík*

In computing of non-parametric confidence intervals for differences of location parameters of k populations, based on large sample approximation, the quantile $t(k, 1-\alpha)$ of range of k-dimensional $N_k(\mathbf{0}, \mathbf{I}_k)$ distribution is used. It is known that this approach yields a conservative results. It is demonstrated by the numerical computation of the precise asymptotic quantile $t(k, p_1, \ldots, p_k, 1-\alpha)$ based on the limits $p_1, \ldots, p_k$ of relative frequencies of the sample sizes, that its use yields asymptotic coverage practically equivalent with this classical approach, because with using $t(k, 1-\alpha)$ the departure from chosen $1-\alpha$ coverage differs only slightly.

Acknowledgments

This work was supported by the Grant APVV-21-0216.

References

[1] Critchlow, D.E and Fligner, M, A. (1991). On distribution free multiple comparisons in the one-way analysis of variance. *Communications in Statistics, Theory and Methods* **20** 127–139.

[2] Genz, A. (1992). Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics* **1** 141— 150.

[3] Genz, A. and Bretz, F. (2009) Computation of Multivariate Normal and t Probabilities. *Springer-Verlag,* Berlin-Heidelberg.

[4] Healy, M.J.R. (1968) Algorithm AS 6: Triangular decomposition of a symmetric matrix. *Journal of the Royal Statistical Society, Series C* **17** 195 –197.

[5] Hayter, A. J. (1984). A proof of the conjecture that the Tukey-Kramer multiple comparison procedure is conservative. *Annals of Statistics* **12** 61–75.

[6] Hollander, M., Douglas, A. W. and Chicken, E.(2014) Nonparametric Statistical Methods. *John Wiley & Sons, Inc.*

*Institute of Measurement Science, Slovakia, umerrubl@savba.sk

Spatio-temporal high-dimensional matrix autoregressive models via tensor decomposition

Seyed Yaser Samadi* Rukayya Ibrahim†
Tharindu P De Alwis‡

With the rapid increase in massive, interactive datasets, including time-dependent big data and spatiotemporal data, various domains such as econometrics, geospatial technologies, and medicine face the challenge of efficiently handling their high dimensionality. To address this complexity, tensor decomposition techniques offer valuable advantages, such as latent structure identification, information extraction, data imputation, and complexity control, making them popular for analyzing, predicting, and forecasting these datasets. In this presentation, we introduce a novel approach for modeling and analyzing matrix-valued spatiotemporal data by formulating it as a tensor regression model based on matrix autoregression. Our method capitalizes on the matrix structure of both the response and predictors while achieving dimension reduction through a low-rank tensor structure. Comparative analyses demonstrate the superior efficiency of our model compared to existing approaches for high-dimensional data. Furthermore, we propose two estimation methods to estimate the transition tensor in both low and high-dimensional scenarios. We also derive the asymptotic and non-asymptotic properties of the proposed estimators, providing a solid theoretical foundation. Simulation studies and real data analysis are conducted to illustrate the advantages of our model over current methodologies.

*Southern Illinois University Carbondale, United States, ysamadi@siu.edu
†Penn State Harrisburg, United States
‡Worcester Polytechnic Institute, Worcester, United States

A unified framework for pattern recovery in penalized estimation and its geometry

Ulrike Schneider* Patrick Tardivel† Tomasz Skalski‡
Piotr Graczyk§

We consider the framework of penalized estimation where the penalty term is given by a polyhedral norm, or more generally, a polyhedral gauge, which encompasses methods such as LASSO and generalized LASSO, SLOPE, OSCAR, PACS and others. Each of these estimators can uncover a different structure or "pattern" of the unknown parameter vector. We define a novel and general notion of patterns based on subdifferentials and formalize an approach to measure pattern complexity. For pattern recovery, we provide a minimal condition for a particular pattern to be detected by the procedure with positive probability, the so-called accessibility condition. We also introduce the stronger noiseless recovery condition which can be shown to play exactly the same role as the well-known irrepresentability condition for the LASSO in that the probability of pattern recovery in our general framework is bounded by 1/2 if the condition is not satisfied, thereby unifying and extending the irrepresentability condition to a broad class of penalized estimators. Finally, we prove that the noiseless recovery condition can indeed be relaxed when turning to so-called thresholded penalized estimation: in this setting, the accessibility condition is already sufficient (and necessary) for sure pattern recovery provided that the signal of the pattern is large enough. We demonstrate how our findings can be interpreted through a geometrical lens throughout the talk and illustrate our results for LASSO and SLOPE.

References

[1] Ewald, K. and Schneider, U. (2020). Model Selection Properties and Uniqueness of the Lasso Estimator in Low and High Dimensions. *Electronic Journal of Statistics* **14**, 944–969.

[2] Schneider, U. and Tardivel, P. (2022). The Geometry of Uniqueness, Sparsity and Clustering in Penalized Estimation. *Journal of Machine Learning Research* **23**, 1–36.

[3] Graczyk, P., Schneider, U., Skalski, T. and Tardivel, P. (2023). A Unified Framework for Pattern Recovery in Penalized and Thresholded Estimation and its Geometry. arXiv: `https://arxiv.org/pdf/2307.10158`

*TU Wien, Vienna, Austria, `ulrike.schneider@tuwien.ac.at`
†Université de Bourgogne, Dijon, France, `patrick.tardivel@u-bourgogne.fr`
‡Politechnika Wrocławska, Wrozław, Poland, `tomasz.skalski@pwr.edu.pl`
§Université d'Angers, Angers, France, `piotr.graczyk@univ-angers.fr`

A randomized exchange algorithm for optimal experimental design problems with general elementary information matrices

Pál Somogyi* Radoslav Harman† Samuel Rosa‡

The available randomized exchange algorithm for computing optimal approximate designs of experiments ([1]) is simple and efficient, yet it is limited to the basic optimality criteria and rank-one elementary information matrices. This paper proposes a generalization of the REX algorithm, expanding its capabilities to compute optimal approximate designs with respect to all Kiefer's optimality criteria and elementary information matrices of any rank. In addition to the generalization of the REX algorithm, we also present some applications of optimal design problems with general-rank elementary information matrices, such as multivariate response models. Numerical results confirm the stability and rapid convergence of the proposed algorithm.

Acknowledgments

This work was supported by the Slovak Scientific Grant Agency VEGA, under grant VEGA 1/0362/22 and by the Collegium Talentum Programme of Hungary.

References

[1] Harman, R. and Filová, L. and Richtárik, P. (2020). A Randomized Exchange Algorithm for Computing Optimal Approximate Designs of Experiments. *Journal of the American Statistical Association* **115(529)** 348–361.

*Comenius University Bratislava, Slovakia, pal.somogyi@fmph.uniba.sk
†Comenius University Bratislava, Slovakia, radoslav.harman@fmph.uniba.sk
‡Comenius University Bratislava, Slovakia, samuel.rosa@fmph.uniba.sk

On determining the minimal sample size in balanced 3-way ANOVA models where no exact F-test exists

Bernhard Spangl* Norbert Kailblinger†

We consider balanced three-way ANOVA models to test the hypothesis that the fixed factor A has no effect. The other factors are fixed or random. For most of these models (including all balanced 1-way and 2-way ANOVA models) an exact F-test exists. Details on the determination of the minimal sample size and on an in-depth structural result can be found in [1].

For the two models

$$A \times \boldsymbol{B} \times \boldsymbol{C} \qquad \text{and} \qquad (A \succ \boldsymbol{B}) \times \boldsymbol{C} \tag{1}$$

(bold letters indicate random factors), however, an exact F-test does not exist. Approximate F-tests, which can be obtained by Satterthwaite's approximation, involves mean squares to be simulated. To approximate the power of the test, we simulate data such that the null hypothesis is false and the rate of rejections then approximates the power of the test.

In this talk we aim to determine the minimal sample size of the two models mentioned above, given a prespecified power, and we

1. determine the active and inactive variance components for both ANOVA models by using a surrogate fractional factorial model with variance components as factors.

2. determine the worst combination of active variance components for both models by using a surrogate response surface model based on a Box-Behnken design. The special structure of the Box-Behnken design ensures that the used models have similar total variance.

Additionally we compare practical methods that help reducing the number of simulations required to determine the minimal sample size.

References

[1] Spangl, B. and Kaiblinger, N. and Ruckdeschel, P. and Rasch, D. (2023). Minimal sample size in balanced ANOVA models of crossed, nested, and mixed classifications. *Communications in Statistics – Theory and Methods* **52(6)**, 1728–1743.

*BOKU University, Vienna, bernhard.spangl@boku.ac.at
†BOKU University, Vienna, norbert.kaiblinger@boku.ac.at

Comparison of algorithms for exact optimal designs in the gBLUP model

Alexandra Stadler* Werner G. Müller† Andreas Futschik‡

In contemporary breeding programs, genomic best linear unbiased prediction (gBLUP) models are employed to drive decisions on artificial selection. Experiments are performed to obtain responses on the units in the breeding program. Due to restrictions on the size of the experiment, an efficient experimental design must be found. The proposed poster states the design problem for the gBLUP model and compares classical exchange-type algorithms for exact optimum designs to the TrainSel R package and algorithm by [1]. Particular emphasis is placed on evaluating the computational runtime of algorithms along with their respective efficiencies over different sample sizes. The algorithms are compared for the D-criterion and the CDMin-criterion.

References

[1] Akdemir, D., Rio, S. and Isidro-Sánchez, J. (2021). TrainSel: An R Package for Selection of Training Populations. *Frontiers in Genetics*, **12**. DOI: 10.3389/fgene.2021.655287

[2] Cook, R. D., and Nachtsheim, C. J. (1980). A Comparison of Algorithms for Constructing Exact D-Optimal Designs. *Technometrics*, **22**(3), 315. DOI: 10.2307/1268315

[3] Fedorov, V. V. (1972) Theory of Optimal Experiments. *Academic Press, New York.*

[4] Isidro y Sánchez, J. and Akdemir, D. (2021). Training Set Optimization for Sparse Phenotyping in Genomic Selection: A Conceptual Overview. *Frontiers in Plant Science*, **12**. DOI: 10.3389/fpls.2021.715910

[5] Laloë, D. (1993). Precision and information in linear models of genetic evaluation. *Genetics Selection Evolution*, **25**(6), 557. DOI: 10.1186/1297-9686-25-6-557

*Institute of Applied Statistics, Johannes Kepler University Linz, Austria, alexandra.stadler@jku.at

†Institute of Applied Statistics, Johannes Kepler University Linz, Austria, werner.mueller@jku.at

‡Institute of Applied Statistics, Johannes Kepler University Linz, Austria, andreas.futschik@jku.at

Additive regression with general imperfect variables

Germain Van Bever*

In this talk, we study an additive model where the response variable is Hilbert-space-valued and predictors are multivariate Euclidean, and both are possibly imperfectly observed. Considering Hilbert-space-valued responses allows to cover Euclidean, compositional, functional and density-valued variables. By treating imperfect responses, we can cover functional variables taking values in a Riemannian manifold and the case where only a random sample from a density-valued response is available. Dealing with imperfect predictors allows us to cover various principal component and singular component scores obtained from Hilbert-space-valued variables. For the estimation of the additive model having such variables, we use the smooth backfitting method originated by Mammen et al. (1999). We provide full non-asymptotic and asymptotic properties of our regression estimator and present its wide applications via several simulation studies and real data applications. This is a joint work with Jeong Min JEON.

*Department of Mathematics, University of Namur, Namur, Belgique, Université libre de Bruxelles, Brussels, Belgium, germain.vanbever@unamur.be

Depth-based detection of outliers for functional data

Ondřej Vencálek* Daniel Hlubinka† Stanislav Nagy‡

Functional data analysis has developed greatly in the last decade. In this contribution, we deal with detection of functional outliers.

We introduce a new procedure for detection of outliers among continuous functions on a compact domain $[0,1]$ or $[0,1]^d$. Outlyingness of a function may be due to its location but also due to its unusual shape. Newly defined procedure enables detection of outlying functions both in location and/or in shape.

The procedure is based on functional data depth, either integrated or infimal. It is an alternative to the outlier detection based on the Jth order depth suggested in [1]. The idea is to use one dimensional functional depth of function x "decomposed" to polynomials.

The suggested procedure is simple and computationally feasible. Moreover, it gives the information about the *order* of outlyingness, i.e., that a function is outlier in the sense of location, trend, convexity and so on.

References

[1] Nagy, S., Gijbels, I. and Hlubinka, D. (2017). Depth-based recognition of shape outlying functions. *Journal of Computational and Graphical Statistics* **26**, 4, 883–893.

*Palacký University, Olomouc, ondrej.vencalek@upol.cz
†Charles University, Prague, daniel.hlubinka@matfyz.cuni.cz
‡Charles University, Prague, nagy@karlin.mff.cuni.cz

Change-point detection for functional time series without dimension reduction

Martin Wendler*

For functional data, it is often recommended to use dimension reduction or regularization techniques. But it is also possible to use the full functional information. We will discuss how the classical CUSUM test statistics and the Wilcoxon type statistic for the detection of change-points in time series can be generalized to observations that take values in a Hilbert-space. In order to obtain critical values, we propose to use Bootstrap methods.

Acknowledgments

This work was supported by the Germany Research Foundation (DFG) project WE 5988/3 and project WE 5988/5.

References

[1] L. Wegner, M. Wendler (2023). Robust Change-Point Detection for Functional Time Series Based on U-Statistics and Dependent Wild Bootstrap. *preprint* arXiv:2206.01458.

[2] O. Sharipov, J. Tewes, M. Wendler (2016). Sequential block bootstrap in a Hilbert space with application to change point analysis. *Canadian Journal of Statistics* **44** 300-322.

*Otto-von-Guericke-Universität Magdeburg, martin.wendler@ovgu.de

Calibration model as a straight-line errors-in-variables model

Gejza Wimmer* Viktor Witkovský†

In metrology is calibration defined as an operation that, under specified conditions, in a first step,establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication. Calibration is central to measurement science and relates to a wide range of industrial sectors and services. We are considering the measurement model of linear comparative calibration which is from the point of view of mathematical statistics a nonlinear regression model of direct measurements. The measurement vectors are normally distributed random vectors, μ and ν are vectors of their mean values and are related by the equation $\nu = a\mathbf{1} + b\mu$. The covariance matrix of the model is a known positive definite matrix. We develop the mean value function of the model using Taylor expansion in appropriately chosen prior values μ_0,a_0,b_0 of the model parameters. In the case under consideration a linear-quadratic regression model of direct measurements is obtained with new parameters $\delta\mu, \delta a, \delta b$. This model is referred to us as a weakly nonlinear model [1]. We aim to determine the conditions under which the nonlinear regression model representing the straight-line calibration model can be treated as a conventional linear regression model. If the second-order terms in the Taylor expansion of the function of mean values are statistically nonsignificant, we get a linear regression model and linear methods can be applied. On the other hand, we recommend utilizing the Weighted Total Least Squares method for parameter estimation and covariance matrix determination. The Monte Carlo method can be used as an alternative approach.

Acknowledgments

The project was supported by projects APVV-21-0216, SK-CZ-RD-21-0109, VEGA 2/0120/24, VEGA 2/0023/22.

References

[1] Kubáček L., Tesaříková, E. (2008). *Weakly Nonlinear Regression Models.* Olomouc.

*Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovak Republic, wimmer@mat.savba.sk

†Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovak Republic, witkovsky@savba.sk

Testing the difference of medians using empirical characteristic functions

Viktor Witkovský*

This presentation was prompted by the prevailing misconception that the Wilcoxon–Mann–Whitney (WMW) procedure tests for equality of medians, which is known to be largely imprecise, as discussed, for example, in [1]. This interpretation of the null hypothesis is true only in exceptional cases where the population distributions of the two groups are merely offsets of each other (i.e., they differ solely in location, not in shape or scale). Thus, for general tests about medians, other alternative test procedures should be considered, such as Mood's median test and the permutation test of differences in medians, which offer valid alternatives, but each has its own limitations.

This presentation proposes a simple alternative approach based on differences in bootstrapped medians. We introduce the exact bootstrap distribution of sample medians based on their characteristic functions as well as the exact bootstrap distribution of their difference as a tool for testing the null hypothesis concerning the equality of medians based on two independent samples.

The suggested test procedure can be easily implemented within the characteristic functions toolbox [2], consisting of a set of algorithms for evaluating selected characteristic functions and numerical inversion of the combined and/or compound characteristic functions, used to evaluate the cumulative distribution function (CDF), the probability density function (PDF), and/or the quantile function (QF).

Acknowledgments

This work was supported by projects APVV-21-0216, VEGA 2/0023/22, and VEGA 2/0120/24.

References

[1] Divine, G. W., Norton, H. J., Barón, A. E., Juarez-Colunga, E. (2018). The Wilcoxon–Mann–Whitney procedure fails as a test of medians. *The American Statistician*, **72(3)**, 278–286.

[2] Witkovský, V. (2024). CharFunTool: The characteristic functions toolbox (MATLAB). Available at GitHub: `https://github.com/witkovsky/CharFunTool`.

*Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovak Republic, `witkovsky@savba.sk`

Feedback on dilution in stochastic gene expression: a comparative study of single-cell and population frameworks

Iryna Zabaikina* Pavol Bokes† Abhyudai Singh‡

We consider a protein that negatively regulates the rate with which a cell grows. Since less growth means less protein dilution, this mechanism forms a positive feedback loop on the protein concentration. We couple the feedback model with a simple description of the cell cycle, in which a division event is triggered when the cell volume reaches a critical threshold. Following the division we either track only one of the daughter cells (single cell framework) or both cells (population framework). We use the piecewise deterministic Markov process to model a single cell; the population is formalised using the measure-valued Markov process, i.e., population is treated as a sequence of Markov processes. For both frameworks, we find an exact stationary joint distribution of protein concentration and cell volume.

We explore the consequences of dilution feedback on ergodicity, population growth rate, and the bias of the population distribution towards faster growing cells with less protein. Finally, we analyse the feedback effect on protein statistics and compare this effect with the unregulated model of gene expression.

Acknowledgments

This work was supported by the Slovak Research and Development Agency under contract no. APVV-18-0308, and VEGA grants 1/0339/21 and 1/0755/22.

References

[1] Zabaikina, I., Zhang, Z., Nieto, C., Bokes, P., Singh, A. (2023). Amplification of noisy gene expression by protein burden: An analytical approach. *American Control Conference (ACC)*, 2861–2866. IEEE.

[2] Zabaikina, I., Bokes, P., Singh, A. (2023). Joint distribution of protein concentration and cell volume coupled by feedback in dilution. *International Conference on Computational Methods in Systems Biology*, 253–268. Springer Nature Switzerland.

*Comenius University, Bratislava, Slovakia, iryna.zabaikina@gmail.com
†Comenius University, Bratislava, Slovakia, pavol.bokes@fmph.uniba.sk
‡University of Delaware, Newark, DE, USA, absingh@udel.edu

Estimators of serial covariance parameters in multivariate linear models

Ivan Žežula* Daniel Klein† Martin Singull‡

The work considers the estimation of unknown covariance parameters in the multivariate linear models when the covariance matrix has serial structure. Especially in the model known as growth curve model there is no closed form of the estimators of unknown parameters. Estimation process is either iterative or the estimators are derived as a root of the polynomial of third or higher degree. The method known as restricted expected multivariate least squares is used to obtain explicit estimators with good statistical properties.

Acknowledgments

This work was supported by the Slovak Research and Development Agency under the Contract No. APVV-21-0369, and by grant VEGA 1/0585/24.

*P. J. Šafárik University, Slovakia, ivan.zezula@upjs.sk
†P. J. Šafárik University, Slovakia, daniel.klein@upjs.sk
‡Linköping University, Sweden, martin.singull@liu.se

List of Participants

Friday I.	Agu	agu1@uniba.sk
Giacomo	Aletti	giacomo.aletti@unimi.it
Jaromír	Antoch	antoch@karlin.mff.cuni.cz
Matej	Benko	Matej.Benko@vutbr.cz
Dominik	Borovský	dominik.borovsky@student.upjs.sk
Andrej	Dvurečenskij	andrej.dvurecenskij@savba.sk
István	Fazekas	fazekas.istvan@inf.unideb.hu
Katarzyna	Filipiak	Katarzyna.filipiak@put.poznan.pl
Christian	Genest	christian.genest@mcgill.ca
Zdeňka	Geršlová	gerslovaz@math.muni.cz
Eva	Gurišová	umersekr@savba.sk
Laura	Hajzoková	laura.hajzokova@savba.sk
Marc	Hallin	mhallin@ulb.ac.be
Jozef	Hanč	jozef.hanc@upjs.sk
Martina	Hančová	martina.hancova@upjs.sk
Radoslav	Harman	radoslav.harman@fmph.uniba.sk
Daniel	Hlubinka	daniel.hlubinka@matfyz.cuni.cz
Martin	Hrba	hrba@karlin.mff.cuni.cz
Zuzana	Hübnerová	hubnerova@fme.vutbr.cz
Šárka	Hudecová	hudecova@karlin.mff.cuni.cz
Marie	Hušková	huskova@karlin.mff.cuni.cz
Markéta	Janošová	janosova@math.muni.cz
Ioannis	Kalogridis	ioannis.kalogridis@kuleuven.be
Stanislav	Katina	katina@math.muni.cz
Barbora	Kessel	barbora.kessel@gu.se
Daniel	Klein	daniel.klein@upjs.sk
Zuzanna	Krysiak	krysiak.zuzanna1@gmail.com
Lukáš	Laffér s	lukas.laffers@umb.sk
Matúš	Maciak	maciak@karlin.mff.cuni.cz
Ján	Mačutek	jmacutek@yahoo.com
Caterina	May	caterina.may@kcl.ac.uk
Ivan	Mizera	imizera@ualberta.ca
Malwina	Mrowińska	malwina.mrowinska@put.poznan.pl
Werner G.	Müller	werner.mueller@jku.at
Alba	Muñoz del Río	alba.munoz@uclm.es
Stanislav	Nagy	nagy@karlin.mff.cuni.cz
Ricardo	Negrete	Ricardo.Negrete@uclm.es
Ostap	Okhrin	ostap.okhrin@tu-dresden.de
Andrej	Pázman	andrej.pazman@fmph.uniba.sk
Michal	Pešta	pesta@karlin.mff.cuni.cz

Zuzana	Prášková	praskova@karlin.mff.cuni.cz
Hedvika	Ranošová	hedvika.ranosova@ovgu.de
Wolf-Dieter	Richter	wolf-dieter.richter@uni-rostock.de
Elena	Říhová	elena.rihova@vse.cz
Samuel	Rosa	samuel.rosa@fmph.uniba.sk
Lívia	Rosová	livia.rosova@fmph.uniba.sk
Zuzana	Rošťáková	zuzana.rostakova@savba.sk
František	Rublík	umerrubl@savba.sk
Seyed Yaser	Samadi	ysamadi@siu.edu
Ulrike	Schneider	ulrike.schneider@tuwien.ac.at
Pál	Somogyi	pal.somogyi@fmph.uniba.sk
Bernhard	Spangl	bernhard.spangl@boku.ac.at
Alexandra	Stadler	alexandra.stadler@jku.at
Gábor	Szűcs	gabor.szucs@fmph.uniba.sk
Germain	Van Bever	germain.vanbever@unamur.be
Ondřej	Vencálek	ondrej.vencalek@upol.cz
Martin	Wendler	martin.wendler@ovgu.de
Gejza	Wimmer	wimmer@mat.savba.sk
Viktor	Witkovský	witkovsky@savba.sk
Iryna	Zabaikina	iryna.zabaikina@gmail.com
Ivan	Žežula	ivan.zezula@upjs.sk