LinStat2020
High Tatras, Slovakia

International Conference on Trends and Perspectives in Linear Statistical Inference

Book of Abstracts

August 30 – September 3, 2021
Będlewo, Poland

Edited by

Daniel Klein
Institute of Mathematics
P. J. Šafárik University, Košice, Slovakia

and

Katarzyna Filipiak
Institute of Mathematics
Poznań University of Technology, Poland

Contents

Part I

Introduction

The International Conference on Trends and Perspectives in Linear Statistical Inference, LinStat'2020, was planned to be held on June 29 - July 3, 2020, in Nový Smokovec, High Tatras, Slovakia. Due to COVID-19 epidemic the conference was postponed to 2021 and will be organized in a hybrid form on August 30 - September 3, 2021. This is the follow-up of the LinStat series held in Będlewo, Poland (2008, 2012, 2018), in Tomar, Portugal (2010), in Linköping, Sweden (2014), and in İstanbul, Turkey (2016).

The purpose of the meeting is to bring together researchers sharing an interest in a variety of aspects of statistics and its applications as well as matrix analysis and its applications to statistics, and offer them a possibility to discuss current developments in these subjects.

The work of young scientists is highly appreciated, it has a special position in the LinStat2020 to encourage and promote them. The Scientific Committee will award the best presentation and best poster. The awarded will be Invited Speakers at the next edition of LinStat.

Special sessions topics:

- Application of Statistics
- Computationally Intensive Statistical Methods
- Design and Analysis of Experiments
- High-Dimensional Statistics
- Miscellaneous
- Multivariate Models
- Ordered Statistics
- Regression Models
- Small Area Estimation

The conference will include invited talks given by

- Adelchi Azzalini (Italy)
- Narayanaswamy Balakrishnan (Canada)
- Somnath Datta (USA)
- Radoslav Harman (Slovakia)
- Marie Hušková (Slovakia)
- Busenur Kızılaslan (Turkey)
- Samuel Rosa (Slovakia)
- Veronika Římalová (Czech Republic)

Organizers

- Institute of Mathematics, Faculty of Science, P. J. Šafárik University, Košice, Slovakia
- Institute of Mathematics, Poznań University of Technology, Poland
- Department of Computer and Information Science, Linköping University, Sweden
- Institute of Measurement, Slovak Academy of Science, Bratislava, Slovakia
- Institute of Plant Genetics, Polish Academy of Sciences, Poznań
- Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poland
- Banach Center, Institute of Mathematics, Polish Academy of Sciences, Poland

Committees

The Scientific Committee for this Conference comprises

- Ivan Žežula (Slovakia) - Chair
- Anthony C. Atkinson (UK)
- Augustyn Markiewicz (Poland)
- João Tiago Mexia (Portugal)
- Simo Puntanen (Finland)
- Dietrich von Rosen (Sweden)
- Müjgan Tez (Turkey)
- Götz Trenkler (Germany)
- Roman Zmyślony (Poland)

The Organizing Committee comprises

- Daniel Klein (Slovakia) - Chair
- Katarzyna Filipiak (Poland)
- Andrej Gajdoš (Slovakia)
- Adam Mieldzioc (Poland)
- Monika Mokrzycka (Poland)
- Jolanta Pielaszkiewicz (Sweden)
- Viktor Witkovský (Slovakia)

Call for papers

We are pleased to announce special issues of *Communications in Statistics – Theory and Methods* as well as *Case Studies, Data Analysis and Applications* (Taylor & Francis) devoted to LinStat2020.

They will include selected papers strongly correlated to the talks of the conference, with emphasis on advances on linear models and inference.

All papers submitted to *Communications in Statistic* must meet the publication standards of the journal:

`http://www.math.mcmaster.ca/bala/comstat/`

and will be subject to normal refereeing procedure. Papers should be prepared according to the instructions found on this page. More theoretical manuscripts should be sent to the *Theory and Methods* series while more computationally or data oriented manuscripts to the *Case Studies, Data Analysis and Applications* series. Please, remember that the purely linear algebraic papers but with no statistical relevance or applications will not be suitable for these two special issues since the journal *Communications in Statistics* does not publish non-statistical papers.

The deadline for paper submission: **31st of October 2021**

Coordinator-editor: N. Balakrishnan

Communications in Statistics - Theory and Methods

Special Issue Title: ADVANCES ON LINEAR MODELS AND INFERENCE

Guest Editors: Katarzyna Filipiak and Ivan Žežula

Topics to be covered: Inference in Linear Models, Mixed Linear Models and Multivariate Linear Models, Designing of Experiments in Linear Models, Matrix Theory and Linear Models, High-dimensional Analysis and Linear Models, Applications of Linear Models, Linear Models from a Bayesian Point of View, Linear Models as Approximations of Non-Linear Models, Linear Models and Nonparametrics.

Papers should be submitted using this portal. The contributors must choose submitting paper for a specific special issue under ”Manuscript Details” and to enter ”LinStat2020” in the box entitled ”Special Issue name”.

Communications in Statistics - Case Studies, Data Analysis and Applications

Special Issue Title: COMPUTATIONAL ISSUES AND APPLICATIONS OF LINEAR STATISTICS

Guest Editors: Simo Puntanen and Daniel Klein

Topics to be covered: Advances relating to Computational Aspects of the above mentioned topics: see Theory and Methods, Simulational Assessments, Graphical Methods, Resampling and other Computationally Intensive Methods, Recent Developments in Statistical Software, Data Analysis.

Papers should be submitted using this portal. The authors should select "Original Paper" as their article type in Step 1 of the submission process and then during Step 5 they will be asked if their paper is a candidate for a special issue. If they say "yes" they are asked to select the title of the special issue "LinStat2021" from a drop-down list.

Part II

Program

Monday, August 30, 2021

7:30 – 8:50 Breakfast

8:50 – 9:00 *Opening*

Plenary Session – chair: **I. Žežula**

9:00 – 9:45 **M. Hušková**: *Detection of instabilities in regression models*

9:45 – 9:50 Intro HyHyve
9:50 – 10:00 Break

Parallel Session A (ROOM A) – Small Area Estimation I
chair: **I. Žežula**

10:00 – 10:30 **N. Longford**: *Audit of neonatal mortality in Great Britain*
10:30 – 11:00 **F. Schirripa Spagnolo**: *Controlling the bias for M-quantile estimators in small area estimation*

Parallel Session B (ROOM B) – Computationally Intensive Statistical Methods I
chair: **U. Beyaztaş**

10:00 – 10:30 **F. Yerlikaya-Özkurt**: *A computational method with density based clustering approach for the data in the existence of outlier*
10:30 – 11:00 **Ö. Çağçağ Yolcu**: *Intuitionistic fuzzy time series: A cascade prediction model*
11:00 – 11:20 **M. Mutiş**: *A robust functional logistic regression method for classification*

11:20 – 12:00 Coffee Break - HyHyve

Special Session (ROOM A) – Miscellaneous I
chair: **M. Fonseca**

12:00 – 12:30 **S. Ferreira**: *Single models and structured families of pair of models*
12:30 – 13:00 **D. Ferreira**: *Inference in non normal mixed models*
13:00 – Lunch

Plenary Session – chair: **A. Markiewicz**

15:00 – 15:45 **N. Balakrishnan**: *Dimension reduction and variable selection for high-dimensional multivariate linear regression*

15:45 – 16:00 Break

Parallel Session A (ROOM A) – Multivariate Models I
chair: **D. von Rosen**

16:00 – 16:20 **E.U. Gasana**: *Moments of the discriminant function*
16:20 – 16:40 **M. John**: *Testing independence under the doubly multivariate models with block compound symmetry covariance structure using Rao score test*
16:40 – 17:00 **M. Janiszewska**: *Matrix approximation by block structures via entropy loss function*

Parallel Session B (ROOM B) – Design and Analysis of Experiments I
chair: **J. Forkman**

16:00 – 16:30 **A. Łacka**: *On a new approach to the ANOVA for experiments with orthogonal block structure. Experiments in row-column designs*
16:30 – 17:00 **R. Singh**: *Design selection for 2-level supersaturated designs*

17:00 – 17:20 Coffee break - HyHyve

Special Session (ROOM A) – Application of Statistics I
chair: **D. von Rosen**

17:20 – 17:40 **M-Z. Spyropoulou**: *A second order two-dimensional Gaussian Markov Random Fields in blood pressure data*

19:00 – Dinner

Tuesday, August 31, 2021

7:30 – 9:00 Breakfast

Plenary Session – chair: **A. C. Atkinson**

9:00 – 9:45 **A. Azzalini**: *A formulation for continuous mixtures of multivariate normal distributions*

9:45 – 10:00 Break

Parallel Session A (ROOM A) – Multivariate Models II
chair: **D. von Rosen**

10:00 – 10:30 **E. Di Nardo**: *Moment symbolic calculus in mathematical statistics*
10:30 – 11:00 **M. Singull**: *Likelihood based classification of growth curves*
11:00 – 11:20 **D. Uwamariya**: *On the ratio of extremal eigenvalues of β-Laguerre ensembles*

Parallel Session B (ROOM B) – Computationally Intensive Statistical Methods II
chair: **M. Tez**

10:00 – 10:30 **U. Beyaztaş**: *Robust function-on-function interaction regression*
10:30 – 10:50 **S. Sarıçam**: *On partial least squares estimation in functional regression models*
10:50 – 11:10 **B. H. Beyaztaş**: *A robust estimation of random effects panel data regression*
11:10 – 11:30 **F. Yerlikaya-Özkurt**: *Feature selection for mean shift outlier model via the conic-fused Lasso*

11:20 – 11:40 Coffee Break - HyHyve

Special Session (ROOM A) – Regression Models
chair: **E. Fišerová**

11:40 – 12:00 **S. Anatolyev**: *Testing many restrictions under heteroskedasticity*
12:00 – 12:20 **Ş. Gökmen**: *Specification of spatial models with measurement error: a simulation study*
12:20 – 12:40 **L. Novais**: *Variable selection in finite mixture of linear mixed models using the EM and CEM algorithms*
12:40 – 13:00 **S. Alica**: *Comparison of performances of heteroskedasticity tests in the presence of measurement error*

13:00 – Lunch

Plenary Session – chair: **S. Puntanen**

15:00 – 15:45 **S. Datta**: *Marginal inference under informative subgroup size induced by a subject level covariate*

15:45 – 16:00 Break

Parallel Session A (ROOM A) – Small Area Estimation II
chair: **D. von Rosen**

16:00 – 16:30 **D. von Rosen**: *Small area estimation using reduced rank regression*
16:30 – 16:50 **V. Pérez Giménez**: *Increasing the sample size by using rolling windows*

Parallel Session B (ROOM B) – High-Dimensional Statistics I
chair: **T. Pavlenko**

16:00 – 16:30 **T. Pavlenko**: *Detection of sparse and weak effects in high-dimensional data*
16:30 – 17:00 **R. Ahmad**: *Efficient correlation tests for high-dimensional data*

17:00 – 17:20 Coffee break - HyHyve

Special Session (ROOM A) – Application of Statistics II
chair: **I. Žežula**

17:20 – 17:40 **E. Derezea**: *A Saddlepoint approximation test for period detection: Application to lightcurves*
17:40 – 18:00 **L. Arnroth**: *A robustness evaluation of Bayesian tests for longitudinal data*

19:00 – Dinner

Wednesday, September 1, 2021

7:30 – 9:00 Breakfast

Plenary Session – chair: **V. Witkovský**

9:00 – 9:45 **R. Harman**: *Computational aspects of optimal experimental designs*

9:45 – 10:00 Break

Parallel Session A (ROOM A) – Application of Statistics III
chair: **V. Witkovský**

10:00 – 10:20 **J. Hauke**: *On multimodal distributions. Estimation methods and examples of applications in socio-economic geography*

10:20 – 10:40 **L. Batra**: *Renyi extropy: A generalization of information extropy*

Parallel Session B (ROOM B) – Design and Analysis of Experiments II
chair: **A. Markiewicz**

10:00 – 10:30 **O. Volkov**: *SARS-CoV-2 variants neutralisation studies to predict vaccines effectiveness*

10:30 – 10:50 **R. Khodsiani**: *Optimality of block designs under a hub correlation structure*

11:00 – 11:30 Coffee Break

Special Session (ROOM A) – Multivariate Models III
chair: **D. von Rosen**

11:30 – 11:50 **R. Zmyślony**: *Jordan algebra in estimation and testing hypotheses in multivariate normal models*

11:50 – 12:10 **A. Szczepańska-Álvarez**: *Likelihood ratio test for the model with block covariance structure*

12:10 – 12:30 **A. Mieldzioc**: *A comparison of the estimators of covariance matrix structured by Toeplitz matrix*

13:00 – Lunch

Plenary Session – chair: **K. Filipiak**

15:00 – 15:45 **S. Rosa**: *Greedy algorithms for computing informative saturated subsets*

15:45 – 16:00 Break

Special Session (ROOM A) – Miscalleneous II
chair: **S. Puntanen**

16:00 − 16:20 **Ł. Smaga**: *Projection tests for linear hypothesis in the functional response model*
16:20 − 16:40 **A. Markiewicz**: *Linearly sufficient and admissible estimators*

16:40 − 17:20 Poster Session (HyHyve)

S. Faria: *Fitting mixtures of linear mixed models: a case study involving fishery data*
M. Graczyk: *D-optimal designs with correlated errors*
A. Lukman: *K-L estimator in Beta regression model*
M. Mokrzycka: *The influence of the variance-covariance structure on the GWAS results*

19:00 − Conference Dinner

Thursday, September 2, 2021

7:30 – 9:00 Breakfast

Special Session (ROOM A) – Multivariate Statistics IV
chair: **D. von Rosen**

9:00 – 9:20 **I. Žežula**: *Testing the mean in three-level data with doubly-exchangeable covariance structure*
9:20 – 9:40 **K. Filipiak**: *How to measure discrepancy from structured matrices*

9:40 – 10:00 Break

Parallel Session A (ROOM A) – Ordered Statistics I
chair: **A. Goroncy**

10:00 – 10:30 **T. Rychlik**: *Conditions for the mixture representation of system lifetime distribution*
10:30 – 11:00 **M. Bieniek**: *Optimal linear combination of adjacent order statistics in quantile estimation*

Parallel Session B (ROOM B) – Computationally Intensive Statistical Methods III
chair: **Ö. Çağçağ Yolcu**

10:00 – 10:20 **V. Witkovský**: *Computational methods for probability distributions*
10:20 – 10:40 **Ş. Gökmen**: *Parameter estimation method of latent variable models: SIMEX approach*
10:40 – 11:00 **A. Gajdoš**: *Fast and accurate calculations of gamma difference distribution characteristics*

11:00 – 11:30 Coffee Break - HyHyve

Special Session (ROOM A) – Design and Analysis of Experiments III
chair: **O. Volkov**

11:30 – 12:00 **J. Forkman**: *Testing components of two-way interaction in bilinear models*
12:00 – 12:20 **M. Cavus**: *A new computational approach test for testing equality of several two-parameter exponential distributed means under unequal scale parameters*
13:00 – Lunch

Plenary Session – chair: **E. Fišerová**

15:00 – 15:45 **V. Římalová**: *Inference for spatial regression models with functional response using a permutational approach*

15:45 – 16:00 Break

Special Session (ROOM A) – Ordered Statistics II
chair: **K. Jasiński**

16:00 – 16:30	**A. Dembińska**: *Discrete time series-parallel systems*
16:30 – 17:00	**M. Szymkowiak**: *Signature conditions for distributional properties of system lifetimes if component lifetimes are iid exponential*

19:00 – Dinner

Friday, September 3, 2021

7:30 – 9:00	Breakfast

Special Session (ROOM A) – Ordered Statistics III
chair: **A. Dembińska**

10:00 – 10:30	**A. Goroncy**: *Upper estimates of kth record values based on the decreasing generalized failure rate samples*
10:30 – 11:00	**K. Jasiński**: *The number of failed components in a coherent working system when the lifetimes are discretely distributed*
11:00 – 11:30	Coffee Break - HyHyve

Special Session (ROOM A) – High-Dimensional Statistics II
chair: **J. Pielaszkiewicz**

11:30 – 12:00	**A. Tillander**: *Supervised feature selection with structure learning for sparse and weak data in high dimensional setting*
12:00 – 12:20	**D. Dai**: *Regularized Mahalanobis distance for high-dimensional data*
12:20 – 12:40	**J. Pielaszkiewicz**: *Comparison of chosen covariance structure tests with regards to sensitivity to outliers and violation of normality assumption in high-dimensional regime*
12:40 – 13:00	**D. Klein**: *Testing the compond symmetry structure in large- and high-dimensional setting*

13:00 – Closing

Part III

Invited Speakers

A formulation for continuous mixtures of multivariate normal distributions

Adelchi Azzalini

University of Padua, Italy

Abstract

Several formulations have long existed in the literature in the form of continuous mixtures of normal variables where a mixing variable operates on the mean or on the variance or on both the mean and the variance of a multivariate normal variable, by changing the nature of these basic constituents from constants to random quantities. More recently, other mixture-type constructions have been introduced, where the core random component, on which the mixing operation operates, is not necessarily normal.

The main aim of the present work is to show that many existing constructions can be encompassed by a formulation where normal variables are mixed using two univariate random variables. For this formulation, we derive various general properties, with focus on the multivariate context.

Within the proposed framework, it is also simpler to formulate new proposals of parametric families, and we provide a few such instances. As a side product, the exposition provides a concise compendium of the main constructions of continuous normal-mixtures type, although a full overview of this vast theme is not attempted.

Dimension reduction and variable selection for high-dimensional multivariate linear regression

Narayanaswamy Balakrishnan

McMaster University, Hamilton, Ontario, Canada

Abstract

This talk will have two parts. In the first part, I will develop reduced rank regression with matrix projections for high-dimensional multivariate linear regression model. After introducing the basic setup of the model, I will discuss the estimation methods and present some of their key properties, and then illustrate their performance through simulation studies as well as a real-life data analysis. Then, in the second part, I will develop envelope-based sparse reduced rank regression for high-dimensional multivariate linear model. I will describe the method of estimation and then present some optimality properties of the estimates. I will finally illustrate their performance through Monte Carlo simulation and a real-life data analysis. I will also compare the performance of the developed methods with some other well-known methods existing in the literature.

Keywords

High-dimensional multivariate linear regression model, Reduced rank regression, Envelope-based sparse reduced rank regression, Monte Carlo simulation.

Marginal inference under informative subgroup size induced by a subject level covariate

Somnath Datta

University of Florida, Gainesville, USA

Abstract

The Wilcoxon rank-sum test is a popular nonparametric test for comparing two independent populations (groups). In recent years, there have been renewed attempts in extending the Wilcoxon rank sum test for clustered data, some of which addresses the issue of informative cluster size, i.e., when the outcomes and the cluster size are correlated. We are faced with a situation where the group specific marginal distribution in a cluster depends on the number of observations in that group (i.e., the intra-cluster group size). We develop a novel extension of the rank-sum test for handling this situation. We compare the performance of our test with the Datta-Satten test, as well as the naive Wilcoxon rank sum test. Using a naturally occurring simulation model of informative intra-cluster group size, we show that only our test maintains the correct size. We also compare our test with a classical signed rank test based on averages of the outcome values in each group paired by the cluster membership. While this test maintains the size, it has lower power than our test. Extensions to multiple group comparisons and the case of clusters not having samples from all groups are also discussed. We apply our test to determine whether there are differences in the attachment loss between the upper and lower teeth and between mesial and buccal sites of periodontal patients. The last part of the talk considers an extension to informativeness caused by a subject specific continuous covariate.

Computational aspects of optimal experimental designs

Radoslav Harman

Comenius Univeristy, Bratislava, Slovakia

Abstract

In the talk, we will review various classes of algorithms proposed to compute efficient experimental designs. We also introduce our R package OptimalDesign which provides a toolbox for the computation of D-, A-, I-, and c-efficient exact and approximate designs of experiments on finite domains, for regression models with real-valued, uncorrelated observations, with multiple linear constraints on the vector of design weights. We will demonstrate how to apply the procedures to selected practical optimal design situations. In addition, we will show that some well-known problems from computational geometry, matrix theory and graph theory are special cases of the general problem of optimal experimental design. We will discuss to which extent these problems can be solved via existing algorithms used for computing optimum experimental designs.

Keywords

Experimental design, Minimum volume enclosing ellipsoid, Outliers, Hadamard hypothesis.

Detection of instabilities in regression models

Marie Hušková

Charles University, Prague, Czech Republic

Abstract

The talk will focus on detection of changes (more generally instabilities) in regression models and partially in time series. It starts with a toy example including an application to change detection in dependence of discharges on precipitation.
Then the talk will focus on polynomial regression with dependent errors and than continues with extensions to time series with trending regression. Finally, certain class of functional linear models when data are arriving sequentially (he so called on-line procedures) are presented.
The talk will cover theoretical results, simulation results as well as applications for all considered setups.

Keywords

Change point problem, Off-line and on-line procedures, Finite dimensional data and functional data, Regression models.

References

[1] Aue, A., Horváth, L., Hušková, M., and Kokoszka, P. (2008). Testing for changes in polynomial regression. *Bernoulli 14*, 637–660.
[2] Aue, A., Horváth, L., and Hušková, M. (2012). Segmenting mean-nonstationary time series via trending regressions.*Journal of Econometrics 168*, 367–381.
[3] Aue, A., Hoermann, S., Horváth, L., and Hušková, M. (2014). Dependent functional linear models with applications to monitoring structural change. *Statistica Sinica 24*, 1042–1073.

A new deep intuitionistic fuzzy time series forecasting method

Busenur Kızılaslan[1], **Eren Baş**[2] **and Erol Eğrioğlu**[2]

[1] Marmara University, Istanbul, Turkey
[2] Giresun University, Turkey

Abstract

Forecasting methods based on intuitionistic fuzzy sets are popular in fuzzy research communities. The intuitionistic fuzzy inference systems employ membership and non-membership values in the modelling of time series. In this study, a new intuitionistic fuzzy time series is introduced. The proposed method uses intuitionistic fuzzy c-means clustering algorithm to create intuitionistic fuzzy time series. Simple deep recurrent neural networks are employed for the modelling of intuitionistic fuzzy time series. The training of the deep recurrent neural network is made by using a modified particle swarm optimization. The performance of proposed method is investigated by using maximum and minimum temperature time series in Giresun, Turkey. The performance of the new method is compared with forecasts of "Turkish State Meteorological Service", machine learning and classical methods.

Keywords

Intuitionistic fuzzy sets, Deep learning, Particle swarm optimization, Intuitionistic fuzzy time series, Recurrent neural networks.

References

[1] Atanassov, K.T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems 20*, 87–96.
[2] Chaira, T. (2011). A novel intuitionistic fuzzy c-means clustering algorithm and its application to medical images. *Applied Soft Computing 11*, 1711–1717.
[3] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. *In Proceedings of the IEEE International Conference on Neural Networks*, 1942–1948.

Greedy algorithms for computing informative saturated subsets

Radoslav Harman and Samuel Rosa

Comenius Univeristy, Bratislava, Slovakia

Abstract

Optimal design or minimum-volume enclosing ellipsoid algorithms often require an initial solution that consists of a saturated subset of points – i.e., a subset whose size is equal to the number of dimensions of the points. The construction of informative saturated subsets is a non-trivial problem, as it deals with subsets that are in a sense "singular", and is often done by a random selection of points or by a regularized heuristic.
In this talk, which is based on [2], we examine greedy algorithms for constructing saturated subsets of points that attain high values of D-optimality criterion. We propose a computationally efficient form of the method described by [1] and a modification of the Kumar-Yıldırım method [3]. We provide theoretical properties and geometrical interpretations of the methods, and we show that they outperform the commonly used random and regularized initialization methods.

Keywords

Experimental design, D-optimal design, Minimum-volume enclosing ellipsoid, Subsampling, Greedy heuristic.

Acknowledgements

The author was supported by the Slovak Scientific Grant Agency [grant VEGA 1/0341/19].

References

[1] Galil, Z. and Kiefer, J. (1980). Time- and space-saving computer methods, related to Mitchell's DETMAX, for finding D-optimum designs. *Technometrics 22*, 301–313.
[2] Harman, R. and Rosa, S. (2020). On greedy heuristics for computing D-efficient saturated subsets. *Oper. Res. Lett. 48*, 122–129.
[3] Kumar, P. and Yıldırım, E.A. (2005). Minimum volume enclosing ellipsoids and core sets. *J. Optimiz. Theory App. 126*, 1–21.

Inference for spatial regression models with functional response using a permutational approach

Veronika Římalová and Eva Fišerová

Palacký University, Olomouc, Czech Republic

Abstract

The aim of this contribution is to introduce an approach to hypotheses testing in a functional linear model for spatial data. The inferential procedure will firstly be demonstrated on a functional-on-functional case of homoscedastic data recorded in time and space, with a real world application to the transportation research. This case study is focused on modelling the driving speed on an expressway interchange in Brno, Czech Republic. Then, the potential of permutation testing will be demonstrated under the presence of spatial correlation among functional observations, with all necessary considerations in methodology. The proposed method can deal with the spatial structure of data by building a permutation testing procedure on spatially filtered residuals of a spatial regression model. Indeed, due to the spatial dependence existing among the data, the residuals of the regression model are not exchangeable, breaking the basic assumptions of the Freedman and Lane permutation scheme. Instead, it is proposed here to base the permutation test on approximately exchangeable spatially filtered residuals, i.e. the variance-covariance structure of the residuals is estimated by variography and then the correlation of the residuals is removed by a spatial filtering. A simulation study is conducted, intended to evaluate the performance of the proposed method in terms of empirical size and power under different covariance settings. It will be shown that neglecting the residuals spatial structure in the permutation scheme, i.e., permuting the correlated residuals directly, yields a very liberal testing procedures, whereas the proposed procedure based on spatially filtered residuals is close to the nominal size of the test. The methodology is demonstrated on a real world data set on the yearly production of municipal waste in the Venice province in Italy, collected between years 1997 and 2011.

Keywords

Functional geostatistics, Functional-on-functional regression model, Permutation tests, Spatial correlation, Spatial functional regression model.

Acknowledgements

The author was supported by the project no. IGA_PrF_2020_015.

References

[1] Menafoglio, A., Secchi, P. and Dalla Rosa, M. (2013). A Universal Kriging predictor for spatially dependent functional data of a Hilbert Space. *Electronic Journal of Statistics 7*, 2209–2240.
[2] Ramsay, J.O. and Silverman, B.W. (2013). *Functional data analysis.* New York, NY, Springer.
[3] Římalová, V., Fišerová, E., Menafoglio, A., and Pini, A. (2021). Inference for spatial regression models with functional response using a permutational approach. *Journal of Multivariate Analysis.*

Part IV

Special Sessions

Design and Analysis of Experiments

Rosemary Bailey

University of St Andrews, UK

Abstract

Experiments are planned with the aim of making good use of resources to obtain answers to questions of interest: for example, which variety of wheat will give us the most bread, or which drug will best alleviate a medical condition. Thus design is inseparable from the proposed method of data analysis. Design questions include: how should we recruit participants into an observational study? If our treatment factors are quantitative, with larger quantities being more costly, which quantities should we test? If we have very many treatment factors but cannot test all combinations of their levels, which combinations should we include? How should we block, in order to be aware of inherent differences? Analysis questions include: How can interaction be distinguished from noise? How does the design affect the data analysis?

Invited speakers:

- Johannes Forkman (Sweden)
- Agnieszka Łacka (Poland)
- Rakhi Singh (USA)
- Oleg Volkov (UK)

Other speakers:

- Mustafa Cavus (Turkey)
- Razieh Khodsiani (Iran)

Ordered Statistics

Agnieszka Goroncy

Nicolaus Copernicus University, Toruń, Poland

Abstract

The session is devoted to various aspects of the ordered statistical random variables, in particular order statistics, record values, censored order statistics, sequential order statistics, generalized order statistics, their linear combinations and more. Recent results including, but not restricted to, characterizations, approximations, reliability theory and survival analysis, relations, bounds, distribution theory, testing, censoring and applications of ordered data will be presented.

Invited speakers:

- Mariusz Bieniek (Poland)
- Krzysztof Jasiński (Poland)
- Tomasz Rychlik (Poland)
- Anna Dembińska (Poland)
- Magdalena Szymkowiak (Poland)

Other speakers:

- Agnieszka Goroncy (Poland)

Small Area Estimation

Nicholas Longford

Imperial College London, UK

Invited speakers:

- Dietrich von Rosen (Sweden)
- Volodymyr Sarioglo (Ukraine)
- Francesco Schirripa Spagnolo (Italy)

Other speakers:

- Nocholas Longford (UK)
- Virgilio Pérez Giménez (Spain)

High-Dimensional Statistics

Tatjana Pavlenko

KTH Royal Institute of Technology, Stockholm, Sweden

Invited speakers:

- Rauf Ahmad (Sweden)
- Annika Tillander (Sweden)

Other speakers:

- Deliang Dai (Sweden)
- Daniel Klein (Slovakia)
- Tatjana Pavlenko (Sweden)
- Jolanta Pielaszkiewicz (Sweden)

Multivariate Models

Deitrich von Rosen

Swedish University of Agricultural Sciences, Uppsala, Sweden

Invited speakers:

- Elvira Di Nardo (Italy)
- Emelyne Umunoza Gasana (Sweden)
- Martin Singull (Sweden)
- Denise Uwamariya (Sweden)

Other speakers:

- Katarzyna Filipiak (Poland)
- Malwina Janiszewska (Poland)
- Mateusz John (Poland)
- Adam Mieldzioc (Poland)
- Anna Szczepańska-Álvarez (Poland)
- Ivan Žežula (Slovakia)
- Roman Zmyślony (Poland)

Computationally Intensive Statistical Methods

Müjgan Tez

Marmara University, Istanbul, Turkey

Abstract

This title covers some modern data analysis approaches including heuristic algorithms, linear and non-linear regression, parametric and non-parametric regression, artificial intelligence and classification techniques, soft computing, bootstrap and outlier detection methods.

Invited speakers:

- Ufuk Beyaztaş (Turkey)
- Fatma Yerlikayz-Özkurt (Turkey)
- Özge Çağçağ Yolcu (Turkey)

Other speakers:

- Beste Hamiye Beyaztas (Turkey)
- Andrej Gajdoš (Slovakia)
- Şahika Gökmen (Turkey)
- Müge Mutiş (Turkey)
- Semanur Sarıçam (Turkey)
- Viktor Witkovský (Slovakia)
- Fatma Yerlikaya-Özkurt (Turkey)

Applications of Statistics

Speakers:

- Lukas Arnroth (Sweden)
- Olgun Aydin (Poland)
- Luckshay Batra (India)
- Efthymia Derezea (UK)
- Jan Hauke (Poland)
- Robert Sebastian Castellanos Rodriguez (Colombia)
- Maria Zafeiria Spyropoulou (UK)

Regression Models

Speakers:

- Selin Alica (Turkey)
- Stanislav Anatolyev (Czech Republic)
- Şahika Gökmen (Sweden)
- Luísa Novais (Portugal)

Miscellaneous

Speakers:

- Dário Ferreira (Portugal)
- Sandra Ferreira (Portugal)
- Augustyn Markiewicz (Poland)
- Łukasz Smaga (Poland)

Part V

Abstracts

Efficient correlation tests for high-dimensional data

Rauf Ahmad

Uppsala University, Sweden

Abstract

A modified multivariate measure of correlation, composed of computationally efficient estimators, is suggested for high-dimensional low sample size scenario. The measure is subsequently used to construct a test of zero correlation for vectors of large dimension. Both the measure and the test are defined under a general multivariate model, with normality as a special case whence the test corresponds to that of independence. Simulations are used to demonstrate the size and power properties of the test, and comparison with distance correlation is shown.

Comparison of performances of heteroscedasticity tests under measurement error

Selin Alica[1], Şenay Açikgöz[1], Rukiye Dağalp[2] and Şahika Gökmen[1,3]

[1] Ankara Haci Bayram University, Turkey
[2] Ankara University, Turkey
[3] Uppsala University, Sweden

Abstract

One of the important assumptions of the classical regression analysis is homoscedasticity. It simplifies the standard error formulas of the ordinary least square (OLS) estimators and ensures that these estimators are efficient among the other linear estimators. Heteroskedastic errors disturb the efficiency properties of the OLS estimators and the estimated standard errors would be biased and wrong that make inferences misleading. Under the heteroscedastic errors, the OLS estimators could still be unbiased and consistent by weighting the data. However, estimations with errors-in-variables yield biased and inconsistent OLS estimators in either way. Specifically, measurement error on the explanatory variables causes biased and inconsistent OLS estimators that yield mistaken conclusions for hypothesis testing [2]. This study aims to compare performances of mostly used heteroscedasticity tests under the presence of measurement error on either the explanatory and/or the dependent variables. [4] made theoretical conclusions about the properties of heteroscedasticity tests under measurement error [4]. He did not perform a simulation study. [1, 3] compared the powers of heteroscedasticity tests while comparisons of [1] are under different scenarios for the functional patterns of conditional variance [1, 3]. To the best of our knowledge, this paper will be the first study that comprehensively examines the performances of heteroscedasticity tests under measurement error. Preliminary Monte Carlo simulations under different heteroscedasticity forms and sample sizes show that the Goldfeld-Quandt test has better performance when there is no measurement error in the explanatory variable in the simple linear regression model. These results are compatible with [1, 3] results. Simulations will be extended to show the performances of the heteroscedasticity tests under the presence of measurement error either the explanatory and/or dependent variables.

Keywords

Heteroscedasticity, Measurement error models, Error-in-variables, Model specification.

References

[1] Adamec, V. (2017). Power of heteroskedasticity tests in presence of various types of skedastic function and sample size. *AIP Conference Proceedings (Vol. 1863, No. 1)*, AIP Publishing LLC.

[2] Buonaccorsi, J.P. (2010). *Measurement error: models, methods, and applications.* Chapman and Hall/CRC.

[3] Uyanto, S.S. (2019). Monte Carlo power comparison of seven most commonly used heteroscedasticity tests. *Communications in Statistics - Simulation and Computation*, 1–18.

[4] Wooldridge, J. (1996). Solutions: Asymptotic Properties of Tests for Heteroskedasticity under Measurement Error. *Econometric Theory 12(2)*, 402–403.

Testing many restrictions under heteroskedasticity

Stanislav Anatolyev[1,2] and Mikkel Sølvsten[3]

[1] CERGE-EI, Prague, Czech Republic
[2] New Economic School, Moscow, Russia
[3] University of Wisconsin, Madison, USA

Abstract

We propose a hypothesis test that allows for many tested restrictions in a heteroskedastic linear regression model. The test compares the conventional F statistic to a critical value that corrects for many restrictions and conditional heteroskedasticity. The correction utilizes leave-one-out estimation to correctly center the critical value and leave-*three*-out estimation to appropriately scale it. Large sample properties of the test are established in an asymptotic framework where the number of tested restrictions may be fixed or may grow with the sample size and can even be proportional to the number of observations. We show that the test is asymptotically valid and has non-trivial asymptotic power against the same local alternatives as the exact F test when the latter is valid. Simulations corroborate the relevance of these theoretical findings and suggest excellent size control in moderately small samples also under strong heteroskedasticity.

Keywords

Linear regression, Ordinary least squares, Many regressors, Leave-out estimation, Hypothesis testing, High-dimensional models.

Acknowledgements

The first author acknowledges support by the grant 20-28055S from the Czech Science Foundation.

A robustness evaluation of Bayesian tests for longitudinal data

Lukas Arnroth and Rauf Ahmad

Uppsala University, Sweden

Abstract

Linear mixed models are standard models to analyse repeated measures or longitudinal data under the assumption of normality for random components in the model. Although the mixed models are often used in both frequentist and Bayesian inference, their evaluation from robustness perspective has not received as much attention in Bayesian inference as in frequentist. The aim of this study is to evaluate Bayesian tests in mixed models for their robustness to normality. We use a general class of exponential power distributions, EPD, and particularly focus on testing fixed effects in longitudinal models. The EPD class contains both light and heavy tailed distributions, with normality as a special case. Further, we consider a new paradigm of Bayesian testing decision theory where the hypotheses are formulated as a mixture model, with subsequent testing based on the posterior distribution of the mixture weights. It is shown that the EPD class provides a flexible alternative to normality assumption, particularly in the presence of outliers. Real data applications are also demonstrated.

A novel outlier detection approach for univariate datasets using deep neural networks

Olgun Aydin

Gdańsk University of Technology, Poland

Abstract

An outlier is a data point that appears distinctly different from other values in the dataset.Outliers occur due to changes in system behavior, human error, machine error, or natural deviations in populations.

Dealing with outliers is one of the key steps before moving forward with having any type of predictions. Before having predictions, building predictive models, outliers should be carefully analyzed in order not to reveal situations that could lead to potentially disastrous consequences [3]. Initially, the question of whether it is appropriate to remove outliers should be addressed [4]. Outliers in univariate datasets can be detected by using methods such as: location & scale based intervals, boxplot based methods, statistical tests and fitting mixture models to the dataset [1, 2].

Main goal of this study is proposing a different perspective on outlier detection methods for univariate datasets. The method is combination of location & scale based methods and deep neural networks.The proposed method and traditional methods have been performed on simulated data and it is observed that the proposed method perform better for detecting extreme values and outliers than the traditional methods.

Keywords

Outliers, Deep neural networks, Outlier detection.

References

[1] D'Orazio, M. (2017). Outlier Detection in R: Some Remarks, uRos 2017, Bucharest/ROMANIA

[2] D'Orazio, M. (2017). univOutl: Detection of Univariate Outliers. `https://cran.r-project.org/web/packages/univOutl/univOutl.pdf`

[3] Hodge, V.J. and Austin J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review 22*

[4] Scheff, S.W. (2016). Chapter 9 - Outliers and Missing Data, Editor(s): Stephen W. Scheff, *Fundamental Statistical Principles for the Neurobiologist*, Academic Press, 183–199, doi:10.1016/B978-0-12-804753-8.00009-9

Renyi extropy: A generalization of information extropy

Luckshay Batra and H.C. Taneja

Delhi Technological University, India

Abstract

In this paper, we have introduced Renyi extropy, the dual concept of one parametric Renyi entropy and have discussed its several meaningful properties like it is both positive and negative with specified range of its parameter; Renyi extropy reduces to information extropy for the limiting values of its parameter; if a probability space has two events, then Renyi extropy equals to Renyi entropy. We have also studied the relationship between Renyi extropy with information extropy and one parametric Tsallis extropy measures. In order to highlight the effectiveness of extropy measure, we plan to work out its application to analyse the uncertainty in the cryptocurrency market.

Keywords

Uncertainty, Extropy, Maximum extropy, Tsallis extropy, Renyi extropy, Cryptocurrency.

References

[1] Almarashi, S.K.G.E.A.M., Algarni, A., and Raqab, M. (2020). Quantum extropy and statistical properties of the radiation
eld for photonic binomial and even binomial distributions. *Journal of Russian Laser Research 41*, 334–343. doi:10.1007/s10946-020-09883-9

[2] Balakrishnana, M.L.N. and Buonob, F. (2021). On Tsallis extropy with an application to pattern recognition. arXiv.org. `doi:arXiv:2103.07168v1`

[3] Batra, L. and Taneja, H. (2020). Evaluating volatile stock markets using information theoretic measures. *Physica A: Statistical Mechanics and its Applications 537 (1).* doi:10.1016/j.physa.2019.122711

[4] Brody, D.C., Buckley, I.R.C., and Constantinou, I.C. (2007). Option price calibration from renyi entropy. *Physics Letters A 366 (4-5)*, 298–307. doi:10.1016/j.physleta.2007.01.088

[5] Johnson, D. and Glantz, R. (2004). When does interval coding occur? *Neurocomputing 58 (60)*, 13–18.

[6] Krishnan, A., Sunoj, S., and Sankaran, P. (2020). Some reliability properties of extropy and its related measures using quantile function. *Statistica 80(4)*, 413–437. doi:10.6092/issn.19732201/9887

[7] Lad, F., Sanfilippo, G., and Agro, G. (2015). Extropy: Complementary dual of entropy. *Statistical Science 30 (1)*, 40–58. doi:10.1214/14-STS430

[8] Qiu, G. (2017). The extropy of order statistics and record values. *Statistics and Probability Letters 120*, 52–60. doi:10.1016/j.spl.2016.09.016

[9] Qiu, L.C.W.G.X. and Wang, X.Y. (2019). On extropy properties of mixed systems. *Probability in the Engineering and Informational Sciences 33 (3)*, 471–486. doi:10.1017/S0269964818000244

[10] Raqab, M. and Qui, G. (2017). On extropy properties of ranked set sampling. *Statistics A Journal of Theoretical and Applied Statistics 53 (1)*, 210–226. doi:10.1080/02331888.2018.1533963

[11] Renyi, A. (1961). *On measures of entropy and information*. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 1, 547–561.

[12] Sathar, E. and Nair, R. (2021). On dynamic weighted extropy. *Journal of Computational and Applied Mathematics (113507)*. doi:10.1016/j.cam.2021.113507

[13] Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal 27 (3)*, 379–423, 623–656. doi:10.1002/j.1538-7305.1948.tb01338.x

[14] Schneider, T. (1997). Information content of individual genetic sequences. *Journal of Theoretical Biology 189 (4)*, 427–441.

[15] Tsallis, C. (1988). Possible generalization of boltzman-gibbs statistics. *Journal of Statistical Physics 52 (1-2)*, 479–487. doi:10.1007/BF01016429

[16] Tsallis, C. and Brigatti, E. (2004). Non-extensive statistical mechanics: a brief introduction. *Continuum Mechanics and Thermodynamics 16*, 223–235.

[17] Varma, R. (1966). Generalization of renyi entropy of order α. *Journal of Mathematical Sciences 1 (180)*, 34–48.

A robust estimation of random effects panel data regression

Beste Hamiye Beyaztaş

Istanbul Medeniyet University, Turkey

Abstract

Panel data regression models have become a general framework in empirical economics to model complex relationships between the economic variables by including unobserved individual-specific heterogeneity. Ordinary Least Squares (OLS) is one of the most commonly used techniques to estimate parameters of such models. However, OLS may produce biased and inefficient estimates in case of any departure from the model assumptions and in the presence of outliers. In this study, we propose a robust estimation strategy based on the M-estimator along with Tukey's bisquare loss function and a data-dependent tuning parameter selection algorithm to obtain efficient estimates within the random effects panel data models framework. The finite-sample performance of the proposed estimator is investigated by means of several Monte-Carlo simulations and a real-world data analysis. Our results reveal that the proposed estimator has significantly better performance over existing traditional and robust methods in the presence of outliers.

Keywords

Random effects, Panel regression, Robustness.

Robust function-on-function interaction regression

Ufuk Beyaztaş[1], **Han Lin Shang**[2] **and Abhijit Mandal**[3]

[1] Marmara University, Istanbul, Turkey
[2] Macquarie University, Sydney, Australia
[3] University of Texas at El Paso, USA

Abstract

We consider a function-valued response variable and multiple functional-valued predictor variables in a function-on-function regression. We assume the response and predictors reside in square-integrable Hilbert space with finite second-order moments. The function-on-function regression encompasses quadratic and interaction effects of the functional predictors; this model provides a more flexible form compared with the standard function-on-function regression. Despite the flexibility, the quadratic and interaction effects may be erroneous in the presence of outliers and affect the estimation accuracy in them and the main effect. We propose a robust method to effectively estimate the coefficient of the function-on-function quadratic and interaction regression. Some of its theoretical properties, such as consistency and influence function, are investigated. The finite-sample performance of the proposed method is also evaluated through a series of Monte Carlo studies and an empirical data analysis. The estimation and forecast accuracies of the proposal are compared favourably with several existing methods.

Keywords

τ-estimator, Quadratic term, Robust estimation.

Acknowledgements
This work was supported by The Scientific and Technological Research Council of Turkey (TUBITAK) (grant no: 120F270).

Optimal linear combination of adjacent order statistics in quantile estimation

Mariusz Bieniek and **Luiza Pańczyk**

Maria Curie Skłodowska University, Lublin, Poland

Abstract

We consider the classical problem of quantile estimation by appropriately chosen L–statistics, i.e. linear combinations of order staistics. We are especially interested in linear combinations of two adjacent order statistics. The most often used procedure is to form linear interpolation of the unknown quantile function. We introduce new criterion of optimality based on sharp bounds on the bias of the estimation of the quantile of given order by fixed order statistics. Then we show that the classical approach of linear interpolation is not the best possible.

Keywords

Quantile function, Quantile estimation, L–statistics.

A new computational approach test for testing equality of several two-parameter exponential distributed means under unequal scale parameters

Mustafa Cavus[1,2] **and Berna Yazıcı**[1]

[1] Eskişehir Technical University, Turkey
[2] Warsaw University of Technology, Poland

Abstract

Testing equality of the several populations' mean is one of the main statistical problems. Classical F test is the most powerful test for the solution of the problem when the assumptions, normality and variance homogeneity, hold. These assumptions may be violated in practice. For such cases, the researchers improved alternative solutions using normality transformation techniques and robust estimators. However, these solutions may not give powerful results in all cases. The problem has a large set of possible solutions when the normality assumption is violated. There are some tests which have been improved using Generalized p-Value [6], Parametric Bootstrap [2] and Fiducial Approach [3] methods for log-normal, inverse-normal and two-parameter exponential distribution in the presence of nuisance parameter in the literature [4]. In this study, a new test based on the Computational Approach method [5], is proposed to solve the problem of testing the equality of two-parameter exponentially distributed populations' means. The performance of the proposed test is compared with the alternatives in terms of penalized power [1]. As a result, the proposed test performed better especially for small samples. In addition, the proposed test has been applied on the real data sets, and thus its advantages over the alternatives have been demonstrated.

Keywords

ANOVA, Non-normality, Two-parameter exponential distribution, Computational approach test.

References

[1] Cavus, M, Yazici, B., and Sezer, A. (2019). Penalized power approach to compare the power of the tests when Type I error probabilities are different, *Communications in Statistics - Simulation and Computation*, doi:10.1080/03610918.2019.1588310

[2] Krishnamoorthy, K., Lu, F., and Mathew, T. (2007). A parametric bootstrap approach for anova with unequal variances: Fixed and random models. *Computational Statistics and Data Analysis 51*, 5731–5742.

[3] Li, X., Wang, J., and Liang, H. (2011). Comparison of several means: a fiducial based approach. *Computational Statistics and Data Analysis 55*, 1993–2002.

[4] Malekzadeh, A. and Jafari, A.A. (2019). Inference on the equality means of several two-parameter exponential distributions under progressively Type II censoring. *Communications in Statistics - Simulation and Computation*, doi: 10.1080/03610918.2018.1538452

[5] Pal, N., Lim, W.K., and Ling, C.H. (2007). A computational approach to statistical inferences. *Journal of Applied Probability and Statistics 2(1)*, 13–35.
[6] Weerahandi, S. (1995). ANOVA under unequal error variances. *Biometrics 51*, 589–599.

Regularized Mahalanobis distance for high-dimensional data

Deliangу Dai[1], Yuli Liang[2] and Jianxin Pan[3]

[1] Linnaeus University, Växjö, Sweden
[2] Örebro University, Sweden
[3] The University of Manchester, UK

Abstract

Estimating inverse covariance matrix (precision matrix) is an essential part of many statistical methods. This paper proposes a regularized estimator for the inverse covariance matrix. Modified Cholesky decomposition (MCD) is utilized to construct positive definite estimators. Instead of directly regularizing the inverse covariance matrix itself, we impose regularization on the Cholesky factor. The estimated inverse covariance matrix is used to build Mahalanobis distance (MD). The proposed method is evaluated by detecting outliers through simulations.

Keywords

Modified Cholesky decomposition, Mahalanobis distance, Shrinkage, Robust analysis.

References

[1] Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd ed). Wiley.
[2] Chouakria, A., Diday, E., and Cazes, P. (1998). Vertices principal components with an improved factorial representation. In: A. Rizzi, M. Vichi, and H.-H. Bock (Eds.), *Advances in Data Science and Classification* (pp. 397–402). Springer.
[3] Kushner, H.B. and Meisner, M. (1980). Eigenfunctions of expected value operators in the Wishart distribution. *Ann. Statist. 8*, 977–988.
[4] Olkin, I. (1997). A determinantal proof of the Craig-Sakamoto theorem. *Linear Algebra Appl. 264*, 217–223.

Discrete time series-parallel systems

Anna Dembińska[1] **and Serkan Eryilmaz**[2]

[1] Warsaw University of Technology, Poland
[2] Atılım University, Turkey

Abstract

The talk will be concerned with reliability properties of series-parallel systems when the component lifetimes are independent and have discrete failure time distributions. Under the assumption that each subsystem have identical components while different subsystems may have different types of components, we will present in particular exact distributions of the number of failed components at the time when the system fails. For the special case when the components have phase-type failure time distributions, matrix-based expressions will be given for the quantities under concern. The results will be used to obtain optimal configuration of the series-parallel system which is replaced at failure.

Keywords

Discrete lifetime distribution, Phase-type distribution, Series-parallel system, Reliability theory.

References

[1] Dembińska, A. and Eryilmaz, S. (2021). Discrete time series-parallel system and its optimal configuration. *Reliability Engineering & System Safety 215*, 107832. doi:10.1016/j.ress.2021.107832

A saddlepoint approximation test for period detection: application to lightcurves

Efthymia Derezea, **Alfred Kume and Dirk Froebrich**

University of Kent, UK

Abstract

Estimating accurately the period of time series is a commonly occurring problem. When the sampling of the data occurs at irregular time points, an extra level of complexity is added. Data like that occur in many fields such as biology, economics or as we will examine here, astronomical light curves. These are complex time series at irregularly spaced time observations. For a particular star, there is also additional information related to the measurement accuracy for each observation data point. In order to determine their cyclic behavior (since it is of high interest to the astronomy community), we examine a type of empirical periodogram described as follows:
1. Select a sequence of trial periods,
2. Fit each time a preferable model with a periodic assumption,
3. Evaluate some goodness of fit measure, which defines the periodogram entries for each period.
The values generated in step three above define some measure of validity for each period. As a rule of thumb, extreme values in such periodograms point to the valid periods. However, some understanding of the distribution of the periodogram is needed in order to determine what is extreme and what is not. Here we will focus on non parametric periodograms, in particular a weighted Gaussian process regression periodogram. Most importantly we will introduce a test for quadratic forms in order to detect valid periods using saddlepoint approximations, as a faster and more accurate alternative to simulation-based methods. Moreover, we will extend this approach to correlated noise backgrounds and finally we will present our results for some real light curves.

Keywords

Periodogram, Normal quadratic forms, Hypothesis testing, Gaussian process regression.

References

[1] Azzalini, A. and Bowman, A. (1993). On the use of nonparametric regression for checking linear relationships. *Journal of the Royal Statistical Society: Series B (Methodological) 55(2)*, 549–557.
[2] Kuonen, D. (1999). Miscellanea. Saddlepoint approximations for distributions of quadratic forms in normal variables. *Biometrika 86(4)*, 929–935.
[3] Thieler, A.M., Backes, M., Fried, R., and Rhode, W. (2013). Periodicity detection in irregularly sampled light curves by robust regression and outlier detection. *Statistical Analysis and Data Mining: The ASA Data Science Journal 6(1)*, 73–89.
[4] Wang, Y., Khardon, R., and Protopapas, P. (2012). Nonparametric Bayesian estimation of periodic light curves. *The Astrophysical Journal 756:67.*
[5] Williams, Ch.K. and Rasmussen, C.E. (2006). *Gaussian processes for machine learning.* Vol. 2. No. 3. Cambridge, MA: MIT press.

Moment symbolic calculus in mathematical statistics

Elvira Di Nardo

University of Turin, Italy

Abstract

In the last ten years, the employment of symbolic methods has substantially extended both the theory and the applications of statistics and probability. This contribution [1] reviews the development of a symbolic technique arising from classical umbral calculus, as introduced by Rota and Taylor in 1994 [3]. The usefulness of this symbolic technique is twofold. The first is to show how new algebraic identities drive in discovering insights among topics apparently very far from each other and related to probability and statistics. One of the main tools is a formal generalization of the convolution of identical probability distributions, which allows us to employ compound Poisson random variables in various topics that are only somewhat interrelated. Having got a different and deeper viewpoint, the second goal is to show how to set up algorithmic processes performing efficiently algebraic calculations. Some of the algorithms have been implemented in `R` [2]. In particular, the challenge of finding these symbolic procedures should lead to new methods, and it poses new problems involving both computational and conceptual issues. Evidence of efficiency in applying this symbolic method will be shown within statistical inference, parameter estimation and, more generally, problems involving multivariate functions. Recent connections within random matrices have extended the applications of the symbolic method.

Keywords

Umbral calculus, Composition of formal power series, Cumulants, Symmetric polynomials.

References

[1] Di Nardo, E. (2015). Symbolic calculus in mathematical statistics: a review. *Seminaire Lotharingien de Combinatoire 67*, Article B67a, pp. 72.

[2] Di Nardo, E. and Guarino, G. (2021). k-Statistics: Unbiased Estimators for Cumulant Products and Faà Di Bruno's Formula. Documents for `R`. `R` package version 2.1
`https://CRAN.R-project.org/package=kStatistics`

[3] Rota, G.-C. and Taylor, B.D. (1994). The classical umbral calculus. *SIAM J. Math. Anal. 25*, 694–711.

Inference in non normal mixed models

Dário Ferreira[1], **Sandra S. Ferreira**[1], **Célia Nunes**[1] **and João T. Mexia**[2]

[1] University of Beira Interior, Portugal
[2] Nova University of Lisbon, Portugal

Abstract

Linear models play a central role in statistical methods. A linear model that presents fixed effects factors and random effects factos is called mixed model. It is usually assumed that the random part of these models is normally distributed. However, in many situations, the assumption of normality is only for convenience. In this presentation we will show how to make inference in mixed models whose factors of random effects may or may not follow a normal distribution. A numerical example is included.

Keywords

Mixed models, Inference, Non normal distributions.

Acknowledgements

The author was supported by the Portuguese Foundation for Science and Technology through the projects UIDB/00212/2020 and UIDB/00297/2020.

References

[1] Antunes, P., Ferreira, S.S., Ferreira, D., and Mexia, J.T. (2020). Multiple additive models. *Communications in Statistics - Theory and Methods.* Published online: 08 Feb 2020. doi:10.1080/03610926.2020.1723636
[2] Demidenko, E. (2013). *Mixed Models: Theory and Applications with R.* Wiley, New York.
[3] Ferreira, D., Ferreira, S.S., Nunes, C., Fonseca, M., and Mexia, J.T. (2019). Chisquared and related inducing pivot variables: an application to orthogonal mixed models. *Communications in Statistics - Theory and Methods 48(22)*, 5445–5466.
[4] LaMotte, L.R. (1983). *Fixed, random and mixed-effects models in S.* Kotz and N. L. Johnson, Encyclopedia of Statistical Sciences, Wiley, New York, 137–141.
[5] Lindsey, J.K. (1997). *Normal Models, In: Applying Generalized Linear Models.* Springer Texts in Statistics, Springer, New York, NY.
[6] Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS.* Springer, New York.
[7] Sahai, H. and Ageel, M.I. (2000). *Analysis of Variance: Fixed, Random and Mixed Models.* Birkhäuser, Cambridge.
[8] R: The R Project for Statistical Computing, `https://www.r-project.org/`

Single models and structured families of pair of models

Sandra S. Ferreira[1], Célia Nunes[1], Dário Ferreira[1] and João T. Mexia[2]

[1] University of Beira Interior, Portugal
[2] Nova University of Lisbon, Portugal

Abstract

We present a theoretical framework to study individual additive models. Furthermore, we extend the results to the case where we consider d individual additive models. We are thus able to get a multiple model. We explain the different cases and use generalized least squares estimators for estimable vectors in the individual models. We also test hypotheses in multiple models.

Keywords

Additive models, Cumulants, Moments, Single models.

Acknowledgements

This work was partially supported by the Portuguese Foundation for Science and Technology through the projects:
UIDP/MAT/00212/2020 and UIDP/MAT/00297/2020.

References

[1] Antunes, P., Ferreira, S.S., Ferreira, D., and Mexia, J.T. (2020). Multiple additive models. *Communications in Statistics - Theory and Methods*. Published online: 08 Feb 2020. doi:10.1080/03610926.2020.1723636
[2] Craig, C.C. (1931). On A Property of the Semi-Invariants of Thiele. *Ann Math Statist 2(2)*, 154–164.
[3] Fonseca, M., Mexia, J.T., and Zmyślony, R. (2006). Binary operations on Jordan algebras and orthogonal normal models. *Linear Algebra Appl. 417*, 75–86.
[4] Mexia, J.T. (1987). Multi-treatment regression designs. Trabalhos de investigação N. 1, Mathematics Department, Faculty of Sciences and Technology, Nova University of Lisbon.
[5] Mexia, J.T. (1988). Standardized Orthogonal Matrices and the Decomposition of the sum of Squares for Treatments, Research Studies N.2, Mathematics Department, Faculty of Sciences and Technology, Nova University of Lisbon.
[6] Moreira, E., Ribeiro, A., Mateus, E., Mexia, J., and Ottosen, L. (2005). Regressional modeling of electrodialytic removal of Cu, Cr and As from CCA treated timber waste: application to sawdust. *Wood Science and Technology 39(4)*, 291–305.
[7] Ribeiro, A. and Mexia, J.T. (1997). A dynamic model for the electrokinetic removal of copper from a polluted soil. *Journal of Hazardous Materials 56*, 257–271.
[8] Scheffé, H. (1959). *The Analysis of Variance*. John Willey & Sons, New York.
[9] Stuart, A. and Ord, K. (1994). *Kendall's Advanced Theory of Statistics, Volume1: Distribution Theory*. 6^{th} Edition. New Jersey:Wiley-Blackwel.

How to measure discrepancy from structured matrices?

Katarzyna Filipiak[1], Daniel Klein[2] and Monika Mokrzycka[3]

[1] Poznań University of Technology, Poland
[2] P. J. Šafárik University, Košice, Slovakia
[3] Polish Academy of Sciences, Poznań, Poland

Abstract

In this talk doubly-multivariate model of experiment, in which for every individual several features are measured repeatedly, is considered. Two measures of discrepancy - entropy and quadratic loss functions - are proposed and compared in the context of power comparison, when the hypothesis about separable structure of covariance is studied. The likelihood ratio test and Rao score test statistics are used for the verification of hypothesis of interest.

Keywords

Doubly-multivariate model, Separable structure, Entropy loss function, Quadratic loss function, Likelihood ratio test, Rao score test, Power.

Testing components of two-way interaction in bilinear models

Johannes Forkman

Swedish University of Agricultural Sciences, Uppsala, Sweden

Abstract

Experiments with two factors are commonly analysed using two-way analysis of variance, where testing significance of interaction is straightforward. However, using bilinear models, interaction can be analysed further. The "additive main effects and multiplicative interaction" (AMMI) model uses singular value decomposition for partitioning interaction into multiplicative terms, such that the first terms typically account for a large portion of the sum of squares, whereas the last terms are of minor importance. A recurring question is how to determine the number of terms to retain in the model. If data is replicated, which is usually the case, the F_R test [3] can be used for this purpose. The simple parametric bootstrap method [2] is another option, although this test was developed for unreplicated data. Since both these tests of significance may be applied in cases with replication, researchers need advice on which of the methods to use. It will be shown that the two methods address slightly different questions.

The AMMI model is related to the "genotype main effects and genotype-by-environment interaction effects" (GGE) model and principal component analysis. However, when variables are standardized to unit variance, a full parametric bootstrap method [1] is required for testing the significance of the principal components.

Keywords

Analysis of experiments, Genotype-by-environment interaction, Principal component analysis.

References

[1] Forkman, J., Josse, J., and Piepho, H.P. (2019). Hypothesis tests for principal component analysis when variables are standardized. *J. Agric. Biol. Environ. Stat. 24*, 289–30.

[2] Forkman, J. and Piepho, H.P. (2014). Parametric bootstrap methods for testing multiplicative terms in GGE and AMMI models. *Biometrics 70*, 639–647.

[3] Piepho, H.P. (1993). Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials. *Theor. Appl. Genet. 90*, 438–443.

Fast and accurate calculations of gamma difference distribution characteristics

Martina Hančová, Andrej Gajdoš and Jozef Hanč

P. J. Šafárik University, Košice, Slovakia

Abstract

The distribution of the difference of two independent gamma random variables, in general with different shape parameters, is called gamma difference distribution (GDD).
It can be found in many recent applications such as controlling a measurement accuracy of optical detectors [3], detecting radar sensor threshold [7], setting optimal performance of wifi networks [4], chemotherapy cancer treatment [8], detecting eye glaucoma [1], etc.
Despite the wide applicability of GDD, the fast and precise calculation of particular values of probability density function and cumulative distribution function with unequal GDD shape parameters is still quite challenging. However, the need for accurate, reliable, numerically stable, and fast computations arises naturally in any real data analysis or computational research (e.g. Monte Carlo or bootstrap methods).
Therefore, in this work we investigate the four principal computational ways for GDD in the plethora of currently available computing tools. We pay special attention to open digital tools based on programming languages Python and R, which became significant during the last decade, in the light of enormous advances of open data science. Our focus is on GDD with unequal shape parameters resulting from time series kriging [2], a forecasting approach based on the best linear unbiased prediction and linear mixed models.
We proposed a combination of numerical inversion of the characteristic function [9] and the trapezoidal rule with the double exponential oscillatory transformation (DE quadrature) [6] and we implemented our open-source tool in high-performance Python (with Numba). The results of our numerical study, with emphasis on using open data science tools, demonstrate that it is exponentially fast, highly accurate, and very reliable.
Our approach could be potentially used in the future, e.g. for distributions expressible by linear combinations of characteristic functions in fields such as multidimensional statistics, measurement uncertainty analysis in metrology as well as in financial mathematics and risk analysis.

Keywords

Numerical inversion of the characteristic function, Double exponential quadrature, Computational tools, High-performance Python, Econometrics, Time series, Kriging.

Acknowledgements

This work was supported by the Slovak Research and Development Agency under the contract no. APVV-17-0568 and the Internal Research Grant System of Faculty of Science, P. J. Šafárik University in Košice - project vvgs-pf-2020-1423.

References

[1] Belghith, A., Bowd, C., Medeiros, F.A., et al. (2015). Learning from healthy and stable eyes: A new approach for detection of glaucomatous progression. *Artif Intell Med. 64(2)*, 105–115.

[2] Hančová, M., Gajdoš, A., and Hanč, J. (2017). Kriging Methodology and Its Development in Forecasting Econometric Time Series. *Statistika 97(1)*, 59–73.

[3] Hendrickson, A.J. (2017). Centralized inverse-Fano distribution for controlling conversion gain measurement accuracy of detector elements. *JOSA A. 34(8)*, 1411–1423.

[4] Khan, T.A., Heath, R.W. Jr, and Popovski, P. (2017). Wirelessly Powered Communication Networks With Short Packets. *IEEE Transactions on Communications 65(12)*, 5529–5543.

[5] Klar, B. (2014). A note on gamma difference distributions. *J Stat Comput Simul 85(18)*, 3708–3715.

[6] Mori, M. and Sugihara, M. (2001). The double-exponential transformation in numerical analysis. *Journal of Computational and Applied Mathematics 127(1-2)*, 287–296.

[7] Ranney, K., Tom, K., Tadas, D., et al. (2021). An efficient pulse detector and pulse width estimator for waveform characterization. *Radar Sensor Technology XXV*; Vol. 11742. International Society for Optics and Photonics. p. 117421E.

[8] Sekhavati, F. (2015). Dynamic response of individual cells in heterogeneous population Ph.D. Thesis. München: LMU München.

[9] Witkovský, V. (2016). Numerical inversion of a characteristic function: An alternative tool to form the probability distribution of output quantity in linear measurement models. *ACTA IMEKO 5(3)*, 32–44.

Moments of the discriminant function

Emelyne Umunoza Gasana[1,2], **Martin Singull**[1] **and Dietrich von Rosen**[3]

[1] Linköping University, Sweden
[2] University of Rwanda, Rwanda
[3] Swedish University of Agricultural Sciences, Uppsala, Sweden

Abstract

Let $\mathbf{x}$ be a p-dimensional random vector $\mathbf{x} = (x_1, \ldots, x_p)^\intercal$. In classification methodology we wish to allocate an observation $\mathbf{x}$ to one of populations π_i, $i \in \{1, \ldots, q\}$. In this paper, classification rules and their moments for a multivariate normal population are established.

Keywords

Classification, Linear discriminant analysis, Moments of discriminant function.

Parameter estimation method of latent variable models: SIMEX approach

Şahika Gökmen[1,2] and Johan Lyhagen[2]

[1] Ankara Hacı Bayram Veli University, Turkey
[2] Uppsala University, Sweden

Abstract

It is widely accepted that economic variables cannot be measured, directly observed or only measured with an indirect variable. In any of case, errors in the measurement process are inevitable too. These situations cause problems from the very beginning (compilation of data) for a valid analyzes of economic data. If only indirect measurement is possible for any non-measurable variables such as the productivity of a worker, permanent income, etc, these variables are described as "latent"[6]. Thereby, according to [6], latent variable consists of the observed variable and its difference from its latent counterpart. On the other side, the measurement error definition is more general and refers to the difference between the observed variable and its real value. With regard to measurement error, the observed variable in place of the true variable is named errors-in-variable or error-prone variable.

In the literature, latent variables are incorporated in the analysis via Structural Equation Models (SEM). However, since the error-prone variables cause biased parameter estimations in regression models, ignoring measurement error can lead to serious problems. In this context the statistical models used to analyze mismeasured data are called measurement error models or error-in-variables model. Among the measurement error models, Simulation-Extrapolation (SIMEX) method is considered as one of the most successful one in the lately related field. SIMEX method is a unbiased parameter estimation technique based on simulation improved by [1].

Among the current studies in the literature, the differences between these concepts are frequently mentioned and the importance of this emphasized. Especially [2] noted in his study that the relationship between latent variables and their measures are important in defining the observed and unobserved relationships in economic models. Although not through SEM, [3] used the SIMEX method for binary variables with a focus on classification success. In addition, even though [5] considers the processing of estimation methods in measurement error models on SEM, there is no methodological improvement was made such as updating SIMEX. [4], on the other hand, deal with latent variables and measurement error separately, but their research only contains an application on the education system through the study of [3]. Overall, study of [2] draws attention as the first social science study that takes into account business economics in this field. In his study, latent variables considered as mismeasured and then some measurement error estimation methods are applied. However, a concrete adaptation of SIMEX for the estimation of SEM has not been proposed yet. Thereby, in this study, the SIMEX method, which is very effective and popular estimation method in measurement error models, is adapted to a new field. It is aimed to obtain unbiased estimations for regression models with a latent explanatory variable in which it is taken into consideration that latent variable also is an error-prone variable. Combining measurement errors and latent variables together plays an important role in terms of providing more reliable results for both survey studies and economic analyzes.

The main purpose of this study is that the theoretical adaptation of SIMEX method in terms of SEM will be discussed. Also, the adaptation will be supported by simulation studies under different conditions. In this way, it is expected that the study will serve as a bridge between SEM and measurement error models and it will contribute to the estimation process of SEM.

Keywords

Structural equation models, Error-prone variables, SIMEX.

Acknowledgements

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under the TUBITAK BIDEB 2219-International Postdoctoral Research Scholarship Programme.

References

[1] Cook, J.R. and Stefanski, L.A. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical association 89*, 1314–1328.

[2] Hu, Y. (2017). The Econometrics of Unobservables: Applications of Measurement Error Models in Empirical Industrial Organization and Labor Economics. *Journal of Econometrics 200*, 154–168.

[3] Küchenhoff, H., Mwalili, S.M., and Lesaffre, E. (2006). A General Method for Dealing with Misclassification in Regression: The Misclassification SIMEX. *Biometrics 62(1)*, 85–96.

[4] Rutkowski, L. and Zhou, Y. (2015). Correcting Measurement Error in Latent Regression Covariates via the MC-SIMEX Method. *Journal of Educational Measurement 52(4)*, 359–375.

[5] Simonetto, A. (2008). *Estimation Procedures for Latent Variable Models with Psychological Traits*. PhD Thesis, Universita Degli Studi Di Milano - Bicocca, Milano.

[6] Wansbeek, T. and Meijer, E. (2000). *Measurement Error and Latent Variables in Econometrics.* North Holland, Elsevier.

Specification of spatial models with measurement error: A simulation study

Anıl Eralp[1], Şahika Gökmen[2] and Rukiye Dağalp[3]

[1] Bolu Abant İzzet Baysal University, Turkey
[2] Ankara Hacı Bayram University, Turkey
[3] Ankara University, Turkey

Abstract

In the literature, the response and explanatory variables examine over the observations and random errors that arise during the observation process. The variables that cannot exactly measure are called error-in-variables and are analyzed with measurement errors models. Besides some variables are not measurable directly with a measurement unit due to their nature and are measured indirectly and so-called latent variables, and in these kinds of variables, measurement error is also encountered [9]. Therefore, measurement errors could be encountered in both direct and indirect measurements of the relevant variable. When the measurement error is generally mentioned, it is that arises from the measurement in observing the true value of variables. In this context, the difference between the true value of observation and its observed values gives the measurement error. As a simple explanation, the random variable X^*, called error-prone, is observed with a measurement error U instead of the real variable X that is $X^* = X + U$ defined by the classical error model. The measurement error U is assumed a random variable with a distribution of $N(0;\sigma_U^2)$ and unrelated to X [2]. Therefore, the effects of measurement error cause inconsistent parameter estimation and so the inconsistent prediction in a regression model analyzing with error-prone covariates [7, 3]. The effect of the measurement error in the explanatory variables is much more important than the measurement error in the response variable on the parameter estimation. The results of statistical analysis are reliable, useful, and meaningful depending on the measurements of the data that is used, although there is an opinion that measurement errors are small enough to be neglected and can be ignored in the analysis in practice. In contrast, the view of measurement errors ignored by studies in this area has been decreasing in recent years. There are lots of studies for both indicate the effects of measurement error through the different models and improve new methodologies to estimate measurement error models.

Observations for a random variable are assumed to be independent of each other in conventional regression models using cross-section data sets. However, the observations obtained from points in regional locations, spatial dependence is often encountered between observations. Spatial dependence is that observation values in a location tend to be similar to observation values in nearby locations; spatial regression methods are used to explain the cause of this dependence [6]. In parallel with the increase in the production of spatial data, the importance of measurement errors has raised with the increase in the use of spatial models. However, it is seen that error-in-variables are handled in relatively few studies investigating spatial models using spatial data have stressed (e.g., [5, 4, 8]. These studies investigate spatial regression models using spatial data with the classical measurement error models. In case, [1] discuss the model estimation presence of measurement error that occurs actually the unobservable covariate but is measured by spatial interpolation. This study can be taken as an example of the measurement error caused by a latent variable.

Keywords

Measurement error, Spatial regression models, Simulation.

References

[1] Anselin, L. and Lozano-Gracia, N. (2009). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Spatial Econometrics Methods and Applications* (editors Giuseppe Arbia and Badi H. Baltagi) Physica-Verlag Heidelberg, 5–34.
[2] Carroll, R.J., Ruppert, D., Stefanski, L.A., and Crainiceanu, C.M. (2006). *Measurement error in nonlinear models: a modern perspective.* (2nd ed) CRC press.
[3] Fuller, W.A. (1987). *Measurement error models.* John Wiley & Sons.
[4] Huque, Md H., Bondell, H.D., and Ryan, L. (2014). On the impact of covariate measurement error on spatial regression modelling. *Environmetrics 25(8)*, 560–570.
[5] Le Gallo, J. and Fingleton, B. (2012). Measurement errors in a spatial context. *Regional science and urban economics 42(1-2)*, 114–125.
[6] LeSage, J.P. (2008). An Introduction to Spatial Econometrics. *Revue d'économie industrielle 123(3e)*, 19–44.
[7] Stefanski, L.A. (1985). The effects of measurement error on parameter estimation. *Biometrika 72(3)*, 583–592.
[8] Suesse, T. (2018). Estimation of spatial autoregressive models with measurement error for large data sets. *Computational Statistics 33(4)*, 1627–1648.
[9] Wansbeek, T. and Meijer, E. (2000). *Measurement Error and Latent Variables in Econometrics.* Elsevier.

Upper estimates of kth record values based on the decreasing generalized failure rate samples

Agnieszka Goroncy

Nicolaus Copernicus University, Toruń, Poland

Abstract

We present the upper nonpositive bounds on expectations of the standardized kth record values, $k \geq 1$, which are based on the decreasing generalized failure rate distributions (DGFR(α)). Such a family of distributions is defined in terms of the convex transform order of the baseline distribution function with respect to the generalized Pareto distribution (GPD) function.

Keywords

Record values, Optimal bound, Generalized failure rate, Generalized Pareto distributions.

References

[1] Bieniek, M. and Goroncy, A. (2020). Sharp lower bounds on expectations of gOS based on DGFR distributions. *Stat. Papers 61(3)*, 1027–1042.
[2] Goroncy, A. (2020). On the upper bounds on expectations of gOSs based on DFR and DFRA distributions. *Statistics 54(2)*, 402–414.
[3] Goroncy, A. (2017). Upper non-positive bounds on expectations of generalized order statistics from DD and DDA populations. *Comm. Statist.-Theory Meth. 46(24)*, 11972–11987.

On multimodal distributions. Estimation methods and examples of applications in socio–economic geography

Jan Hauke

Adam Mickiewicz University, Poznań, Poland

Abstract

The distribution modality may be unimodal or bimodal, less often multimodal, depending on the frequency of the values of the analyzed random variable. In his classification system (AJUS) for distributions, Galtung [3] distinguished two following types of multimodal distributions: U - bimodal, with peaks at both ends, and S - bimodal or multimodal, with multiple peaks. The problem of the statistical analysis of phenomena for which the empirical distributions could suggest departure from unimodality was noticed already at the end of the nineteenth century, when theoretical bases for statistics began to be constructed. Pearson [9] was the first who in 1894 proposed the procedure of the dissection of abnormal frequency-curves into normal curves. From then, during the next several decades, bimodal (also multimodal) distributions were estimated mainly via the mixture of normal distributions. The situation began to change with the development of computerization and the possibilities of numerical analyses of even very complicated statistical models. During the last twenty years, different bimodal and skew distributions have been proposed. The aim of the paper is to expand the idea proposed by [1] during Linstat2018 conference (based on a specific exponential distribution) concerning U and S type of distributions and compare with the results of other authors (as [4, 5, 6, 11]) taking into account some empirical applications of such models (e.g. in [2, 7, 8, 10, 12]), including applications in socio–economic geography.

Keywords

Multimodal distribution, Bimodal distribution, Exponential distribution.

Acknowledgements

The author was supported by the project ID-UB UAM no. 001/06/POB5/0025.

References

[1] Bivand, R. and Hauke, J. (2018). On bimodal exponential distributions. *Linstat2018, Book of Abstract*, 83–84.

[2] Deo, N. and Trivedi, M.M. (2018). Convolutional social pooling for vehicle trajectory prediction. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1468–1476.

[3] Galtung, J. (1967). *Theory and methods of social research.* Oslo: Universitetsforlaget.

[4] Grasman, R.P.P.P., van der Maas, H.L.J., and Wagenmakers, E. J. (2009). Fitting the Cusp Catastrophe in R: A cusp Package Primer. *J. Stat. Softw. 32(8)*, 1–28.

[5] Hassan, M.Y. and Hijazi, R.H. (2010). A bimodal exponential distribution. *Pak. J. Statist. 26*, 379–396.

[6] Hassan, M.Y. and El-Bassiouni, M.Y. (2016). Bimodal skew-symmetric normal distribution. *Commun. Stat. Theory Methods 45*, 1527–1541.

[7] Lötsch, J., Dimova, V., Lieb, I., Zimmermann, M., Oertel, B. G., and Ultsch, A. (2015). Multimodal distribution of human cold pain thresholds. *PLoS One, 10(5)*, e0125822.

[8] Nikezić, D. and Stevanović, N. (2007). Room model with three modal distributions of attached 220Rn progeny and dose conversion factor. *Radiation protection dosimetry, 123(1)*, 95–102.

[9] Pearson, K. (1894). Contributions to the mathematical theory of evolution: On the dissection of asymmetrical frequency-curves. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 185*, 71–90.

[10] Rodriguez, N.B., Benettin, P., and Klaus, J. (2020). Multimodal water age distributions and the challenge of complex hydrological landscapes. *Hydrological Processes 34(12)*, 2707–2724.

[11] Vila, R. and Çankaya, M.N. (2021). A bimodal Weibull distribution: properties and inference. *Journal of Applied Statistics*, 1–19.

[12] Wua, Y., Zhoud, C., Xiao, J., Kurthsa, J., and Schellnhuber, H.J. (2010). Evidence for a bimodal distribution in human communication. *Proc. Natl. Acad. Sci. USA 107*, 18803–18808.

Matrix approximation by block structures via entropy loss function

Malwina Janiszewska[1], **Augustyn Markiewicz**[1] **and Monika Mokrzycka**[2]

[1] Poznań University of Life Sciences, Poland
[2] Polish Academy of Sciences, Poznań, Poland

Abstract

The need to study the relationship between different groups of features leads to the analysis of multivariate data sets. This association is described by the covariance matrix, but in case when the number of observations is not large enough, the covariance matrix is singular or ill-conditioned. To overcome this problem we use structural restrictions to covariance matrix.
The aim of this talk is to present a procedure of approximation of symmetric positive definite matrix by symmetric block partitioned matrices with structured off-diagonal blocks. We study two structures of off-diagonal blocks: as a part of compound symmetry or first-order autoregression matrix. As an approximation criterion the entropy loss function is used. The approximation procedure and statistical problem of covariance structure identification is discussed in [1].

Keywords

Matrix approximation, Block covariance structure, Entropy loss function.

References

[1] Janiszewska, M., Markiewicz, A., and Mokrzycka, M. (2020). Block matrix approximation via entropy loss function. *Applications of Mathematics 65(6)*, 829–844.

The number of failed components in a coherent working system when the lifetimes are discretely distributed

Krzysztof Jasiński

Nicolaus Copernicus University, Toruń, Poland

Abstract

We study the number of failed components of a coherent system. We consider the case when the component lifetimes are discrete random variables that may be dependent and non-identically distributed. Firstly, we compute the probability that there are exactly $i=0,\dots,n-k$, failures in a k-out-of-n system under the condition that it is operating at time t. Next, we extend this result to other coherent systems. In addition, we show that, in the most popular model of independent and identically distributed component lifetimes, the obtained probability corresponds to the respective one derived in the continuous case and existing in the literature. All details can be found in [1].

Keywords

Coherent system, K-out-of-n system, Discrete lifetime distribution, Reliability, Order statistics.

References

[1] Jasiński, K. (2021). The number of failed components in a coherent working system when the lifetimes are discretely distributed. *Metrika*, 1–14. doi:10.1007/s00184-021-00817-2

Testing independence under the doubly multivariate models with block compound symmetry covariance structure using Rao score test

Katarzyna Filipiak[1], Mateusz John[1] and Daniel Klein[2]

[1] Poznań University of Technology, Poland
[2] P. J. Šafárik University, Košice, Slovakia

Abstract

The goal of this talk is to verify the hypothesis related to independence of features between any two time points in block compound symmetry structure in doubly multivariate model.
The Rao score test was determined for such hypothesis. It was compared with the test which is based on the likelihood ratio, F-test [1] and Roy's Largest Root test. These tests were compared in terms of their power and asymptotic convergence of their distributions for various number of individuals, features or time points.
Finally, Rao score test was applied to real data example.

Keywords

Testing, Rao score test, Likelihood ratio test, F-test, Roy's test.

Acknowledgements

This work was supported by the project no. 0213/SIGR/2154 (Mateusz John) and by the Slovak Research and Development Agency under the contract no. APVV-17-0568 (Daniel Klein).

References

[1] Fonseca, M., Kozioł, A., and Zmyślony, R. (2018). Testing hypotheses of covariance structure in multivariate data. *Electronic J. Linear Algebra 33*, 53–62.

Optimality of block designs under a hub correlation structure

Razieh Khodsiani[1] **and Saeid Pooladsaz**[2]

[1] K. N. Toosi University of Technology, Tehran, Iran
[2] Isfahan University of Technology, Isfahan, Iran

Abstract

Block experiments have been widely used in sciences, medicine and engineering. Some correlation structures can be considered in block experiments. The observations in different blocks are usually assumed to be uncorrelated but that observations within blocks are correlated with the same correlation structure.
The hub correlation is a special correlation structure with applications to experiments in genetics, networks and other areas in industry and agriculture. There is a known correlation between a hub plot (typically the first plot) and each of other plots (see [3, 5]). In general, the correlation matrix of each block for hub structure is as below,

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & \boldsymbol{\rho}'_{k-1} \\ \boldsymbol{\rho}_{k-1} & \mathbf{I}_{k-1} \end{bmatrix} \tag{1}$$

where $\boldsymbol{\rho}'_{k-1} = (\rho_2, \rho_3, \ldots, \rho_k)$, and ρ_i is the correlation between plots 1 and i, $i = 2, \ldots, k$. Note that there is no correlation between the plots $2, 3, \ldots, k$. Moreover, one can typically assume that the correlation between the first and i-th plots decays as i increases. For example, [2] considered the case that $\rho_2, \rho_3, \ldots, \rho_k$ decrease from $\rho_{\max}$ to $\rho_{\min}$.
Some researchers have obtained the optimal designs under hub correlation structure with $\rho_i = \rho$, $i = 2, \ldots, k$. For example see [1, 4].
In a block experiment with hub correlation, if one of the correlation values in ρ_{k-1}, say $\rho_{i'}$, depends on some elements such as location of i'-th plot in a block or the distance between the first and i'-th plots, then $\rho_{i'}$ may be different from ρ_i's, $i \neq i'$. So, we define $\mathbf{\Lambda}_a^{(i')}$ as the correlation matrix in (1) such that for specified plot i', $\rho_{i'} = \gamma\rho$, $0 \leq \gamma < 1$, and $\rho_i = \rho$ for all $i \neq i'$. In this presentation, by a decomposition of concurrence matrix of designs, we theoretically obtain the universally optimal block designs, binary or non-binary, under $\mathbf{\Lambda}_a^{(i')}$ for every i'.

Keywords

Universal optimality, Hub correlation structure, Generalized binary block design, Balanced block design.

References

[1] Chang, L. and Coster, D.C. (2015). Construction of weak universal optimal block designs with various correlation structures and block sizes. *Journal of Statistical Planning and Inference 160*, 1–10.

[2] Hardin, J., Garcia, S.R. and Golan, D. (2013). A method for generating realistic correlation matrices. *The Annals of Applied Statistics 7(3)*, 1733–1762.

[3] Hero, A. and Rajaratnam, B. (2012). Hub discovery in partial correlation graphs. *IEEE Transactions on Information Theory 58*, 6064–6078.

[4] Khodsiani, R. and Pooladsaz, S. (2017). Universal optimal block designs under hub correlation structure. *Statistics and Probability Letters 129*, 387–392.

[5] Zhu, Z. (2004). Optimal experimental designs with correlated observations. PhD thesis, Department of Mathematics and Statistics, Utah State University.

Testing the compound symmetry structure in large- and high-dimensional setting

Daniel Klein[1], Jolanta Pielaszkiewicz[2] and Katarzyna Filipiak[3]

[1] P. J. Šafárik University, Košice, Slovakia
[2] Linköping University, Sweden
[3] Poznań University of Technology, Poland

Abstract

In this paper the Rao score test and likelihood ratio test for testing the hypothesis related to compound symmetry structure of multivariate data covariance matrix are studied. The normal approximation of the Rao score test statistic distribution under large-dimensionality setting as well as the exact and approximate distributions of the likelihood ratio test are derived. The tests are compared via simulations. The Rao score test is available also in the case of high-dimensionality, and it is shown that the normal approximation matches well its distribution in this case. Thus, the normal approximation could be recommended for practice.

Keywords

Rao score test, Likelihood ratio test, Large dimensional asymptotics, High-dimensional asymptotics, Compound symmetry structure.

Acknowledgements

This work was supported by the Slovak Research and Development Agency under the contract no. APVV-17-0568 (Daniel Klein).

On a new approach to the ANOVA for experiments with orthogonal block structure. Experiments in row-column designs

Tadeusz Caliński, Agnieszka Łacka and Idzi Siatkowski

Poznań University of Life Sciences, Poland

Abstract

The main estimation and hypothesis testing procedures are presented for experiments conducted in row-column designs of a certain desirable type. It is shown that, under appropriate randomization, these experiments have the convenient orthogonal block structure. Due to this property, the analysis of experimental data can be performed in a comparatively simple way. Relevant simplifying procedures are indicated.

The main advantage of the presented methodology concerns the analysis of variance and related hypothesis testing procedures. Under the adopted approach one can perform these analytical methods directly, not by combining results from analyses based on some stratum submodels.

Practical application of the presented theory is illustrated by examples of real experiments in the relevant row-column designs.

Keywords

Analysis of variance, Estimation, Hypothesis testing, Orthogonal block structure, Randomization-derived model, Row-column designs.

References

[1] Caliński, T., Łacka, A., and Siatkowski, I. (2019). On a new approach to the analysis of variance for experiments with orthogonal block structure. III. Experiments in row-column designs. *Biometrical Letters 56(2)*, 183–213.

[2] Houtman, A.M. and Speed, T.P. (1983). Balance in designed experiments with orthogonal block structure. *Annals of Statistics 11*, 1069–1085.

[3] Nelder, J.A. (1965). The analysis of randomized experiments with orthogonal block structure. *Proceedings of the Royal Society of London, Series A 283*, 147–178.

Audit of neonatal mortality in Great Britain

Nicholas T. Longford

Imperial College London, UK

Abstract

The National Neonatal Audit Programme (NNAP) is an annual assessment of the performance of the neonatal units and their networks in Great Britain. It uses the data collected in the National Neonatal Research Database (NNRD). Newborns are cared for in neonatal units for three main reasons: preterm birth (poorly developed vital organs) injuries sustained during birth and congenital anomalies. Mortality rate of preterm-born babies is one of the audit items. Until recently it has been analysed by logistic regression (with random effects). From 2020 on, it is be analysed by indirect standardisation, described in [1]. In the method, the babies cared for in a network of units are matched with babies from the entire domain of the audit so that the two groups have nearly identical distributions of an extensive list of background variables. The two groups are then compared straightforwardly, without any involvement of the background variables, because their influence has been diminished by matching. Matching is accomplished by propensity analysis. The results show that the networks have similar casemixes — their matched mortality rates are in a relatively narrow range, but the observed rates are in a much wider range that could not be accounted for by sampling variation. For background and some of the outputs of NNAP, see `https://nnap.rcpch.ac.uk`. For a related application, see [2].

Keywords

Causal inference, Health-care audit, Indirect standardisation, Small-area estimation.

References

[1] Longford, N.T. (2020). Performance assessment as an application of causal inference. *Journal of the Royal Statistical Society, Series A 183*, 1363–1385.

[2] Longford, N.T. (2018). Decision theory for comparing institutions. *Statistics in Medicine 37*, 437–456.

Linearly sufficient and admissible estimators

Augustyn Markiewicz

Poznań University of Life Sciences, Poland

Abstract

The problem of admissible linear estimation under the Gauss-Markov model was extensively studied by Rao in [7]. Rao's results were developed in [3] and [2] under a possibly singular model. However, the linear sufficiency was studied and characterized in [1] and [4]. The subject of linear sufficiency and admissibility was considered in [5] and [6].
The purpose of this paper is to present estimators that satisfy both of these properties, along with some new supplementary results.

Keywords

Admissibility, Linear sufficiency.

References

[1] Baksalary, J.K. and Kala, R. (1981). Linear transformations preserving best linear unbiased estimators in a general Gauss-Markoff model. *Ann. Statist. 9*, 913–916.
[2] Baksalary, J.K., Markiewicz, A., and Rao, C.R. (1995). Admissible linear estimation in the general Gauss-Markov model with respect to an arbitrary quadratic risk function. *J. Statist. Plann. Inference 44*, 341–347.
[3] Baksalary, J.K., Rao, C.R., and Markiewicz, A. (1992). A study of the "natural restrictions" on the estimation problems in the singular Gauss-Markov model. *J. Statist. Plann. Inference 31*, 335–351.
[4] Drygas, H. (1983). Sufficiency and completeness in the general Gauss-Markov model. *Sankhya, Ser. A 45*, 88–98.
[5] Markiewicz, A. (1996). Characterization of general ridge estimators. *Stat. Probab. Lett 27*, 145–148.
[6] Markiewicz, A. and Puntanen, S. (2009). Admissibility and linear sufficiency in linear model with nuisance parameters. *Stat Pap 50*, 847–854.
[7] Rao, C.R. (1976). Estimation of parameters in a linear model. *Ann. Statist. 4*, 1023–1037.

A comparison of the estimators of covariance matrix structured by Toeplitz matrix

Adam Mieldzioc

Poznań University of Life Sciences, Poland

Abstract

The estimation of the covariance matrix plays a key role in the statistical analysis. Determination of a well-conditioned estimator of the covariance matrix with a given structure and good statistical properties can be challenging and time-consuming, especially in the high-dimensional case, where many characteristics are observed. In many cases, the standard estimators of the unstructured covariance matrix are ill-conditioned. Therefore, *shrinkage method* is used along with an appropriately selected loss function.
The purpose of the work is to extend *shrinkage method* to the structured covariance matrix estimation. We focus on a Toeplitz matrix structure. The obtained shrinkage estimators of the covariance matrix are compared with the maximum likelihood estimators in the model, where the data comes from the matrix normal distribution.

Keywords

Banded Toeplitz matrix, Shrinkage, MLE, Frobenius norm.

References

[1] Chen Y., Wiesel, A., Eldar, Y.C., and Hero, A.O. (2010). Shrinkage Algorithms for MMSE Covariance Estimation. *IEEE Transactions on Signal Processing 58*, 5016–5029.
[2] John, M. and Mieldzioc, A. (2019). The comparison of the estimators of banded Toeplitz covariance structure under the high-dimensional multivariate model. *Communications in Statistics - Simulation and Computation.*
[3] Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis 88(2)*, 365–411.
[4] Won, JH., Lim, J., Kim, SJ., and Rajaratnam B. (2013). Condition-Number-Regularized Covariance Estimation. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 75(3)*, 427–450.

A robust functional logistic regression method for classification

Müge Mutiş[1], Ufuk Beyaztaş[2], Gülhayat Gölbaşi Şimşek[1] and Han Lin Shang[3]

[1] Yıldız Technical University, Turkey
[2] Marmara University, Instanbul, Turkey
[3] Macquarie University, Sydney, Australia

Abstract

Functional logistic regression, where the response is a binary outcome and the predictor consists of random curves, has become a general framework for classification. However, parameter estimation is mainly based on a least-squares method, which can be seriously hindered by the presence of outliers. We propose a robust partial least squares method to estimate the regression coefficient function in the functional logistic regression. The regression coefficient function represented by functional partial least squares decomposition is estimated by a weighted likelihood method, which downweights the effect of outliers in the response and predictor. The finite-sample performance of the proposed method is evaluated by Monte-Carlo simulation studies and an empirical dataset

Keywords

Functional data, Robust estimation, Weighted likelihood method.

Acknowledgements

This research was supported by the Scientific and Technological Research Council of Turkey with project no. 120F270.

Variable selection in finite mixture of linear mixed models using the EM and CEM algorithms

Luísa Novais and **Susana Faria**

University of Minho, Guimarães, Portugal

Abstract

Variable selection is an important problem of any modeling study, involving the search for the simplest model that adequately describes the data, which assumes a great importance in the context of mixture models. However, the technological advances of the last decades have led to the use of data of large dimensions and of great complexity. As such, the classic variable selection methods become impracticable with the increasing size of the data, being computationally too demanding to be used in practice.

Therefore, in order to deal with the computational complexity, the need to develop new methods for variable selection has emerged in recent years. Among the new methods, methods based on penalizing functions have received great attention. These methods, unlike the classic methods, can be used in complex data problems since they allow the identification of the subset of the most relevant explanatory variables, by estimating the effect of the non-significant variables to be zero and, consequently, removing them from the model, thus drastically reducing the computational burden.

In this work we analyse the problem of variable selection in finite mixture of linear mixed models in the presence of a large number of explanatory variables. In order to do this, we compare the performance of a penalized likelihood approach for variable selection via the Expectation-Maximization (EM) and the Classification Expectation-Maximization (CEM) algorithms through a simulation study and a real data application.

Keywords

EM algorithm, CEM algorithm, Mixture models, Penalized likelihood, Variable selection.

Acknowledgements

The research of L. Novais was financed by FCT - Fundação para a Ciência e a Tecnologia, through the PhD scholarship with reference SFRH/BD/139121/2018.

References

[1] Celeux, G. and Govaert, G. (1992). A classification EM algorithm for clustering and two stochastic versions. *Computational Statistics & Data Analysis 14(3)*, 315–332.

[2] Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological) 39(1)*, 1–38.

[3] Du, Y., Khalili, A., Nešlehová, J.G., and Steele, R.J. (2013). Simultaneous fixed and random effects selection in finite mixture of linear mixed-effects models. *Canadian Journal of Statistics 41(4)*, 596–616.

[4] Khalili, A. and Chen, J. (2007). Variable selection in finite mixture of regression models. *Journal of the American Statistical Association 102(479)*, 1025–1038.

Detection of sparse and weak effects in high-dimensional data

Tatjana Pavlenko

KTH Royal Institute of Technology, Stockholm, Sweden

Abstract

We present a family of goodness-of-fit statistics based on sup-functionals of weighted empirical processes where weight functions are Erdös-Feller-Kolmogorov-Petrovski (EFKP) upper-class functions of a Brownian bridge. Chibisov-O'Reilly type of weight functions are also considered. In high dimensions, the sup-norms of weighted empirical processes standardized by the EFKP upper-class functions of a Brownian bridge have been recently seen to be good alternatives to the higher criticism statistic which is routinely applied in problems dealing with detection of sparse and weak effects. We show that the proposed family of statistics at hand that can be effectively used to solve a number of high-dimensional inferential problems.

Increasing the sample size by using rolling windows

Virgilio Pérez, Cristina Aybar and Jose M. Pavía

University of Valencia, Spain

Abstract

When studying a population, it is often not feasible to numerate it by contacting every one of its members. A sample is drawn instead, and its responses or measurements are used to predict (or infer) the behaviour of the entire population.
Many aspects of the sampling design influence how well the sample represents the population, and how valid and reliable are our inferences. The size of the sample is a key aspect of the design. In general, greater sample size is associated with more information, and therefore less uncertainty about the parameters of interest. This study focuses on how to increase the sample size, without having to obtain additional data. After all, the purpose of increasing the sample size is to reduce the estimation errors.
Validation is difficult with small samples. We address this problem by using information from members of the population that are adjacent to the subjects in the sample, that is, subjects with similar attributes. We verify this assumption on a database comprising over 150 variables recorded on 700,000 observations over a period of 30 years. We describe on it the use of rolling windows, combining AGE and TIME variables. We propose different scenarios, achieving to increase the initial sample size up to 25 times.

Keywords

Estimation error, Predictions, Rolling windows, Sample size.

Comparison of chosen covariance structure tests with regards to sensitivity to outliers and violation of normality assumption in high-dimensional regime

Jolanta Pielaszkiewicz and **Jiawei Wu**

Linköping University, Sweden

Abstract

Presentation is aimed to introduce listener to the results published in [1]. Jiawei Wu is an author of results, Jolanta Pielaszkiewicz contributed to work in role of supervisor.
We consider classical tests on structure of covariance matrices under the high-dimensional scenario in which dimensionality exceeds the sample size. We are particularly interested in identity structure, sphericity structure and diagonal structure of covariance matrix. Performance of test statistics for each structure is investigated under given assumptions and when the distributional assumption is violated. We compare test's sensitivity to outliers discussing attained significance level, power of test, and goodness of fit of considered tests. Simulation studies indicate that the test statistics testing hypothesis about identity structure of covariance matrix seems to be more sensitive to the changes of distribution assumptions and outliers compared with other considered tests. The opposite conclusion can be given for test considering the diagonal structure instead.

Keywords

Covariance matrix, Structure, Outliers, High-dimension, Normality assumption.

References

[1] Wu, J. (2020). Testing Structure of Covariance Matrix under High-dimensional Regime (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-166227.

Prediction of diabetes mellitus in bogotá using machine learning tools

Robert Sebastian Castellanos Rodriguez

National University of Colombia, Colombia

Abstract

Due to its continuously increasing occurrence, more and more families are influenced by diabetes mellitus. Most diabetics know little about their health quality or the risk factors they face prior to diagnosis. In this study, we have proposed a novel model based on data mining techniques for predicting type 2 diabetes mellitus (T2DM). The main problems that we are trying to solve are to improve the accuracy of the prediction model, and to make the model adaptive to more than one dataset. Based on a series of preprocessing procedures, the model is comprised of two parts, the improved K-means algorithm and the logistic regression algorithm. The colombian Diabetes Dataset and the Waikato Environment for Knowledge Analysis toolkit were utilized to compare our results with the results from other researchers [1].

Keywords

K-means, Health prediction, Machine learning.

References

[1] Catellanos, R.S. (2021). Using machine learning for pedictions of disease.

Small area estimation using reduced rank regression

Dietrich von Rosen

Swedish University of Agricultural Sciences, Uppsala, Sweden

Abstract

Small area estimation techniques have got a lot of attention during the last decades due to their important applications in survey studies. Mixed linear models and reduced rank regression analysis are jointly used when considering small area estimation. Estimates of parameters are presented.

Keywords

Small area estimation, Reduced rank regression.

Conditions for mixture representation of system lifetime distribution

Jorge Navarro[1], Tomasz Rychlik[2] and Fabio Spizzichino[3]

[1] University of Murcia, Spain
[2] Polish Academy of Sciences, Poland
[3] University La Sapienza, Italy

Abstract

Under the assumption that the component lifetimes of a reliability system are independent and identically continuously distributed, [4] represented the system lifetime distribution as a mixture of distributions of order statistics of component lifetimes with the vector of mixture representation, called the Samaniego signature, depending merely on the system structure. [3] extended the Samaniego formula to the case of exchangeable component lifetimes. [1] presented necessary and sufficient conditions assuring the Samaniego representation which were expressed in terms of distributional properties of families of auxiliary indicator random vectors parametrized by positive numbers. In [2] we obtained other necessary and sufficient conditions with a more natural probabilistic interpretation. They are represented in terms of the marginal distributions of component lifetimes and the dependence copula of them.

Keywords

Coherent system, Samaniego signature, Exchangeable distribution, Diagonally dependent copula.

References

[1] Marichal, J.-L., Mathonet, P., and Waldhauser, T. (2011). On signature-based expressions of system reliability. *J. Multivariate Anal. 102*, 1410–1416.
[2] Navarro, J., Rychlik, T., and Spizzichino, F. (2021). Conditions on marginals and copula of component lifetimes for signature representation of system lifetime. *Fuzzy Sets and Systems 415*, 99–117.
[3] Navarro, J., Samaniego, F.J., Balakrishnan, N., and Bhattacharya, D. (2008). Applications and extensions of system signatures in engineering reliability. *Naval Res. Logistics 55*, 313–327.
[4] Samaniego, F.J. (1985). On the IFR closure theorem. *IEEE Trans. Reliab. 34*, 69–72.

On partial least squares estimation in functional regression models

Semanur Sarıçam[1], **Ufuk Beyaztaş**[2], **Barış Aşıkgil**[1] **and Han Lin Shang**[3]

[1] Mimar Sinan Fine Arts University, Instanbul, Turkey
[2] Marmara University, Istanbul, Turkey
[3] Macquarie University, Sydney, Australia

Abstract

Scalar-on-function regression, where the response is scalar-valued and predictor consists of random functions, is one of the most important tools for exploring the relationship between the response and predictor. For estimating the regression coefficient function, the functional partial least squares method produces improved estimation accuracy compared to other existing methods, such as least squares, maximum likelihood, and maximum penalized likelihood. The functional partial least squares method is often based on either SIMPLS or NIPALS algorithm. These algorithms can be computationally slow for analyzing a large dataset. By re-orthogonalizing score and loading vectors, we introduce a modified functional partial least squares method to estimate the regression coefficient function under the scalar-on-function framework efficiently. The finite-sample performance and computational speed are evaluated using a series of Monte Carlo simulation studies and a chemometric dataset.

Keywords

Basis expansion, Orthogonalization, Scalar-on-function regression.

Acknowledgements

This research was supported by the Scientific and Technological Research Council of Turkey with project no. 120F270.

Controlling the bias for M-quantile estimators for small area

Francesco Schirripa Spagnolo[1], Gaia Bertarelli[2], Raymond Chambers[3], David Haziza[4] and Nicola Salvati[1]

[1] University of Pisa, Italy
[2] Sant'Anna School of Advanced Studies, Pisa, Italy
[3] University of Wollongong, Australia
[4] University of Ottawa, Canada

Abstract

Representative outlier units occur frequently in surveys. As a result, several methods have been proposed to mitigate the effects of them in survey estimates. If outliers are a concern for estimation of population quantities, it is even more necessary to pay attention to them in a small area estimation (SAE) context, where sample size is usually very small and the estimation in often model-based. [3] explicitly addressed this issue of outlier robustness in SAE, using an approach based on fitting outlier robust M-quantile models to the survey data. More recently, [7] also addressed this issue from the perspective of linear mixed models. Both these approaches, however, use plug-in robust prediction, i.e. they replace parameter estimates in optimal, but outlier-sensitive, predictors by outlier robust versions (a robust-projective approach). Unfortunately, these predictors are efficient under correct model specification and assumptions, but they may be sensitive to the presence of outliers because they use plug-in robust prediction which usually leads to a low prediction variance and a considerable prediction bias. [4, 5] proposed bias corrected method to reduce the prediction bias when the response variable is continuous. In this talk, we focus on M-quantile approach and we propose two general methods (i.e., for continuous and discrete data) to reduce the prediction bias of the robust M-quantile predictors in SAE context. The first estimator is based on the concept of conditional bias and extends the results of [1, 6]. Then, we propose an unified approach to M-quantile predictors for continuous and discrete data which is based on a full bias correction and it could be viewed as a generalization of [2] approach. A Monte-Carlo simulation study is conducted and its results suggest that our approaches mainly improve the efficiency and they control the bias prediction error of M-quantile predictors when the population contains units that may be influential if selected in the sample.

Keywords

Small area estimation, Robust estimators, Bias correction.

References

[1] Beaumont, J.-F., Haziza, D., and Ruiz-Gazen, A. (2013). A unified approach to robust estimation infinite population sampling. *Biometrika 100*, 555–569.

[2] Chambers, R.L. (1986). Outlier robust finite population estimation. *J. Am. Statist. Ass. 81*, 1063–1069.

[3] Chambers, R. and Tzavidis, N. (2006). M-quantile models for small area estimation. *Biometrika 93(2)*, 255–268.
[4] Chambers, R., Chandra, H., Salvati, N., and Tzavidis, N. (2014). Outlier robust small area estimation. *J. Roy. Stat. Soc. B 76(1)*, 47–69.
[5] Dongmo Jiongo, V., Haziza, D., and Duchesne, P. (2013). Controlling the bias of robust small area estimators. *Biometrika 100*, 843–858.
[6] Favre-Martinoz, C. (2015). Estimation robuste en population finie et infinite. PhD Thesis.
[7] Sinha, S.K. and Rao, J.N.K. (2009). Robust small area estimation. *Canadian Journal of Statistics 37(3)*, 381–399.

Design selection for 2-level supersaturated designs

Rakhi Singh

UNC Greensboro, USA

Abstract

The commonly used design optimality criteria are inadequate for selecting supersaturated designs. As a result, there is extensive literature on alternative optimality criteria within this context. Most of these criteria are rather ad hoc and are not directly related to the primary goal of experiments that use supersaturated designs, which is factor screening. Especially, unlike almost any other optimal design problem, the criteria are not directly related to the method of analysis.

An assumption needed for the analysis of supersaturated designs is the assumption of effect sparsity. Under this assumption, a popular method of analysis for 2-level supersaturated designs is the Gauss-Dantzig Selector (GDS), which shrinks many estimates to 0. We develop new design selection criteria inspired by the GDS and establish that designs that are better under these criteria tend to perform better as screening designs than designs obtained using existing criteria.

This presentation is based on joint work with John Stufken, University of North Carolina at Greensboro.

Likelihood based classification of growth curves

Dietrich von Rosen[1,2] and Martin Singull[2]

[1] Swedish University of Agricultural Sciences, Uppsala, Sweden
[2] Linköping University, Sweden

Abstract

This presentation considers the discrimination between two populations of growth curves. We establish a classification procedure that is likelihood based, in that sense that we compare the two likelihoods given that the new observation belongs to respectively population. We also discuss the possibility that we classify the new observation to an unknown population, which we show is natural when considering growth curves.

Keywords

Discriminant analysis, Growth Curve model, Likelihood based.

Projection tests for linear hypothesis in the functional response model

Łukasz Smaga

Adam Mickiewicz University in Poznań, Poland

Abstract

The paper concerns the linear hypothesis testing problem in the functional response model, which is one of the regression models considered in the functional data analysis. In this model, the response is a function represented as a random process, while the predictors are random variables. To test the linear hypothesis, the projection tests are constructed and theoretically justified. Namely, a kind of equivalence between the functional null hypothesis and its projected version is established. Different Gaussian processes and numbers of projections are considered in the implementation of new solutions. Moreover, as there is no one test having the best power for all correlation cases, a simple combining test is also proposed. It has satisfactory power in all cases. In simulation studies, the new tests are compared with existing methods in terms of size control and power. The real data example is also provided to illustrate the results.

Keywords

Functional regression model, Gaussian process, Linear hypothesis, Projection method, Statistical test.

References

[1] Chiou, J.M., Müller, H.G., and Wang, J.L. (2004). Functional response models. *Statist. Sinica 14*, 675–693.

[2] Cuesta-Albertos, J.A. and Febrero-Bande, M. (2010). A simple multiway ANOVA for functional data. *Test 19*, 537–557.

[3] Cuesta-Albertos, J.A., García-Portugués, E., Febrero-Bande, M., and González-Manteiga, W. (2019). Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes. *Ann. Statist. 47*, 439–467.

[4] Smaga, Ł. (2019). General linear hypothesis testing in functional response model. *Comm. Statist. Theory Methods*. doi:10.1080/03610926.2019.1691233

[5] Zhang, J.T. (2011). Statistical inferences for linear models with functional responses. *Statist. Sinica 21*, 1431–1451.

A second order 2-dimensional intrinsic Gaussian Markov random fields in blood pressure data

Maria Zafeiria Spyropoulou and James Bentham

University of Kent, UK

Abstract

Raised blood pressure is a key risk factor for non-communicable diseases, and is estimated to affect 1.13 billion people worldwide. The focus is on estimating trends in systolic (SBP) and diastolic (DBP) blood pressure but with an emphasis on their interaction (INT).

We separate the globe into groups of countries, and apply a Bayesian hierarchical model. Each country is a member of a region and super-region, in which there is a smaller and a larger group of countries respectively. This structure allows us to borrow strength across regions and super-regions when there is not adequate information from the countries' own data. Within each country, data are correlated temporally and within each region and super-region data have temporal and between-countries correlation [2].

We begin with a linear model over time. Three separate linear models for DBP, SBP and INT are created, with each taking advantage of the hierarchical structure and borrowing information. To allow the temporal correlation we add a non-linear model over time. A two-dimensional Intrinsic Gaussian Markov Random Field (IGMRF) is implemented for the model. By using the IGMRF of two dimensions for DBP and SBP, the INT is included implicitly [1]. At last, we introduce an age model using B-splines, and fixed effects for covariates such as diet types, urbanization and studies' coverage, which help the model fitting for DBP, SBP and INT.

Applying these methods has shown that IGMRF approximations can be used for estimating the correlations and for facilitating MCMC algorithms. For the computational process of the model, we use canonical parametrisation for the Metropolis' updates, Cholesky factorisation for the Gibbs' sampler updates and parallel computing for accelerating the code. By using the aforementioned methodology, our model is mixing well and converges, hence is able to estimate the variation of the interaction between SBP and DBP across countries, over time in different age groups and gender.

Keywords

MCMC, Intrinsic Gaussian Markov random fields, 2^{nd} order 2-dimensional random walk

Acknowledgements

The author was supported by the University of Kent.

References

[1] Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications. Vol. 104. Monographs on Statistics and Applied Probability.* London: Chapman & Hall.

[2] Danaei, G., Finucane, M.M., Lin JK, et al. (2011). *National, regional, and global trends in systolic blood pressure since 1980:systematic analysis of health examination surveys and epidemiological studies with786 country-years and 5Âů4 million participants.* Lancet.

Likelihood ratio test for the model with block covariance structure

Malwina Janiszewska[1], Anna Szczepańska-Álvarez[1] and Adolfo Álvarez[2]

[1] Poznań University of Life Sciences, Poland
[2] O-I Business Service Center, Poznań, Poland

Abstract

In this talk we present the study about the association between two sets of characteristics. We assume that the observations have the multivariate normal distribution with the block-structure of the covariance matrix. We verify the separability of top left block of the covariance matrix and test independence of two groups of characteristics. Moreover, we consider how the distribution of test statistic depends on the sample size and the number of analyzed characteristics. We perform the simulations studies.

Keywords

Block covariance structure, Separable covariance structure, Maximum likelihood estimation, Likelihood ratio test.

Signature conditions for distributional properties of system lifetimes if component lifetimes are iid (exponential)

Tomasz Rychlik[1] and Magdalena Szymkowiak[2]

[1] Polish Academy of Sciences, Poland
[2] Poznań University of Technology, Poland

Abstract

First we consider coherent and mixed systems built of components with independent identically exponentially distributed lifetimes. We establish various sufficient conditions on the system signatures for asserting strong unimodality of the system lifetime as well as the monotonicity of its failure rate and mean residual life. Later on we show that the assumption of exponentiality of the component lifetimes by the increacing (decresing) failure rate condition can be relaxed.

Keywords

Coherent and mixed system, Signature, Strong unimodality, Failure rate, Mean residual life, Convex transform order.

Acknowledgements
The second author was partially supported by PUT under grant 0211/SBAD/0121.

References

[1] Arnold, B.C., Rychlik, T., and Szymkowiak, M. (2021). Preservation of distributional properties of component lifetimes by system lifetimes. Submitted.
[2] Rychlik, T. and Szymkowiak, M. (2021a). Properties of system lifetime in the classical model with i.i.d. exponential component lifetimes. In: *Advances in Statistics – Theory and Applications. Honoring the Contributions of Barry C. Arnold in Statistical Science* (I. Ghosh, N. Balakrishnan, H.K.T. Ng, eds.). Springer, Cham, Switzerland, 4–66.
[3] Rychlik, T. and Szymkowiak, M. (2021b). Signature conditions for distributional properties of system lifetimes if component lifetimes are iid exponential. Submitted.
[4] Samaniego, F.J. (1985). On closure of the IFR class under formation of coherent systems. *IEEE Trans. Reliab. 34*, 69–72.

Supervised feature selection with structure learning for sparse and weak data in high dimensional setting

Annika Tillander[1] and Tatjana Pavlenko[2]

[1] Linköpings University, Sweden
[2] KTH Royal Institute of Technology, Stockholm, Sweden

Abstract

In a high-dimensional setting where the informative features are few and the information as such is weak, it is difficult to successfully differentiate between classes. By identifying the relation between features the information can be enhanced. A known dependence structure between features enables block-diagonal approximation of the inverse covariance matrix, then block-wise information strength can be estimated and selection of blocks of features instead of single features. Using only informative blocks in an additive classifier has shown improved classification accuracy [1]. Hence the need of learning the structure of the covariance matrix. A novel block-identifying procedure is suggested and applied to simulated data as well as real data.

Keywords

Feature selection, High dimensionality, Supervised classification, Structure learning.

References

[1] Pavlenko, T., Björkström, A. and Tillander, A. (2012). Covariance structure approximation via gLasso in high-dimensional supervised classification. *Journal of applied statistics 39*, 1643–1666.

On the ratio of extremal eigenvalues of β-Laguerre ensembles

Denise Uwamariya and Xiangfeng Yang

Linköping University, Sweden

Abstract

Classical β-Laguerre ensembles consist of three special matrix models taking the form $\mathbf{X}\mathbf{X}^T$ with $\mathbf{X}$ denoting a random matrix having i.i.d. entries being real ($\beta = 1$), complex ($\beta = 2$) or quaternion ($\beta = 4$) normal distribution. It had been actually believed that no other choice of $\beta > 0$ (besides $1, 2$ and 4) would correspond to a matrix model $\mathbf{X}_\beta \mathbf{X}_\beta^T$ which can be constructed with entries from a classical distribution until the work [5]. Since then the spectral properties of general β-Laguerre ensembles have been extensively studied dealing with both the bulk case (involving all the eigenvalues) and the extremal case (addressing the (first few) largest and smallest eigenvalues). However, the ratio of the extremal eigenvalues (equivalently the condition number of $\mathbf{X}_\beta$) has not been well explored in the literature. In this paper, we study such ratio in terms of large deviations.

Keywords

β-Laguerre ensembles, Extremal eigenvalues, Large deviations.

References

[1] Anderson, G., Guionnet, A., and Zeitouni, O. (2010). *An introduction to random matrices.* Cambridge University Press, Cambridge.

[2] Baker, T., Forrester, P., and Pearce, P. (1998). Random matrix ensembles with an effective extensive external charge. *J. Phys. A 31*, 6087–6101.

[3] Dembo, A. and Zeitouni, O. (2010). *Large deviations techniques and applications.* Corrected reprint of the second (1998) edition. Springer-Verlag, Berlin.

[4] Dumitriu, I. (2003). *Eigenvalue statistics for beta-ensembles.* Ph.D. Thesis, Massachusetts Institute of Technology.

[5] Dumitriu, I. and Edelman, A. (2002). Matrix models for beta ensembles. *J. Math. Phys. 43*, 5830–5847.

[6] Gupta, A. and Nagar, D. (1999). *Matrix variate distributions.* Chapman & Hall/CRC, Boca Raton, FL.

[7] James, A. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist. 35*, 475–501.

[8] Jiang, T. and Li, D. (2015). Approximation of rectangular beta-Laguerre ensembles and large deviations. *J. Theoret. Probab. 28*, 804–847.

[9] Johansson, K. (2000). Shape fluctuations and random matrices. *Comm. Math. Phys. 209*, 437–476.

[10] Johnstone, I. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist. 29*, 295–327.

[11] Ledoux, M. and Rider, B. (2010). Small deviations for beta ensembles. *Electron. J. Probab. 15*, 1319–1343.

[12] Petz, D. and Hiai, F. (1997). *Logarithmic energy as an entropy functional.* Advances in differential equations and mathematical physics (Atlanta, GA, 1997), 205–221, Contemp. Math., 217, Amer. Math. Soc.

[13] Rudelson, M. (2008). Invertibility of random matrices: norm of the inverse. *Ann. of Math. 168*, 575–600.

[14] Ramírez, J., Rider, B., and Virág, B. (2011). Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc. 24*, 919–944.

[15] Trefethen, L. and Bau, D. (1997). *Numerical linear algebra.* SIAM, Philadelphia, PA.

SARS-CoV-2 variants neutralisation studies to predict vaccines effectiveness

Oleg Volkov

Stanford University, USA

Abstract

Response to emerging SARS-CoV-2 variants could be amplified by correlates of immune protection. They would allow predicting vaccine effectiveness from in-vitro data, without months-long clinical trials and observational studies. Despite intensive research and an accute need, few correlates frameworks have been established for SARS-CoV-2 vaccines.
This talk introduces such a framework which utilises neutralisation studies with SARS-CoV-2 variants. We consider how it could boost esablished vaccines and accelarate new vaccine development. We also consider study designs for improved predictions and vaccination policies.

Computational methods for probability distributions

Viktor Witkovský

Slovak Academy of Sciences, Bratislava, Slovakia

Abstract

The contribution deals with numerical methods for calculating the exact probability distributions using the numerical inversion of the characteristic function. We shall present the general concept of *The Characteristics Functions Toolbox* (CharFunTool) [1] – the MATLAB repository of characteristic functions and tools for their combinations and numerical inversion. We shall motivate its applicability by computing the exact probability distribution of estimators and test statistics in linear models with constraints (we consider known covariance structure) and in specific multivariate models.
Moreover, here we shall illustrate that the method is suitable for fast calculation of the exact bootstrap distribution of the sample mean (as well as other linear functions, e.g., the sample moments) on the lattice distributions (i.e. such discrete distributions that every possible value can be represented in the form $a + bn$, where $b \neq 0$ and n is an integer.

Keywords

Characteristic function approach, Linear regression model with constraints, Best linear unbiased estimator, Exact bootstrap distribution.

Acknowledgements

The work was supported by the Slovak Research and Development Agency, project APVV-18-0066, and by the projects VEGA 2/0081/19 and VEGA 2/0096/21.

References

[1] Witkovský V. (2021). *CharFunTool: The Characteristic Functions Toolbox (MATLAB).* `https://github.com/witkovsky/CharFunTool`.

A computational method with density based clustering approach for the data in the existence of outlier

Fatma Yerlikaya-Özkurt

Atılım University, Ankara, Turkey

Abstract

Recently, collection of huge amount of data and analysis of that much data have vital importance for human activities in many different application areas. Advanced statistical methods play crucial role for modeling of such data when the data contains outliers. Although there are number of outlier detection methods for revealing outlier observations in data, most of them may not be reasonable and appropriate for prediction purposes due to structural and requirements of modeling. In this study, density based clustering algorithms are considered in order to detect the location of outlier observations effectively with respect to form of the model in high-dimensional data. Based on obtained results, the Mean Shift Outlier Model (MSOM) is constructed as a robust linear model. This newly proposed computational method use power of data clustering and also minimize the impact of the outlier observations. The numerical example is also presented to reveal the performance of the method in this study.

Keywords

Linear models, Outlier observation, High-dimensional data, Density based clustering.

References

[1] Gan, G., Ma, C., and Wu, J. (2020). *Data clustering: theory, algorithms, and applications.* SIAM.
[2] Xia, J., Gao, L., Kong, K., Zhao, Y., Chen, Y., Kui, X., and Liang, Y. (2018). Exploring linear projections for revealing clusters, outliers, and trends in subsets of multi-dimensional datasets. *Journal of Visual Languages and Computing 48* 52–60.
[3] Xu, R. and Wunsch, D. (2008). *Clustering.* John Wiley & Sons.
[4] Xu, X., Liu, H., Li, L., and Yao, M. (2018). A comparison of outlier detection techniques for high-dimensional data. *International Journal of Computational Intelligence Systems 11(1)*, 652–662.

Feature selection for mean shift outlier model via the conic-fused Lasso

Pakize Taylan[1] **and Fatma Yerlikaya-Özkurt**[2]

[1] Dicle University, Diyarbakır, Turkey
[2] Atılım University, Ankara, Turkey

Abstract

Outlier detection and variable selection are among main objectives of statistical analysis. In our study, we address the outlier problem for classification by using the Mean Shift Outlier Model (MSOM). Since the MSOM has more coefficients than the linear regression model, the complexity of the MSOM is high. Therefore, we consider future selection for MSOM by using fused Lasso which is useful when the number of features is much larger than the sample size. Fused Lasso is penalizing the L1- norm of both the coefficients and their successive differences and it allows sparsity for both of them while Lasso only allows the coefficients. In this study, we take into account Iterated Ridge approximation for fused Lasso problem and it is solved by using continuous optimization technique which is permitting the use of interior point methods. The newly developed method is called C-Fused Lasso (C-FL) and is applied to real world data set to show the performance of C-FL.

Keywords

Outlier, Lasso, Mean shift outlier model, Classification problem, Convex optimization.

References

[1] Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications*, MPS-SIAM Series on Optimization.
[2] Renegar, J.A. (2001). *Mathematical View of Interior-Point Methods in Convex Programming*, SIAM, Philadelphia.
[3] Rao, C.R., Toutenburg, H., and Fieger, A. (1999). *Linear models: Least squares and alternatives*, Springer.
[4] Tibshirani, R., Saunders, M., Rosset, S., and Zhu, J .(2005). Sparsity and smoothness via the fused lasso. *J. R. Statist. Soc. B 67*, 91–108.

Intuitionistic fuzzy time series: A cascade prediction model

Özge Cağcağ Yolcu and Ufuk Yolcu

Marmara University, Istanbul, Turkey

Abstract

Time series prediction models have both theoretical and practical significance in a wide research area such as finance, health care, environment, and energy. The main purpose of time series analysis is to get an accurate picture of the future. Obtaining satisfactory prediction results is only possible with a suitable and competent prediction tool. Prediction tools introduced in the literature can be evaluated under two subtitles as probabilistic and non-probabilistic tools. While the traditional statistical methods constitute probabilistic methods, fuzzy-based systems, and computational-based systems including machine learning, deep learning, etc. methods constitute the non-probabilistic methods. Fuzzy-based prediction models are effectively used as a non-probabilistic prediction tool when the data sets are vague and linguistic terms. Fuzzy-based prediction models firstly proposed by [2], as based on Zadeh's fuzzy set theory [3], do not take into consideration both the neutrality degree and the non-membership degrees. In order to overcome this issue, intuitionistic fuzzy time series prediction models (IFTS-PM), based on Atanassov's intuitionistic fuzzy sets [1], have been put forward. Although current IFTS-FMs use non-membership values as inputs in the prediction process as well as memberships, they only model linear or non-linear relationships between inputs and outputs. However, simultaneous modeling of linear and non-linear relationships between inputs and outputs, which contain valuable information for time-series prediction, will improve prediction accuracy. In this study, we aim to consider both linear and non-linear intuitionistic fuzzy relationships between inputs and outputs together in an IFTS-PM. For this purpose, a cascade forward neural network is used to get these relations and the model is called as intuitionistic fuzzy time series-cascade prediction model (IFTS-CPM). Cascade forward neural network is also capable of accommodating the nonlinear relationship between input and output by not eliminating the linear relationship between the two. By using the intuitionistic fuzzy C-means clustering algorithm, the membership and non-membership values are obtained for each time-series observation. And these membership and non-membership values are used as inputs of the cascade neural network besides the lagged observations of real time-series. The target values are also composed of real observation of time-series at t time. The performance of IFTS-CPM has been discussed on test sets of some real-world time-series, comparatively via the error criteria, in addition, the convergence time has been examined, and also the fitting of predicts and observations has been presented with different demonstrations.

Keywords

Intuitionistic fuzzy time series, Cascade forward neural network, Time series prediction, Intuitionistic fuzzy relation.

References

[1] Atanassov, K.T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets Systems 20(1)*, 87–96.

[2] Song, Q. and Chissom, B.S. (1993). Fuzzy time series and its models. *Fuzzy Sets Systems 54*, 269–277.
[3] Zadeh, L.A. (1965). Fuzzy sets. *Information and Control 8(3)*, 338–353.

Testing the mean in three-level data with doubly-exchangeable covariance structure

Ivan Žežula[1], Daniel Klein[1] and Anuradha Roy[2]

[1] P. J. Šafárik University, Košice, Slovakia
[2] University of Texas at San Antonio, Texas, USA

Abstract

We consider matrix-valued multivariate observation model with three-level doubly-exchangeable covariance structure. Using known optimal estimators of unknown parameters under multivariate normality, we propose two different methods of testing the mean value. Both one-sample and two-sample tests are constructed, and exact distributions of the test statistics are derived. Possible methods of computing p-values and critical values of the distributions are compared using two real data examples.

Keywords

Multivariate observations, Three-level data, Special variance structures, Mean testing.

Acknowledgements

Žežula's and Klein's research was supported by the Slovak Research and Development Agency under the Contract No. APVV-17-0568.

References

[1] Žežula, I., Klein, D. and Roy, A. (2020). Mean Value Test for Three-Level Multivariate Observations with Doubly Exchangeable Covariance Structure. In: Holgersson, T., Singull, M. (Eds.), *Recent Developments in Multivariate and Random Matrix Analysis - Festschrift in Honour of Dietrich von Rosen* (pp. 335–349). Springer.

Jordan algebra in estimation and testing hypotheses in multivariate normal models

Roman Zmyślony and **Arkadiusz Kozioł**

University of Zielona Góra, Poland

Abstract

In this presentation estimation of fixed and covariance parameters will be considered in multivariate normal models with special covariance structure i.e. block compound symmetric (BCS). The properties of the estimators will be analyzed. Under multivariate normality, the free-coordinate approach is used to obtain unbiased linear and quadratic estimates for the model parameters. Optimality of these estimators follows from sufficiency and completeness of their distributions. As unbiased estimators with minimal variance, they are consistent.
Also in presentation will be given new approach for testing hypotheses on the structure of covariance matrices in double multivariate data with block compound symmetric covariance structure. It is proved that ratio of positive and negative parts of best quadratic unbiased estimators (BQUE) provide an F-test for independence of blocks variables in double multivariate models. Simulation studies for comparison of powers between F-test and LRT will be presented.

Keywords

Best unbiased estimator, Block compound symmetric covariance structure, Double multivariate data, Positive and negative part of estimator, Structure of covariance matrices, Structure of mean vector, Testing hypotheses.

References

[1] Drygas, H. (1970). *The Coordinate-Free Approach to Gauss-Markov Estimation.* Berlin, Heidelberg: Springer.
[2] Jordan, P., von Neumann, J., and Wigner, E. (1934). On an algebraic generalization of the quantum mechanical formalism. *The Annals of Mathematics 35(1)*, 29–64.
[3] Johnson, R.A. and Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis.* Sixth ed., Pearson Prentice Hall, Englewood Cliffs, NJ.
[4] Kozioł, A., Roy, A., Fonseca, M., Zmyślony, R., and Leiva, R. (2018). Free-coordinate estimation for doubly multivariate data. *Linear Algebra Appl.*, doi:10.1016/j.laa.2018.02.019.
[5] Michalski, A. and Zmyślony, R. (1999). Testing hypotheses for linear functions of parameters in mixed linear models. *Tatra Mountains Mathematical Publications 17*, 103–110.
[6] Seely, J.F. (1971). Quadratic subspaces and completeness. *The Annals of Mathematical Statistics 42(2)*, 710–721.
[7] Zmyślony, R. (1978). A characterization of best linear unbiased estimators in the general linear model. *Lecture Notes in Statistics 2*, 365–373.

Part VI

Posters

Fitting mixtures of linear mixed models: a case study involving fishery data

Susana Faria and Luísa Novais

University of Minho, Portugal

Abstract

Finite mixture regression models have been extensively used for modelling regression relationships arising from a clustered and heterogeneous population. Within the family of mixtures of regression models, mixtures of linear mixed models also allow to take into account the correlation between observations from the same individual.

In this study, we discuss briefly the procedure for fitting mixtures of linear mixed models to fishery data, by means of maximum likelihood. We pretend to study the relationship between the price per unit of mackerel caught in the coast of Portugal and other covariates, like as, region of fishing, body size of mackerel and year of fishing.

The number of components in the mixture model was deemed to be unknown and estimated from the data.

A mixture of linear mixed models with two components was fitted to the data. The first component contains the highest number of vessels and concerns vessels with a smaller quantity caught, whose price per unit of mackerel is low and which landed almost entirely in North and Center regions of Portugal. A second component has a slightly lower number of vessels and contains vessels that have caught more mackerel whose price per unit of mackerel is high. These vessels have landed mostly in the South region of Portugal.

Keywords

Mixtures of linear mixed models, EM algorithm, Fishery data.

Acknowledgements

The research of L. Novais was financed by FCT - Fundação para a Ciência e a Tecnologia, through the PhD scholarship with reference SFRH/BD/139121/2018.

References

[1] Celeux, G., Martin, O., and Lavergne, C. (2005). Mixture of linear mixed models for clustering gene expression profiles from repeated microarray experiments. *Stat Model 5(3)*, 243–267.

[2] Novais, L. and Faria, S. (2021). Selection of the number of components for finite mixtures of linear mixed models. *Journal of Interdisciplinary Mathematics*. doi:10.1080/09720502.2021.1889786.

[3] Ogle, D.H. (2016). *Introductory Fisheries Analyses with R*. CRC Press, UK.

D-optimal designs with correlated errors

Małgorzata Graczyk and Bronisław Ceranka

Poznań University of Life Sciences, Poland

Abstract

This paper deals with problematic issues related to the planning of experiments. In the theory of experiments, considering the planning of experiments and, afterwards, the analysis of results and drawing of conclusions, a significant role is played by the way in which experiments are planned. This research helps to determine unknown measurements of p objects in n measurements according to the plan of the spring balance weighing design. Such experimental plans and their modifications are widely applied in different branches of knowledge, including agriculture, optics and physics. Here, issues related to the regular D-optimal spring balance weighing design are considered. These designs are considered under the assumption that the measurement errors are correlated. The upper bound of the determinant of the information matrix of estimators is given, and the conditions for this upper bound to be attained are proved. Moreover, conditions for the existence of a regular D-optimal design are presented, as well as new construction methods, based on the set of incidence matrices of some known block designs.

Keywords

Block design, D-optimality, Spring balance weighing design.

References

[1] Hudelson, M.G., Watkins, W., and Zeitlin, J. (1998). Notes on D-optimal designs. *Linear Algebra and its Applications 24*, 519–598.

[2] Neubauer, M., Klee, V., and Larman, D. (2000). Largest j-simplices in d-cubes: Some relatives to the Hadamard determinant problem. *Algebra and its Applications 30*, 109–127.

The influence of the variance-covariance structure on the GWAS results

Monika Mokrzycka and Paweł Krajewski

Polish Academy of Sciences, Poznań,Poland

Abstract

Genome-Wide Association Study (GWAS) is an analysis, which aims at finding relationships between genetic and phenotypic properties of plants. This study was completed with 509 wheat varieties and breeding lines observed during 2 vegetation seasons: 2017/18 and 2018/19 at four experimental stations belonging to two breeding companies. Genetic structure of the considered set of wheat accessions was analyzed by [1] and phenotypic data analyses are under preparation.
We perform a mixed model marker-trait association analysis in the context of multiple environments. This statistical analysis allows to specify four variance-covariance models for the set of environments as identity, diagonal, compound symmetry (cs) or uniform covariances with unequal variances (hcs). The problem is about the influence of the specified variance-covariance structure on significance of fixed marker effects. Smaller differences are observed between pairs diagonal–identity and cs–hcs. All analyses were completed with Genstat 19 software [2].

Keywords

Variance-covariance structure, Genome-Wide Association Study.

Acknowledgements

This work was supported by the National Centre for Research and Development (www.ncbr.gov.pl) as a part of the BIOSTRATEG program, project HYBRE no. BIOSTRATEG3/343665/6/NCBR/2017.

References

[1] Tyrka, M., Mokrzycka, M., Bakera, B., Tyrka, D., Szeliga, M., Stojałowski, S., Matysik, P., Rokicki, M., Rakoczy-Trojanowska, M., and Krajewski, P. (2021). Evaluation of genetic structure in European wheat cultivars and advanced breeding lines using high-density genotyping-by sequencing approach. *BMC Genomics 22*, 81.

[2] VSN International. Genstat for Windows. 19th ed. Hemel Hempstead: VSN International; 2017. Web page: Genstat.co.uk.

Part VII

Happy Birthday Augustyn!

My private impressions for the 65th anniversary of professor Augustyn Markiewicz

Jan Hauke

Adam Mickiewicz University in Poznań, Poland

Augustyn started working at the Agricultural University in Poznań (now the University of Life Sciences) in 1979. At the Department of Mathematical and Statistical Methods, where we worked together, the relationships were friendly, so my contact with Augustyn was quite close from the very beginning (as with other colleagues). However, it was only after I completed my Ph.D. in 1984 that our cooperation became closer in various fields: together we participated in English courses at Adam Mickiewicz University (AMU), we conducted classes in the same fields, and perhaps most importantly, we had a mutual scientific mentor, professor Jerzy K. Baksalary. Frequent and long discussions provoked by Jerzy (not only scientific) were a cornerstone for our further cooperation. We also met on a social basis (Figure 1 is a photo from a joint event in my apartment in mid-eighties).

Figure 1. Joint event in Jan Hauke's apartment in mid-eighties.

Shortly after Augustyn completed his Ph.D. (in 1988), our mentor left for a one-year research stay in Tampere. Augustyn continued the research into the problems he started in collaboration with J. K. Baksalary, including with professor C.R. Rao (who received an honorary doctorate at AMU in 1991 - a photo of Augustyn's participation in this ceremony on Figure 2).

Figure 2. Augustyn participating in the ceremony of C.R. Rao honorary doctorate at Adam Mickiewicz University.

Then Augustyn and me (on his initiative) began to cooperate more closely scientifically. What made it easier was the fact that we were transferred to a joint office room, where we worked together until 1994, when I moved to AMU. Despite this change, our scientific cooperation continues, resulting in joint publications. Together, we carry out a grant on matrix ordering, we regularly meet at a seminar on algebraic-statistical problems (PLAG), participate in conferences (a photo from the IWMS conference in Shrewsburry in 1996 on Figure 3).
Our intensive cooperation lasted over 17 years (which is over 25% of Augustyn's current jubilee). Its effect has also been a jointly elaborated idea of the MATTRIAD conference, which was held for the first time in 2005 in Będlewo. Augustyn, regardless of the algebraic and statistical topics constituting the basis for our cooperation, also dealt with theoretical statistical research, which was later an important part of his post-doctoral thesis. After obtaining the post-doctoral degree (in

Figure 3. IWMS conference in Shrewsburry in 1996.

Polish "doktor habilitowany") in 1999, Augustyn began to form his own group of students who under his supervision have been preparing Ph.D. dissertations. And they are now the core of the PLAG seminar group. Augustyn's intense activity means that these meetings are still taking place and he gives them a special tone, inspiring discussions, commenting on presentations and encouraging creative thinking. I am very pleased to be part of this experience, and I hope it will continue for years to come. Augustyn, **the jubilee of 65 is only a certain stage** for you. You have accomplished a lot, but I know that thanks to your youthful scientific energy and research fervour, you will achieve much more.

My adventures with A.M.

Simo Puntanen

Faculty of Information Technology and Communication Sciences
FI-33014 Tampere University, Finland
email: simo.puntanen@tuni.fi

Abstract.
A.M., Augustyn Markiewicz, Professor at the Poznań University of Life Sciences, will make it to 65 on 29 November 2021. That made me thinking. Which should happen more often, as only my best friends dare to say.

My humble thinking results in the following pages, illustrated with some photographic glimpses. Let me mention, in passing, that the title of this talk is inspired by Saul Bellow's novel *The Adventures of Augie March*, Viking Press (1953), which I read when Augustyn had just reached the age of 20 years.

1 Polish–Scandinavian Party, 7 June 1984

Let's get going with a real stuff. [Please don't mind if you have heard this or any other forthcoming stuff earlier. Who cares.]

I Know What You Did in That Party in Poznań, 7 June 1984!

That was the title of my talk given in LinStat-2014, Linköping, Sweden (hello Martin!). The venue for this Fundamental Party was Jan Hauke's Suite; Jan was very much a bachelor those days.

Before that, on 3 June 1984, my colleague Erkki Liski and I stepped out of the Berlin–Poznań train at the Poznań Railway Station, and were friendly welcomed by Barbara Bogacka, who kindly guided us to the conference site: *International Statistical Conference in Linear Inference*, held in Poznań, 4–9 June 1984. That event was my first opportunity to enjoy live performances of some famous Polish stars of linear statistical models—among them young fresh Augustyn, with only 27 years' existence. In particular, I was an eager beaver to see alive the Dynamic Duo, Jerzy K. Baksalary & Radosław Kala.

Back to the Party where among the contributors were, under the leadership of Jerzy K. Baksalary, Katarzyna B (JKB's daughter), Augustyn Markiewicz, Barbara Bogacka, Dietrich von Rosen, Kenneth Nordström, Erkki Liski and some others (not forgetting myself which I actually almost forgot during the long night). In a nutshell: that was an unforgettable Party.

Thinking of the Party and the Poznań conference in 1984, I can too clearly imagine the terrible difference what it would have made to have missed all face-to-face conferences!

The *Proceedings* were published by Springer in 1985, edited by Tadeusz Caliński and Witold Klonecki. The Local Organizing Committee was headed by Radosław Kala and Roman Zmyślony. In particular I recall Roman's discussions with me and Erkki as we were travelling under some agreement between the Academies of Finland and Poland and Roman was then employed by the Polish Academy of Sciences. Roman, as we all know, has a remarkable career as a conference organizer, among them the StatLin/LinStat-2003 in Będlewo. (You may ask Roman why the name StatLin was used in 2003!) Roman has some 20–30 conferences and workshops on his belt, and in a way, Augustyn has been following Roman's trail in his organizing activities.

That Poznań meeting in 1984 is one predecessor of the LinStat series, but I'm not sure whether it could be called the first LinStat (there were three related meeting earlier in Wisła) but I guess that it was then so far the most international and influential LinStat.

2 1988: Email, Augustyn's PhD Thesis in the air

On 13 January 1988, I wrote a letter, a *p-letter*, to Jerzy K. Baksalary to Poznań (who in summer 1988 spent several weeks visiting C.R. Rao in Pennsylvania State University and before that George P.H. Styan in Montréal):

> Dear Jerzy: ... I have an extremely convenient and fast communication way with George [Styan]: *electronic mail*. Since last summer we have been using it almost daily. As you may know, it is a computer network, very simple to use. and quite reliable. Sometimes my memos go to Montréal only in some minutes but usually it takes an hour or so; during the weekends the systems are, however, sometimes down. ... On best days I have sent a memo to George in the morning, he has received it and replied and once again I have replied to him—and all this within one day!

Jerzy stepped on the North American continent first time on 3 May 1988. I was then visiting George in Montréal. And so George and I went to the Montréal–Mirabel Airport to welcome him.

Jerzy was rather exhausted but otherwise in excellent mood after the long LOT flight, telling that through all his flight he'd been reading and correcting the thesis of Augustyn Markiewicz. "*Oh, those youngsters and their writings ...*", was Jerzy's comment. Jerzy was Augustyn's supervisor.

3 1990: IWMS #1 in Tampere

It took some time until I met Augustyn after 1984: that was in Tampere, in December 1989. What happened was that Jerzy was visiting our department for one year, starting September 1989. One consequence was that his three hard-working young students/collaborators, Augustyn, Paweł Pordzik and Idzi Siatkowski, visited Tampere for several weeks around December 1989 – January 1990.

In August 1990, Augustyn returned to Tampere to attend the first *International Workshop on Matrices and Statistics*, IWMS, 6–8 August, see the program. Soon it became a series of workshops, so that the next one, the 28th IWMS, will be held in Manipal Academy of Higher Education, MAHE (formerly known as the Manipal University), Karnataka State, South India, 13–15 December 2021.

Over the years Augustyn has been an active attendant of the IWMS meetings and has been a member of the International Organizing Committee for several years. The next

IWMS after Tampere in 1990 for Augustyn was in Montréal, 15–16 July 1995 and I attended that as well. Augustyn was also in the IWMS in Shrewsbury, UK, July 1996. I had to skip that and thus I have no full complete set of IWMSs in my diary. (The other I missed was Voorburg in 2001.) In 1990s I met Augustyn in two further IWMSs, Fort Lauderdale (Florida) 1998, and Tampere 1999. To Fort Lauderdale, George's introduction was the following:

> The meeting will begin in the early evening of Friday, December 11, with a *reception* (in the Styan's suite at Rolling Hills) and end on the Monday, December 14, with a *cruise*, and a *gourmet dinner* (in the Armadillo); a *banquet* (in the Waterfall Cafe at Rolling Hills) is planned for the Saturday evening.

Videos from the conferences in statistics in Tampere in 1987 and 1990, prepared by Jarmo Niemelä and Reijo Sund, are available in YouTube. C.R. Rao's talks have been watched more than 7000 times. Further info on the IWMS series is in the IWMS-website and in *A short history of the IWMS*.

Augustyn was the main host of the IWMS #13: Będlewo, 18–21 August 2004, to celebrate Ingram Olkin's 80th birthday. Augustyn was heavily occupied by the endless aspects of the meeting—it was clearly his biggest so far. I was lucky to get some support for the workshop from Nokia (a remarkable Finnish company, whose cell phone Augustyn is still using) and thereby Ingram Olkin was performing as a *Nokia Lecturer*. In passing I cannot resist pointing out that I kinda live in a cell phone. Anyways, in spite of that, let's continue. In July 2004 Augustyn was having good time in the ILAS Conference in Coimbra (celebrating Richard Brualdi's 65th birthday!). In the middle of all the fun he sent me desperate emails asking me to do anything I can ever do to accelerate the arrival of Olkin's funding from Nokia: "If it is not soon in the workshop's account I don't dare to return to Poznań".

Well, everything went smoothly and the IWMS #13 was a successful meeting. In particular, we were happy to have the three Stanford Kings among us: T.W. Anderson (1918–2016), Gene H. Golub (1932–2007) and Ingram Olkin (1924–2016).

4 2007: Będlewo, Montréal, Windsor, Dortmund

The year 2007 was a good one: several joint meetings with Augustyn. First the MatTriad #2 in Będlewo in March, then *McGill Matrix Wednesday* in Montréal in May, and IWMS #16, in Windsor, Ontario, Canada, in June. In addition, in September 2007 Augustyn visited Tampere and was present in Jarkko Isotalo's thesis defence, where Götz Trenkler was the opponent; photo in Photo Section. Please don't let me forget Augustyn's several visits, often including lectures, to Tampere & Nokia. (Though less lectures in Nokia.) One visit conveniently took place in May 2000 when George received his Honorary PhD from our University.

One highlight in our spring tour in Canada in 2007 was the train trip Montréal–Toronto–Windsor on Thursday, 31 May. We had a good crowd, let me say that. On the previous evening the preparatory dinner was enjoyed at the Chinese Kam Fung Restaurant. George had organized our train tickets, and the trip was colourful, needless to say. The train was supposed to run from 09:40–20:05 staying in Toronto 15:15–16:00, thus making total

10h25, and then for our arrival, Ejaz Ahmed had booked a dinner for us at Escape Café at 21:00 (was it a Lebanese one?). Well, the train was delayed almost two hours, but no worries in such a great company, Bertie being much responsible for the entertainment division.

But that was not all for 2007. In November 23–24, 2007, Götz Trenkler and his team had organised the *Fourth Autumn Symposium of the Research Training Group on "Statistical Modelling"*; so called Bommerholz meeting, in Dortmund. Among the invitees were Augustyn, Jarkko Isotalo, Oskar Baksalary, and me as well.

On Saturday evening, 24 November, Birgit and Götz invited our quartet for dinner at Château Trenkler. Wunderbar, wasn't it! Next day, after returning back home, I sent the following email to Götz:

> The return trip of the Finnish delegation went fluently with no problems. Jarkko and I found it completely harmless that while being excessively concentrated in the flaws of our proof for a recent conjecture, we missed the train station of Düsseldorf Flughafen. Fortunately we realised this (in the same morning) and immediately stepped out of the train (waiting of course until it was fully stopped in another railway station). We were extremely lucky to find a train going back to the opposite direction and so we were breathing almost normally.
>
> I may add that after some time Jarkko was so attentive that he concluded that our current train had turned to a sidetrack and was in full speed towards Essen. This was sehr Scheiße news since even though we had nothing against Essen, it was still further away from Düsseldorf than we ever dreamed. Again we stepped out in the next possible station and in spite of the fact that we agreed that with our travelling technique we could almost certainly in due course reach the Düsseldorf Airport, we had a quick summit and we decided to take a taxi to the airport and remain seated in the taxi until we have been told that this is the requested airport. Boringly, it indeed happened.

Let's go back to MatTriad conferences. As said above the second MatTriad was held in Będlewo, in March 2007, local chair being Augustyn. The first conference in the MatTriad series was held in Będlewo, 3–5 March 2005, see the Będlewo Conference Center. From the first eight MatTriads, four have been organized in Będlewo. According to the original plan, it was supposed to be a rather small conference, concentrating on matrix analysis and its applications mainly with invited speakers, limited to three days. Hence the name MatTriad was adopted. However, since 2009, this three-day-rule has not been anymore followed. The eighth MatTriad was held in Liblice, near Prague, 8–13 September 2019. Chateau Liblice, run by the Czech Academy of Sciences, is a fabulous venue for conferences not to mention weddings/crushes/divorces, whatever, and not forgetting the impressive Princely Lounge providing inspiration for joint scientific adventures, which I enjoyed with Steve Haslett and Augustyn.

Augustyn was the chair of the first MatTriad and since then has been an instrumental person in all MatTriads. He's attended all eight meetings but I regret missing the first one in 2005. For the history of the MatTriads, see the article by Hauke, Markiewicz & Puntanen (2020): PDF, in *Applications of Mathematics*.

5 2011: Indian trips

Now it's time to get serious. Indian trips. I have made a number of them since 1990, and three of them with Augustyn (and Soile, needful to say). There is a lot to tell, India is adventurous, I can tell you. So I have to be selective and leave the rest for another occasion. I say a few words about the three Indian trips:

- 1–18 Jan 2012
  TMP–HEL–Delhi–Goa–Manipal–Kochi–Delhi–HEL–TMP
- 23 Dec 2012 – 8 Jan 2013
  TMP–HEL–Kolkata–Delhi–HEL–TMP
- 29 Nov–18 Dec 2017
  TMP–HEL–Goa–Mangalore–Madikeri–Mysore–Manipal–Goa–HEL–TMP

5.1 January 2012: Delhi–Goa–Manipal–Kochi

On Sunday, 1 January 2012. Soile and I flew Tampere–Helsinki in 35 minutes (currently no flights from Tampere at all ...), flew Helsinki–Delhi in 7 hrs with Tõnu Kollo, then took an IndiGo's 2h30 flight Delhi–Goa on Monday 2 January. Delightfully we were welcomed at Goa Airport at 14:45 by Augustyn who had flown there by Air India from Mumbai. Did good sightseeing, some of us went even swimming. Stayed in Panjim Inn. On Friday 6 January, train 14:25–18:59 from Goa to Manipal. In Manipal we attended the Workshop on *Combinatorial Matrix Theory and Generalized Inverses of Matrices* on 2–7 January, and on the related Conference on 10–11 January.
Overnight sightseeing trip to Mysore (organised by Prasad) on 8 January. Colorful longish bus trip, several jokes told and repeated by our good group. One highlight was the Ranganathittu Birdsanctuary, Karnataka, 9 January, see the group photo in Photo Section.
After the Manipal events, Augustyn, Soile and I took a train to Kochi (formerly known as Cochin). Train #16311, Bkn Kcvl Express, 500km, 10h30. An experience! See the photos on two sleepy travellers. I recall that next day we were in a pleasant beach restaurant and the waitress said to Augustyn "Your Daddy seems to be a bit out of order. Do you think a nice portion of liquid refreshments might do him any good?"

5.2 2012–2013: Kolkata-Delhi

On 23 December 2012, I flew with Kimmo Vehkalahti from Helsinki to Delhi and from there to Kolkata. Everlasting delay in Delhi for the Air India's flight to Kolkata; Kimmo sent a mountain of emails to our contact guy who was supposed to meet us at the CCU airport. After arriving five hours late our guy indeed was at the airport but told that he had not received any emails from Kimmo. Some other Indian must have received interesting emails!
Augustyn arrived on the 26th and Soile on the 27th of December. Stayed ten nights in Park Palace Hotel, whose best feature was the excellent neighbouring Mirch Masala Restaurant. Never failed.
We attended the *Eighth International Triennial Calcutta Symposium on Probability & Statistics*, see the program. We dined with the Dattas' Trio in the Aaheli Restaurant in the Peerless Inn, and were wandering around Kolkata with the competent guide Thomas

Mathew (whom I met first time in Poznań in 1984). Ratan Gasgupta showed the ISI Campus in Barrackpore Trunk Road and Bikas Sinha the great Banyan trees. Celebrated the New Year 2013 in Taj Bengal where I had been 20 years earlier with George and Somesh Dasgupta (1935–2006). (Curiosity: George and Somesh returned their Glenlivets to kitchen as too tiny.)
One day in Kolkata we went into one temple where people where queuing and we thought let's go. When strolling through the temple we admired everything accordingly but when approaching the exit, I was told with a serious tone that I should provide a reasonable valid donation. How much, I asked in a thin voice. The previous customer paid 40 dollars, was the reply. I said: I have only peanuts in my wallet, how can I get out of here, and moreover, my wife has to get out as well, and actually I'm not willing to leave her here. Give all you have in wallet, sir, then perhaps you can go. I did as asked and was terribly nervous but was shown out. Later I was told that Kimmo had provided a decent amount of rupees from his pocket. However, when Augustyn was asked he had said he has no money at all, zero. And smiled. And they let him go. Cosmo is Cosmo.
On 3 January we took the Air India flight to Delhi where we stayed two nights in the Guest House of the ISI Delhi Campus. My room with Soile was excessively cold, it was so bloody cold. Problem with windows. Stayed then three further nights in Blu Marina Hotel in Connaught Place. Did the Taj Mahal excursion, of course.

5.3 2017: Goa–Mangalore-Madikeri-Mysore-Manipal

On 29 November 2017, our Trio, Augustyn, Soile and I, had a direct flight from Helsinki to Goa, 8h10. Very convenient. Our aim was to stay four nights in Goa and then take a train to Mangalore, do some sightseeing and one academic visit and end up to Manipal on 10 December to attend the ICLAA-2017 and the special DAE-BRNS Theme Meeting. Prasad had kindly advised us for planning this round trip and indeed it was a great experience.
On the 4th of December 2017, we took a train from Goa's Madgaon Station to Mangalore (431 km). Waking up at 4am, we had a taxi leaving Panaji (former Panjim) at 5am, arriving at Madgaon at 6am. After a delayed train trip we reached Mangalore ca 4pm, from where two PhD students carried us directly to the best restaurant of Mangalore. Augustyn and I gave talks in the seminar at Mangalore University with two other speakers; see the program.
Let me mention that when entering the Goa–Mangalore express train on the 4th of December, we realised that somebody had carelessly taken Augustyn's reserved seat, and was even more carelessly but apparently fully innocently sleeping in the same. What to do? Simple: with Augustyn, Soile, and one unknown co-passanger, I raised and pushed this Sleeping Beauty into the upper level seat/bed, see the photograph in Photo Section. What we found a kind of a miracle was that while in full sleep, this guy was able to use his smartphone fluently, with no problems. Business is business.
The following episode is from Manipal, 11 December 2017. It's a personal one in the sense that Augustyn is not directly involved. In spirit, he definitely was involved. This is more or less how I put it into my Xmas Letter in December 2017. Some of you might have read it.

On Monday, 11 December, in Manipal, I was scheduled to give a talk in the *Fourth DAE-BRNS Theme Meeting on Generation and use of Covariance Matrices in the Applications of Nuclear Data*. see www. My slot was at 14:30–15:30. After having some light family chat with Soile around noon, I left the good old Fortune Inn Valley View hotel in good time, walking in deep thoughts which is my style having reached this age. The walking distance to the conference venue is ca 5–7 minutes. I knew this precisely as this was my third visit to Manipal and my tenth visit to India. After ten minutes, intensively in my thoughts, then, suddenly, I realised that I had taken a wrong way. What do do? Returned back for five minutes or something, turned right, walked another ten minutes, got a hiccup realising that I had no idea where I was. You agree this is getting interesting? Sweating, terribly: 34°C. Started running, in a cosmopolitan style, pretending to be in no hurry at all. Wanted a Tuk-Tuk express ride but was advised to go straight ahead and then turn left or something. Asked another local whether she's happened to see the University of Manipal. "Straight ahead, Ol' Man!" The bad news was it was already now time when I was allocated to deliver my talk about covariance matrices. After some powerless efforts I ended up at the Department of Statistics, 6th floor, beautiful charming building, but I was not interested in architecture, I was desperate, hopeless, hating any beauty in whatever architecture.

In the Statistics Department, I explained that I should have been speaking already half an hour in a conference which apparently was in another place. After my standing 15 minutes in coma in the Statistics Department, the Emergency Bus Driver drove me—using no siren by my request—to the correct venue to Manipal Institute of Technology. The whole episode, including the bus trip, took ca 1h10 or something. Which is a kind of achievement for a trip that should take max seven minutes by walking.

6 2006–2016: Portugal

Portugal deserves a separate section in any article dealing with Augustyn and the scientific & social meetings of our Old Boys' Gang. (I have no guts to use "girls" nor "ladies" in this context.)

My first trip to Portugal took place in September 2006: *SCRA 2006âĂŞFIM XIII, International Conference on Interdisciplinary Mathematical and Statistical Techniques*, in Tomar, 1–4 Sep 2006. It was much to do by Augustyn and Carlos Coelho that I really found myself in Tomar in September. Since then, Tomar with the leadership of Francisco Carvalho and João Tiago Mexia has been an excellent venue for LinStat, MatTriad and IWMS conferences. The IWMS #25 was organised in Madeira in 2016 by Francisco, and MatTriad #6 in Coimbra in 2015 by Natália Bebiano.

One morning in Tomar in September 2006, actually it was the day when C.R. Rao received his Honorary PhD in Tomar, Augustyn and I were slowly strolling towards the place of the festive ceremonies, wearing our ultra super black suits. You cannot be surprised to hear that we had black ties, shiny black shoes and dark sunglasses. I had even color-washed my problematic hair so that it would be almost as black & shiny as that of Augustyn. Then, of course we lost our way in minutes but to our luck we suddenly met a youngish shining lady. Excuse us, dear, but would you mind telling where we are now, and besides, are we right now in Portugal, was our passionate question. The expression in the lady's face and her mysterious hymning were due to her mixing us with the celebrated movie stars *Men in Black*, that was our conclusion.

7 2003–2020: Będlewo

My latest trip abroad was to Poland, Będlewo, 16–21 February 2020, when Kasia Filipiak & Augustyn organised the biannual MMLM research group meeting. That appeared to be the last-moment opportunity for the old-school meeting for some time. Then, none in our group had any idea how hard the Wuhan phenomenon would hit.

In Będlewo, from 2009 onwards, there has been an International Research Group Meeting on Multivariate and Mixed Linear Models, shortly MMLM, supported by the Stefan Banach International Mathematical Center, and organised by Kasia and Augustyn. The number of participants has been mostly below 20 and the mutual discussions on joint problems are the main and most significant part of the meetings. I have had an opportunity to attend these biannual meetings since 2013. For a bit of history, see Augustyn's report in 2019 meeting's Abstract Book.

As a result of these meetings Springer has new book on pipeline, see the website, editors being Kasia, Augustyn and Dietrich. I will have there a paper, ca. 54 pages, with Steve, Jarkko, Radek Kała and Augustyn. So far my longest paper has been the DLLPS, 54 pages, appearing in *Sankhyā*, in 1992, with four coauthors.

These Będlewo meetings are something unique. No standard meetings at all. Hardly any talks but can be some. Discussions, a lot. Free forum. Nicely motivating for me, indeed. I could give guided walks in the neighbouring forest, warn you about beavers, show the track to the lake and around. Confession: maybe once, perhaps it was in 2013, I missed one afternoon presentation by making an unnecessarily extended after-lunch walk with Augustyn. That could happen; if one has serious plans to revise the manuscript thoroughly.

I almost forgot one thing: why am I doing so well in Poland? Some of you may know but the plain fact is that sometime in 1990s I bought in Delhi Airport *Simple Etiquette in Poland*, by Krystyna Carter, a guide book first published in UK in 1992 but with the first Indian edition in 1994 (for sale in Indian subcontinent only). There are lots of advises, some really old ones and thereby nicely marinated and tested. For example, a sixteenth century traveller to Poland, William Bruce, noted in 1598: "The Poles are large of body, tall, uprighte and personable. The gentry full of ceremonies, civill and courteous in entertainement, bountifull at table, costly in dyett, great gourmandes, and quaffers, not sleepy, nor heavy in their dronkenesse, . . . , high-mynded, and proude . . . "

Acknowledgements

This article was prepared for possible publication in the Abstract Book of the LinStat Conference, held in hybrid form, 30 August – 3 September 2021 in Będlewo, Poland, www. Most photographs of this paper are taken by me but some by Soile, Augustyn, Kimmo, Kasia and Somnath Datta – thank you very much!

And to Augustyn, very many happy returns and adventures for your 65th birthday, 29 November 2021!

8 Photographs

Figure 1. A Pole on a Goan Beach, 3 January 2012.

Figure 2. That’s it. Kolkata, 31 December 2012.

Figure 3. Nordic–Polish Party in Jan Hauke's suite, 7 June 1984. Jerzy K. Baksalary, Katarzyna B (JKB's daughter), Dietrich von Rosen.

Figure 4. Augustyn, Jerzy, Jan, June 1984.

Figure 5. Augustyn, Jerzy, Idzi Siatkowski and Paweł Pordzik in Tampere Railway Station, December 1989.

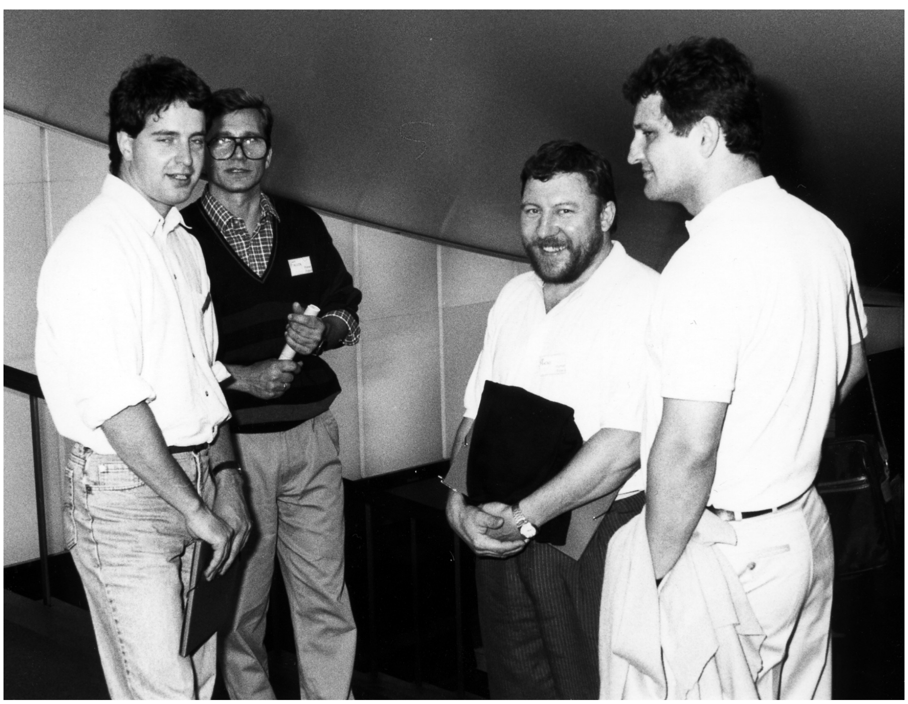

Figure 6. Bernd Schipp, Paweł Pordzik, Jan, Augustyn. IWMS #1, Tampere, 6–8 August 1990.

Figure 7. Jerzy lecturing, IWMS #1, Tampere, August 1990. Front row: Shanti S. Gupta, J.N. Srivastava, Song-Gui Wang, Götz Trenkler, Stanisław Gnot. 2nd & 3rd row: Jürgen Kleffe, Roman Zmyślony, Augustyn Markiewicz, Jan Hauke, Tõnu Kollo.

Figure 8. IWMS #7, Fort Lauderdale, Florida, celebrating T.W. Anderson's 80th birthday, 11–14 December 1995. Program.

Figure 9. Tadeusz Caliński, Radosław Kala, Götz Trenkler, Roman Zmyślony, Augustyn. StatLin conference in Będlewo, 21–27 August 2003.

Figure 10. IWMS #13, Będlewo, 18–21 August 2004. Abstract Book.

Figure 11. Gene H. Golub, X, Bikas K. Sinha, Ingram Olkin, Augustyn, Ludvig Elsner, Yongge Tian. Będlewo, August 2004.

Figure 12. IWMS–IOC meeting in Będlewo, August 2004. George, Götz, Augustyn, Tõnu, Dietrich, Jochen.

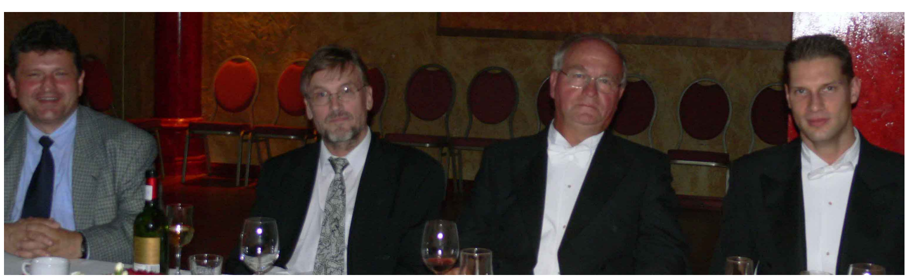

Figure 13. Jarkko Isotalo's Thesis Defence Dinner, Tampere, 7 September 2007. Opponent: Götz Trenkler.

Figure 14. Augustyn, Erkki Liski and Friedich Pukelsheim, IWMS #15 in Uppsala, 13–16 June 2006.

Figure 15. International Conference on Interdisciplinary Mathematical and Statistical Techniques, Tomar, Portugal, 1–4 September 2006. C.R. Rao received an Honorary PhD. Can you spot Augustyn?

Figure 16. MatTriad in Będlewo, 22–24 March 2007. George made it to 70.

Figure 17. Above: *McGill Matrix Wednesday*, Montréal, 20 May 2007. Below: waiting for the train Montréal–Windsor, 31 May 2007.

Figure 18. IWMS #15, in Windsor, Ontario, Canada, 1–3 June 2007.

Figure 19. Shuangzhe Liu, Augustyn, George, Yogendra P. Chaubey; in Ejaz Ahmed's home. Windsor, June 2007. Who made it up to 70 that year?


Figure 20. Martin Singull was in charge of LinStat-2014 in Linköping. There is a wonderful beach in Linköping.


Figure 21. Birsen Eygi Erdogan, Augustyn, Francisco Carvalho etc, planning the LinStat-2016/Istanbul in Linköping, 25 August 2014, at 23:40.
Below: Istanbul, LinStat, 22–25 August 2016.

Figure 22. International Workshop & Conference on Combinatorial Matrix Theory and Generalized Inverses of Matrices, Manipal University, Manipal, India, 2–7 & 10–11 January 2012.

IRCTCs e-Ticketing Service
Electronic Reservation Slip (Personal User)

- This ticket will only be valid along with an ID proof in original. If found travelling without ID Proof, Passenger will be treated as without ticket and charged as per extant Railway rules.
- Valid IDs to be presented during train journey by one of the passenger booked on an e-ticket :- Voter Identity Card / Passport / PAN Card / Driving License / Photo ID card issued by Central / State Govt. / Student Identity Card with photograph issued by recognized School or College for their students / Nationalized Bank Passbook with photograph /Credit Cards issued by Banks with laminated photograph.
- General rules/ Information for e-ticket passenger have to be studied by the customer for cancellation & refund.

PNR No: 8548657629	Train No. & Name: 56641/MAO MAQ PASS	Quota: General
Transaction ID: 0404503473	Date of Booking: 15-Oct-2011 04:47:14 PM	Class: 2S
From: MADGAON(MAO)	Date of Journey: 06-Jan-2012	To: UDUPI(UD)
Boarding: MADGAON(MAO)	Date of Boarding: 06-Jan-2012	Scheduled Departure: 14:25 *
Resv Upto: UDUPI(UD)	Distance: 0350 KM	Adult: 03 Child: 00
Passenger Mobile Number: 9980100886		
Passenger Address :- Sri Kirshna, Kotathattu,KOTA Post Udupi Karnataka - 576221		

FARE DETAILS :

1	Ticket Fare	Rs. 180	Rupees One Hundred and Eighty Only
2	IRCTC Service Charges	Rs. 10	Rupees Ten Only
3	Total	Rs. 190	Rupees One Hundred and Ninety Only

PASSENGER DETAILS :

SNo.	Name	Age	Sex	Concession Code	Booking Status/Current Status/Coach No./Seat No
1	Simo Puntanen	66	Male		CONFIRM D1/ 0049/ WS
2	Soile Puntanen	65	Female		CONFIRM D1/ 0050/
3	Augustyn Markiew	55	Male		CONFIRM D1/ 0051/

Figure 23. Goa–Manipal: 350 km = 0.86 EUR, 6 January 2012. Madgaon = train station in Panaji (Goa), Udupi = train station in Manipal.

Figure 24. Goa, 3 January 2012.

Figure 25. Jeff checking his camera in Manipal banquet, 10 January 2012; Jeff, Hazel, SP, Augustyn, Tonu Kõllo, Steve Kirkland.

Figure 26. Ranganathittu Bird Sanctuary, Karnataka, India, 9 January 2012. Excursion arranged by Prasad when attending the International Workshop & Conference in Manipal.

Figure 27. Jan Hauke's home, July 2012. In the context of LinStat/IWMS. Ivan Žežula, Jan, Prasad, SP, Augustyn.

Figure 28. Above: Kolkata 27 December 2012; Toronto, August 2013. Below: Manipal–Kochi train 12 January 2012, 11:56–22:25.

Figure 29. Bargaining in Kochi; visiting the Kochi University, 16 January 2012.

Figure 30. Augustyn: a celebrated dancer, here in Tomar, LinStat, July 2010. He can dance even Fado. Poor me: a developing but eager dancer, here in Kolkata, Taj Bengal Hotel, 31 December 2012. Below: Manipal University Dance Group, January 2012.

Figure 31. Thomas Mathew, Kimmo Vehkalahti and Augustyn, in the front of Park Palace Hotel, Kolkata, stayed 10 nights, from 24 Dec 2012 to 3 Jan 2013. Below: Augustyn thanking the nearby Mirch Masala Restaurant. Our goal: The Eighth International Triennial Calcutta Symposium on Probability & Statistics. www.

Figure 32. Dancing, *matrix*! How to dance the *matrix*? Ever done that? Kolkata, 29 December 2012.

Figure 33. Above: Soile, Augustyn, Thomas, Kimmo, in the night of Kolkata, 28 Dec 2012. Below: With the Dattas in Aaheli Restaurant in Peerless Inn, 29 Dec 2012.

Figure 34. Kolkata, 31 December 2012. After visiting a temple.

Figure 35. With Bikas Sinha et al., Victoria Memorial, Kolkata, 2 January 2013.

Figure 36. Taj Mahal, 6 January 2013.

Figure 37. With Mika Mattila in Montenegro, MatTriad, 16–20 September 2013.

Figure 38. IWMS #22, Toronto, 12–15 August 2013.

Figure 39. Manipal conference was our goal and the route was Helsinki–Goa–Mangalore–Madikeri–Mysore–Manipal–Goa–Helsinki.

Figure 40. 1 December 2017, Panaji, Goa.

Figure 41. Horizontal travelling, 4 December 2017. Goa-Mangalore train # Mangalore Exp-12133, 5h26, Coach AB1, seat/berth LB/4, LB/1, SL/7; Class: 3A.

Figure 42. Mangalore University, 5 December 2017.

Figure 43. Martin & Augustyn in Manipal. Cosmo is Cosmo!

Figure 44. Cultural event, 13 December 2017, Karanth Bhavan, Kota, near Manipal. Soile, Martin, Hilary, Steve H., Augustyn, Hazel, Jeff, Steve K., Michael Tsatsomeros.

Figure 45. Cultural event, Karanth Bhavan, Kota, 13 December 2017.

Figure 46. Sunset in Manipal, 16 December 2017.

Figure 47. Goddesses of Learning in Nokia, December 2017: the bigger one from Tirunelveli in Dec 1993, the smaller one from Manipal in Dec 2017. Augustyn in Montenegro, 17 September 2017.

Figure 48. Princely Lounge, MatTriad conference in Liblice, Czech Republic, 8–13 September 2019.

Figure 49. Steve Haslett, Dário Ferreira, Hilary Smith, Sanda Ferreira, Augustyn, Soile. Banquet in Liblice, September 2019.

Figure 50. Reflections in Będlewo, 14 November 2019. Can you spot a human being?

Figure 51. MMLM group in Będlewo, 12 November 2015.

Figure 52. Nokia, May 2000.

Figure 53. Tomar, September 2006; Nokia, May 2013.

Part VIII

List of participants

Participants

1. **Rauf Ahmad**
   Department of Statistics, Uppsala University, Uppsala, Sweden
   rauf.ahmad@statistik.uu.se
2. **Selin Alica**
   Department of Econometrics, Faculty of Economics and Administrative Sciences, Ankara Haci Bayram Veli University, Ankara, Turkey
   selin.alica@hbv.edu.tr
3. **Stanislav Anatolyev**
   CERGE-EI, Charles University, Prague, Czech Republic
   stanislav.anatolyev@cerge-ei.cz
4. **Lukas Arnroth**
   Department of Statistics, Uppsala University, Uppsala, Sweden
   lukas.arnroth@statistik.uu.se
5. **Anthony C. Atkinson**
   Department of Statistics, London School of Economics, London, UK
   a.c.atkinson@lse.ac.uk
6. **Olgun Aydin**
   Department of Statistics, Faculty of Management and Economics, Gdansk University of Technology, Gdansk, Poland
   olgun.aydin@pg.edu.pl
7. **Adelchi Azzalini**
   Department of Statistical Sciences, University of Paduay, Padua, Italy
   azzalini@stat.unipd.it
8. **Rosemary Bailey**
   School of Mathematics and Statistics, University of St Andrews, St Andrews, UK
   rab24@st-andrews.ac.uk
9. **Narayanaswamy Balakrishnan**
   Department of Mathematics and Statistics, McMaster University, Hamilton, Canada
   bala@mcmaster.ca
10. **Luckshay Batra**
    Department of Applied Mathematics, Delhi Technological University, Delhi, India
    luckshay25@gmail.com
11. **Beste Hamiye Beyaztaş**
    Department of Statistics, Faculty of Engineering and Natural Sciences, Istanbul Medeniyet University, Istanbul, Turkey
    beste.sertdemir@medeniyet.edu.tr
12. **Ufuk Beyaztaş**
    Department of Statistics, Faculty of Science and Literature, Marmara University, Istanbul, Turkey
    ufuk.beyaztas@marmara.edu.tr

13. **Mariusz Bieniek**
Institute of Mathematics, Maria Curie Skłodowska University, Lublin, Poland
mariusz.bieniek@umcs.lublin.pl

14. **Aylin Çakiroğlu**
Computational Biology and Bioinformatics, The Francis Crick Institute, London, UK
aylin.cakiroglu@crick.ac.uk

15. **Mustafa Cavus**
Department of Statistics, Eskisehir Technical University, Eskisehir, Turkey
mustafacavus@eskisehir.edu.tr

16. **Deliang Dai**
Department of Economics and Statistics, School of Business and Economics, Linnaeus University, Växjö, Sweden
deliang.dai@lnu.se

17. **Rukiye Dağalp**
Department of Statistics, Ankara University, Ankara, Turkey
rukiyedagalp@gmail.com

18. **Somnath Datta**
Department of Biostatistics, University of Florida, Gainesville, USA
somnath.datta@ufl.edu

19. **Anna Dembińska**
Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland
dembinsk@mini.pw.edu.pl

20. **Efthymia Derezea**
School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, UK
ed378@kent.ac.uk

21. **Elvira Di Nardo**
Department of Mathematics, University of Torino, Torino, Italy
elvira.dinardo@unito.it

22. **Anıl Eralp**
Department of Econometrics, Bolu Abant İzzet Baysal University, Bolu, Turkey
anil.eralp@ibu.edu.tr

23. **Susana Faria**
Department of Mathematics, University of Minho, Guimarães,, Portugal
sfaria@math.uminho.pt

24. **Dário Ferreira**
Department of Mathematics, Faculty of Sciences, University of Beira Interior, Covilhã, Portugal
dario@ubi.pt

25. **Sandra Maria Bargão Ferreira**
Department of Mathematics, Faculty of Sciences, University of Beira Interior, Covilhã, Portugal
sandraf@ubi.pt

26. **Katarzyna Filipiak**
Institute of Mathematics, Poznań University of Technology, Poznań, Poland
katarzyna.filipiak@put.poznan.pl

27. **Eva Fišerová**
Department of Mathematical Analysis and Applocations of Mathematics, Palacký University Olomouc, Olomouc, Czech Republic
eva.fiserova@upol.cz

28. **Miguel Fonseca**
Center of Mathematics and its Applications, Nova University of Lisbon, Lisbon, Portugal
fmig@fct.unl.pt

29. **Johannes Forkman**
Department of Crop Production Ecolgoy, Swedish University of Agricultural Sciences, Uppsala, Sweden
johannes.forkman@slu.se

30. **Andrej Gajdoš**
Institute of Mathematics, Faculty of Science, P. J. Šafárik University in Košice, Košice, Slovakia
andrej.gajdos@upjs.sk

31. **Emelyne Umunoza Gasana**
Department of Mathematics, Linköping University, Linköping, Sweden
emelyne.umunoza.gasana@liu.se

32. **Şahika Gökmen**
Department of Statistics, Uppsala University, Uppsala, Sweden
sahika.gokmen@statistik.uu.se
Department of Econometrics, Ankara Haci Bayram Veli University, Ankara, Turkey
sahika.gokmen@hbv.edu.tr

33. **Tomasz Górecki**
Department of Mathematical Statistics and Data Analysis, Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland
tomasz.gorecki@amu.edu.pl

34. **Agnieszka Goroncy**
Department of Mathematical Statistics and Data Mining, Nicolaus Copernicus University, Toruń, Poland
gemini@mat.umk.pl

35. **Małgorzata Graczyk**
Department of Mathematical and Statistical Methods, Faculty of Agronomy, Horticulture and Bioengineering, Poznań Uniwersity of Life Sciences, Poznań, Poland
malgorzata.graczyk@up.poznan.pl

36. **Radoslav Harman**
Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Bratislava, Slovakia
radoslav.harman@fmph.uniba.sk

37. **Jan Hauke**
Department of Regional and Local Studies, Faculty of Human Geography and Planning, Adam Mickiewicz University, Poznań, Poland
jhauke@amu.edu.pl

38. **Marie Hušková**
Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic
huskova@karlin.mff.cuni.cz

39. **Malwina Janiszewska**
Department of Mathematical and Statistical Methods, Faculty of Agriculture, Horticulture and Bioengineering, Poznań University of Life Sciences, Poznań, Poland
malwina.janiszewska@up.poznan.pl

40. **Krzysztof Jasiński**
Faculty of Mathematics and Computer Science, Faculty of Mathematics and Computer Science, Nicolaus Copernicus University in Toruń, Toruń, Poland
krzys@mat.umk.pl

41. **Mateusz John**
Institute of Mathematics, Poznań University of Technology, Poznań, Poland
mateusz.john@put.poznan.pl

42. **Barbora Kessel**
Department of Epidemiology, Helmholtz Centre for Infection Research, Braunschweig, Germany
barbora.kessel@gmail.com

43. **Razieh Khodsiani**
Department of Computer Science and Statistics, Faculty of Mathematics, K. N. Toosi University of Technology, Tehran, Iran
razieh.khodsiani@alumni.iut.ac.ir

44. **Busenur Kızılaslan**
Department of Statistics, Marmara University, Istanbul, Turkey
busenur.sarica@marmara.edu.tr

45. **Daniel Klein**
Institute of Mathematics, P. J. Šafárik University in Košice, Košice, Slovakia
daniel.klein@upjs.sk

46. **Arkadiusz Kozioł**
Institute of Mathematics, Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Zielona Góra, Poland
a.koziol@wmie.uz.zgora.pl

47. **Agnieszka Łacka**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland
agnieszka.lacka@up.poznan.pl

48. **Yuli Liang**
Department of Statistics, School of business, Örebro University, Örebro, Sweden
yuli.liang@oru.se

49. **Nicholas Longford**
School of Public Health, Imperial College London, London, UK
sntlnick@sntl.co.uk

50. **Adewale Lukman**
Department of Mathematics, Landmark University, Omu-Aran, Nigeria
adewale.folaranmi@lmu.edu.ng

51. **Augustyn Markiewicz**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland
amark@up.poznan.pl

52. **Adam Mieldzioc**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland
adam.mieldzioc@up.poznan.pl

53. **Monika Mokrzycka**
Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland
mmok@igr.poznan.pl

54. **Müge Mutiş**
Department of Statistics, Graduate School Of Natural And Applied Sciences, Yıldız Technical University, Istanbul, Turkey
mugemutis@gmail.com

55. **Luísa Novais**
Department of Mathematics and Centre of Molecular and Environmental Biology, University of Minho, Guimarães, Portugal
luisa_novais92@hotmail.com

56. **Tatjana Pavlenko**
Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden
pavlenko@math.kth.se

57. **Virgilio Pérez Giménez**
Applied Economics Department, Faculty of Economics, University of Valencia, Valencia, Spain
virgilio.perez@uv.es

58. **Jolanta Pielaszkiewicz**
Department of Computer and Information Science, The Division of Statistics and Machine Learning, LinkÃűping, Sweden
jolanta.pielaszkiewicz@liu.se

59. **Simo Puntanen**
Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland
simo.puntanen@tuni.fi

60. **Veronika Římalová**
Department of Mathematical Analysis and Applications of Mathematics, Palacký University Olomouc, Olomouc, Czech Republic
veronikarimalova@seznam.cz

61. **Robert Sebastian Castellanos Rodriguez**
Department of Statistics, National University of Colombia, Bogota, Colombia
rscastellanosr@unal.edu.co

62. **Samuel Rosa**
Department of Applied Mathematics and Statistics, Comenius University, Bratislava, Slovakia
samuel.rosa@fmph.uniba.sk

63. **Dietrich von Rosen**
Department of Energy and Technology, Swedish University of Agricultural Sciences, Uppsala, Sweden
dietrich.von.rosen@slu.se

64. **Tomasz Rychlik**
Department of Mathematical Statistics, Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland
trychlik@impan.pl

65. **Semanur Sarıçam**
Department of Statistics, Graduate School Of Natural And Applied Sciences, Mimar Sinan Fine Arts University, Istanbul, Turkey
semanursaricam@gmail.com

66. **Volodymyr Sarioglo**
Department of Socio Economic Processes and Structures Modelling, Ptoukha Institute for Demography and Social Studies, Kyiv, Ukraine
sarioglo@idss.org.ua

67. **Francesco Schirripa Spagnolo**
Department of Economics and Management, University of Pisa, Pisa, Italy
francesco.schirripa@ec.unipi.it

68. **Rakhi Singh**
Department of Informatics and Analytics, UNC Greensboro, Greensboro, USA
r_singh@uncg.edu

69. **Martin Singull**
Department of Mathematics, Linköping University, Linköping, Sweden
martin.singull@liu.se

70. **Łukasz Smaga**
Department of Mathematical Statistics and Data Analysis, Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland
ls@amu.edu.pl

71. **Maria-Zafeiria Spyropoulou**
Department of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, UK
mzs2@kent.ac.uk

72. **Anna Szczepańska-Álvarez**
Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poznań, Poland
anna.szczepanska-alvarez@up.poznan.pl

73. **Magdalena Szymkowiak**
Institute of Automation and Robotics, Poznań University of Technology, Poznań, Poland
magdalena.szymkowiak@put.poznan.pl

74. **Müjgan Tez**
Department of Statistics, Marmara University, Istanbul, Turkey
mtez@marmara.edu.tr

75. **Annika Tillander**
Department of Computer and Information Science, Linköping University, Linköping, Sweden
annika.tillander@liu.se

76. **Denise Uwamariya**
Department of Mathematical Statistics, Linköping University, Linköping, Sweden
denise.uwamariya@liu.se

77. **Oleg Volkov**
Department of Energy Resources Engineering, Stanford University, Stanford, USA
volkovoleg@hotmail.com

78. **Viktor Witkovský**
Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
witkovsky@savba.sk

79. **Fatma Yerlikaya-Özkurt**
Industrial Engineering, Engineering, Atılım University, Ankara, Turkey
fatmayerlikaya@gmail.com

80. **Özge Cağcağ Yolcu**
Department of Statistics, Faculty of Arts and Sciences, Marmara University, Istanbul, Turkey
ozgecagcag@yahoo.com

81. **Ivan Žežula**
Institute of Mathematics, P. J. Šafárik University in Košice, Košice, Slovakia
ivan.zezula@upjs.sk

82. **Roman Zmyślony**
Institute of Mathematics, Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Zielona Góra, Poland
r.zmyslony@wmie.uz.zgora.pl

Index