University of Central Lancashire
UCLan Cyprus

ΣtatMod

ΣtatMod2026

Statistical Modelling with Applications

Book of Abstracts

18–19 September 2026
UCLan Cyprus, Larnaka, Cyprus

Table of Contents

Welcome Message

On behalf of the Organising Committee, it is our great pleasure to welcome you to *ΣtatMod2026: Statistical Modelling with Applications*, held at UCLan Cyprus in Larnaka on 18-19 September 2026.

Statistical modelling continues to play a fundamental role in addressing scientific and societal challenges across a broad spectrum of disciplines, including data science, artificial intelligence, engineering, finance, health sciences, environmental sciences and reliability. The aim of ΣtatMod2026 is to provide an international forum where researchers and practitioners can present recent advances, exchange ideas and establish new collaborations.

This year's conference brings together participants from universities, research centres and industry across Europe and beyond. The programme reflects the breadth of contemporary statistical research, ranging from theoretical developments to methodological innovations and real-world applications.

The Organising Committee would like to express its sincere appreciation to all invited speakers, contributed speakers, members of the Steering Committee, Scientific Committee and reviewers for their invaluable support. We also thank our sponsors for their support and UCLan Cyprus for hosting the conference and providing an excellent environment for scientific exchange.

We hope that you will enjoy both the scientific programme and the opportunity to meet colleagues from around the world. We wish you a productive conference and a pleasant stay in Cyprus.

The Organising Committee

Committees

Steering Committee

- **Vlad Stefan Barbu**, Founding Member, LMRS, Univ. of Rouen-Normandy, France & "Costin C. Kiritescu" National Institute of Economic Research of Romanian Academy, Romania
- **Salim Bouzebda**, Univ. of Technology of Compiègne, France
- **Guglielmo D'Amico**, Univ. "G. d'Annunzio" of Chieti-Pescara, Italy
- **Milto Hadjikyriakou**, Acting Chair, Univ. of Central Lancashire, Cyprus
- **Alex Karagrigoriou**, Univ. of Piraeus, Greece
- **Bojana Milošević**, Univ. of Belgrade, Faculty of Mathematics, Serbia
- **Vasile Preda**, Founding Member, Honorary Chair, "Gheorghe Mihoc-Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of Romanian Academy & "Costin C. Kiritescu" National Institute of Economic Research of Romanian Academy, Univ. of Bucharest, Romania.

Scientific Committee

- **Christophe Chesneau**, LMNO, Univ. of Caen-Normandy, France
- **Luminița Chivu**, INCE, Romania
- **Mitra Fouladirad**, École Centrale Marseille, France
- **Gabriela Marinoschi**, Romanian Academy & ISMMA, Romania
- **Valentin Patilea**, ENSAI Rennes & CREST, France
- **Serguei Pergamenchtchikov**, LMRS, Univ. of Rouen-Normandy, France
- **Ionel Popescu**, Univ. of Bucharest & IMAR, Romania
- **Razvan Sfetcu**, Univ. of Bucharest, Romania

Local Organising Committee

- **Milto Hadjikyriakou**, UCLan Cyprus
- **Andreas Makrides**, University of the Aegean
- **Andreas Artemiou**, University of Limassol
- **Petroula Mavrikou**, Frederick University
- **Haritini Tsangari**, University of Nicosia
- **Costas Christophi**, Cyprus University of Technology
- **Kyriaki Michailidou**, Cyprus Institute of Neurology and Genetics
- **Andreas Anastasiou**, University of Cyprus

International Organising Committee

- Silvia Dedu, Bucharest Univ. of Economic Studies & INCE, Romania
- Carmen Adriana Gheorghe, INCE, Romania
- Emmanouil-Nektarios Kalligeris, Circana, Greece
- Ioannis Mavrogiannis, M2P2, Centrale Méditerranée, France
- Kimon Ntotsis, Department of Respiratory Sciences, Univ. of Leicester, UK

- Nicolas Vergne, LMRS, Univ. of Rouen-Normandy, France

Sponsors

The Organising Committee gratefully acknowledges the generous support of our sponsors. Their contribution has played an important role in the successful organisation of the **StatMod2026 Conference** and is sincerely appreciated.

We thank all our sponsors for their commitment to supporting research, education, and collaboration within the mathematical and statistical sciences.

Gold Sponsors

University of Central Lancashire
UCLan Cyprus

Silver Sponsors

Clinbay

Grant Thornton

Needlestack Technologies

The influence of parameters on multivariate optimization in inverse modelling problems

Adela Ionescu[1]

[1]University of Craiova, Romania
adelajaneta2015@gmail.com

Abstract. Inverse problems represent an area strongly connected with various branches of science and technology, becoming more and more useful not only in mathematics, but also in important areas of technology – such as computer tomography, image analysis etc.
The present paper focuses on approaching the inverse modelling problem for systems arising from excitable media, and analysing the undergoing statistical context. There are a lot of techniques to analyse the so-called observational noise we want to reconstruct from data, and to decide between a "well-posed" or "ill-posed" problem. In finite dimensions, the *matrix inversion* is a feasible appliance, since the *condition number* of a matrix is a measure of how stable a system with observable data can be solved.
In the case of mixing flow model this technique is easy to apply and gives interesting results. It is approached a way to reconstruct the system's matrix depending on the parameters' variance. For 2d mixing flow dynamical system, the condition number produced a function of several variables whose comparative analysis in slightly perturbed cases of the model, is well realized with multivariate optimization tools.
The results will be further used for completing the stability analysis panel of mixing flow dynamical system with statistical features.

References

[1] Benning, M, Burger, M. (2018). Modern regularisation methods for inverse problems, *Acta Numerica*, **27**, 1-111

[2] Fausett, L. (1994). *Fundamentals of Neural Networks*. New York: Prentice Hall

[3] Han, J., Kamber, M. (2000). *Data mining: Concepts and Techniques*. New York: Morgan-Kaufman

[4] Korolev, Y, Latz, J. (2021*). Inverse Problems*. Lect. Notes. Cambridge University Press

[5] Ripley, B.D. (1996*). Pattern Recognition and Neural Networks*. Cambridge University Press

Evaluating Multi-Regime Operational Sensitivity in Urban Transit: A Transition from Discrete to Bayesian Continuous-Time Markov Models

Adyda Ibrahim[1], Mohd Azri Pawan Teh[1] & Dadang Amir Hamzah[2]

[1]Universiti Utara Malaysia, Malaysia, [2]President University, Indonesia.

adyda@uum.edu.my; mohd.azri.pawan@uum.edu.my; dadang.hamzah@president.ac.id

Abstract. Discrete-Time Markov Chains (DTMCs) often fail to capture fluid passenger movements across continuous time intervals. To improve this, a Continuous-Time Markov Chain (CTMC) is used to model passenger flow on the Kuala Lumpur Monorail system. Initial Chi-Squared tests successfully reject independence ($p < 0.05$), indicating a station destination is influenced by the origin. However, since the stationarity tests fail, the discrete model is extended to a continuous one. A Bayesian Markov Chain Monte Carlo (MCMC) is used to estimate the posterior transition intensities. Network sensitivity is evaluated by stratifying empirical data into Weekday (43 days) and Weekend (17 days) regimes to test the impact of planned headways (7, 8, 10, and 12 minutes). This modelling approach quantifies operational risk across different schedules by demonstrating that managing transit risk requires balancing train wait times against passenger flow variance.

References

[1] Cascetta, E. (2009). *Transportation Systems Analysis: Models and Applications*. Springer.

[2] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.

[3] Sono, N. U., Ibrahim, A. & Hamzah, D. A. (2026). Modelling passenger flow of KL monorail using Markov chain. To be published in *IOP Journal of Physics: Conference Series*.

Parameter Estimation and State Decoding in Hidden Drifting Markov Models

Aguêmon Léon Thungue[1], Vlad Stefan Barbu[2], Nicolas Vergne[3]

[1]University of Rouen Normandy, France, [2]University of Rouen Normandy, France, [3]University of Rouen Normandy, France.

aguemon.tehungue@univ-rouen.fr;
vladstefan.barbu@univ-rouen.fr; nicolas.vergne@univ-rouen.fr

Abstract. Hidden Drifting Markov Models (HDMMs) extend classical Hidden Markov Models by allowing the underlying dynamics to evolve over time. They provide a flexible probabilistic framework for modeling non-stationary time series characterized by latent structures and gradually changing regimes.

In this presentation, we introduce the theoretical foundations of HDMMs and discuss the main assumptions underlying these models. We then present the inference and learning procedures associated with this framework. Particular attention is devoted to the Viterbi algorithm, used to reconstruct the most probable sequence of hidden states, and to the Expectation--Maximization (EM) algorithm, employed for parameter estimation.

Several numerical experiments based on simulated data are conducted to evaluate the performance of the proposed methods. In particular, we investigate the stability of the EM algorithm, its sensitivity to initialization, and the accuracy of the parameter estimates as the length of the observation sequence increases.

The results demonstrate that HDMMs are capable of effectively capturing non-stationary hidden dynamics while maintaining robust and reliable inference and estimation procedures. They also highlight the complementary roles of the Viterbi and EM algorithms for state decoding, model interpretation, and parameter learning in drifting hidden Markov models.

References

[1] Nicolas Vergne. Drifting markov models with polynomial drift and applications to dna sequences. Statistical Applications in Genetics & Molecular Biology, 7(1), 2008.

[2] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. Journal of the royal statistical society: series B (methodological), 39(1):1–22, 1977.

Robust tests for moment condition models

Aida Toma[1,2], Amor Keziou[3]

[1]Bucharest University of Economic Studies, Romania [2]"Gheorghe Mihoc – Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy, Romania, [3]Universite de Reims Champagne-Ardenne, CNRS, LMR, France.

aida.toma@csie.ase.ro; amor.keziou@univ-reims.fr

Abstract. We present a wide class of robust minimum empirical divergence estimators and corresponding tests for moment condition models. This class extends the minimum empirical divergence estimators and contains the robust exponential tilting estimator and a robust version of the empirical likelihood estimator, as particular cases. The new proposed estimation method makes use of truncated orthogonality functions and the minimization of divergences in their dual form. We present theoretical robustness properties and asymptotic properties for these estimators and test statistics. An important aspect is that the asymptotic properties of the estimators are valid, both under the model and under alternative hypotheses, including misspecification. The performances of these statistical methods are illustrated through examples based on Monte Carlo simulations in the case of Hall and Horowitz model.

Early stopping for avoiding wasting time with random features

Alain Celisse[1]

[1]University of Paris 1 Panthéon-Sorbonne, France.

alain.celisse@univ-paris1.fr

Abstract. We address the question of estimating a regression function in a non-parametric setting by means of a reproducing kernel (and induced RKHS) [1]. For reducing the computation burden, the canonical feature map is replaced by an low-dimensional approximation resulting from Random Features [2]. However a key ingredient in the final statistical performance achieved by such a (random) map is the number of random features. The purpose of Early Stopping Rule (ESR) [3] is to provide a data-driven strategy for choosing how many random features should be used for reaching an optimal statistical performance. In particular, it strongly differs from alternative model selection strategies by avoiding the useless computation of all candidate estimators. We will discuss how to derive such a rule and detail a few of its statistical properties

References

[1] Gretton, A. (2013). Introduction to rkhs, and some simple kernel algorithms. *Adv. Top. Mach. Learn. Lecture Conducted from University College London*, *16*(5-3), 2.

[2] Rahimi, A., & Recht, B. (2007). Random features for large-scale kernel machines. *Advances in neural information processing systems*, *20*.

[3] Celisse, A., & Wahl, M. (2021). Analyzing the discrepancy principle for kernelized spectral filter learning algorithms. *Journal of Machine Learning Research*, *22*(76), 1-59.

Evaluation of Seismic Activity via Poisson Hidden Markov Models

Alex Karagrigoriou

University of Piraeus, Greece

alex.karagrigoriou@unipi.gr

Abstract. Seismic hazard evaluation is always considered to be a research problem of high scientific importance due to the consequences that are often associated with earthquake occurrence. One such region that attracts scientific attention due to its techniques is the region covering the areas of Central and South America. This work is devoted to the evaluation of earthquake occurrence probabilities in the seismic zones that cover parts of the above region. The evaluation is undertaken by implementing the hidden Markov model, specifically the Poisson HMM (PHMM), which is a dual discrete-time stochastic process. The parameters of interest of the underlying Poisson distributions which represent the hidden states of the process are estimated and the PHMM methodology is implemented. A comparative analysis shows that the classical Poisson model is preferred in only 1 out of 10 seismic zones with the PHMM prevailing in all others. The steady state probabilities are also furnished.

Keywords: earthquake hazard, earthquake occurrence probability, stochastic models, Poisson procedure, hidden Markov process, EM algorithm

***Acknowledgement*:** This work has been partly supported by the University of Piraeus Research Center.

This is a joint work with E. Georgakopoulou, T.M. Tsapanos, E. Scordilis and A. Papadopoulou (Univ. of Thessaloniki), V. Karastathis (National Observatory of Athens) & A. Makridis (Univ. of the Aegean).

Extending Grey-Markov Forecasting through Second-Order Transition Dynamics for Energy Mix Prediction

Vasilis Kalaitzis[1], Alexandra Papadopoulou[1], Andreas Makrides[2], Alex Karagrigoriou[3], Nikolaos G. Galanopoulos[2]

[1]Aristotle University of Thessaloniki, [2] University of the Aegean, Greece, [3] University of Piraeus.

vkalaitd@math.auth.gr, apapado@math.auth.gr, amakridis@aegean.gr, alex.karagrigoriou@unipi.gr, ngalan@aegean.gr

Abstract. Grey forecasting models are particularly suitable for energy applications characterized by limited observations and incomplete information. The integration of GM(1,1) with Markov-chain residual correction has demonstrated improved predictive performance by accounting for stochastic fluctuations around the deterministic Grey trend. An extension of the classical Grey-Markov framework by introducing a second-order Markov correction mechanism is proposed. The resulting higher-order stochastic structure is expected to capture residual-state evolution more accurately, particularly in applications involving long-term forecasting horizons. The methodology is applied to historical primary energy consumption data for Italy obtained from the Our World in Data database. Residuals generated by the GM(1,1) model are partitioned into representative error states and used to construct both first-order and second-order transition probability structures. Particular attention is devoted to addressing the sparsity issues arising from the enlarged state space associated with higher-order Markov dynamics. The proposed framework aims to improve forecast accuracy in long-term energy forecasting.

References

[1] Deng, J. L. (1982). Control problems of grey systems, *Systems & Control Letters* **1**(5), 288–294.

[2] Deng, J. L. (1989) Introduction to Grey system theory, *The Journal of Grey System* **1**(1), 1–24.

[3] D'Amico, G., Karagrigoriou, A., Veronica V. (2024). Forecasting the Power Generation Mix in Italy Based on Grey Markov Models, *Energies* **17** (9), 2184, *https://doi.org/10.3390/en17092184*.

Assesing topic changes using Latent Dirichlet Allocation

Ana Colubi[1] & Louisa Kontoghiorghes[1],

[1]King's College London, UK

ana.colubi@kcl.ac.uk; louisa.1.kontoghiorghes@kcl.ac.uk;

Abstract. The statistical comparison of topic prevalence across multiple corpora is considered. After a brief review of topic modelling techniques, a two-sample bootstrap procedure based on Latent Dirichlet Allocation (LDA) is presented for testing differences in topic prevalence between corpora. A time-varying weighted Latent Dirichlet Allocation (tvwLDA) framework is then described for analysing the evolution of topics in a sequence of corpora. The methodology combines LDA with rolling-window estimation and allows the number of topics to vary over time, accommodating the emergence and disappearance of themes. To characterize these dynamics, a topic dominance index is introduced. Empirical examples illustrate the proposed methods.

Repeatedly spatially indexed multivariate functional data: an application to prostate cancer identification

Mohammad Samsul Alam[1] & **Ana-Maria Staicu**[2]

[1]Duke University, School of Medicine, USA, [2]North Carolina State University, USA.

mohammadsamsul.alam@duke.edu; astaicu@ncsu.edu

Abstract. A common practice in analyzing large input images is to split them into smaller patches for training statistical models. However, the resulting patches may exhibit correlations due to this splitting. Most existing approaches, such as deep learning methods based on convolutional neural network (CNN) architectures, ignore these correlations, which may lead to misleading results. Additionally, these methods often involve a large number of tuning parameters and require substantial training samples to achieve desirable performance. This paper addresses a supervised classification problem aimed at automatic cancer identification in histopathology images, which are typically high-dimensional. We propose a framework that transforms images into spatially indexed multivariate functional data, extracts multiple spatial features, and trains a classifier in the final step. Our approach accounts for both inter- and intra-patch correlations and involves only a few tuning parameters. By leveraging domain knowledge to define functional variables during data transformation and integrating functional data methods with off-the-shelf classifiers during modeling, our framework achieves robust performance even with smaller sample sizes. Simulation studies demonstrate that the proposed method performs well under partially misspecified models. When applied to prostate cancer identification from histology images, the proposed method shows significant improvement over popular competitors.

References

[1] Bulten, W., Bándi, P., Hoven, J., Loo, R.V.D., Lotz, J., Weiss, N., Laak, J.V.D., Ginneken, B.V., Hulsbergen-van de Kaa, C. and Litjens, G. (2019). Epithelium segmentation using deep learning in H&E-stained prostate specimens with immunohistochemistry as reference standard. *Scientific reports*, **9**, 864.

Tensor time series change-point detection in cryptocurrency network data

Andreas Anastasiou[1], Ivor Cribben[2]

[1]University of Cyprus, Cyprus, [2]University of Alberta, Canada.

anastasiou.andreas@ucy.ac.cy; cribben@ualberta.ca

Abstract. Financial fraud has been growing exponentially in recent years. The rise of cryptocurrencies as an investment asset has seen a parallel growth in cryptocurrency scams. To detect possible cryptocurrency fraud, in particular market manipulation, previous research focused on the detection of changes in the network of trades; however, market manipulators are now trading across multiple cryptocurrency platforms, making their detection more difficult. Hence, it is important to consider the identification of changes across a 'network of networks' over time. To this end, in this talk, we propose a new change-point detection method, labeled TenSeg, in the network structure of tensor-variate data. TenSeg first employs a tensor decomposition, and, second detects multiple change-points in the second-order (cross-covariance or network) structure of the decomposed data. It allows for change-point detection in the presence of frequent changes of possibly small magnitudes and is computationally fast. We apply our method, which exhibits great performance, to several simulated datasets and to a cryptocurrency dataset, which consists of network tensor-variate data from the Ethereum blockchain.

References

[1] Anastasiou, A. & Cribben, I. (2026). Tensor time series change-point detection in cryptocurrency network data, *Journal of the Royal Statistical Society Series A: Statistics in Society* 189 (2), 723–743.

Discrete-Time Semi-Markov Modelling of Residual Dynamics in Grey GM(1,1) Forecasting

Nikolaos G. Galanopoulos[1], Alex Karagrigoriou[2] & Andreas Makrides[1]

[1] University of the Aegean, Greece, [2] University of Piraeus.

ngalan@aegean.gr, alex.karagrigoriou@unipi.gr, amakridis@aegean.gr

Abstract. A novel extension of the Grey Model GM(1,1) based on the discrete time semi-Markov process in proposed. Although the conventional combination of Grey model GM(1,1) and Markov chain, termed as Grey-Markov model, can increase the forecasting accuracy owing to the ability to account for residual dynamics by means of first-order Markov chains, it assumes the absence of memory. The proposed model makes no such assumption; instead, it introduces duration dependence of the transitions between states using Markov chain and the corresponding sojourn time distributions in each state. The proposed methodology is examined by comparing it to traditional GM(1,1) and Grey-Markov models, while it provides a flexible theoretical improvement to Grey System Modeling.

References

[1] Barbu, V., Karagrigoriou, A., Makrides, A. (2017). Semi-Markov modelling for multi-state systems, *Meth. & Comput. Appl. Prob.*, **19**, 1011-1028.

[2] D'Amico, G., Karagrigoriou, A., & Vigna, V. (2024). Forecasting the Power Generation Mix in Italy Based on Grey Markov Models, *Energies*, **17(9)**, 2184.

[3] Deng, J., L. (1982). Control problems of grey system, *Syst. Control Lett.*, **1**, 288–294.

[4] He, Y. & Bao, Y. (1992). Grey-Markov forecasting model and its application, *Syst. Eng. Theory Pract.*, **9**, 59–63.

Exponentiality Tail Test: A U-statistics Approach

Angeliki Boutsi and Alex Karagrigoriou

University of Piraeus, Greece

aboutsi@unipi.gr ; alex.karagrigoriou@unipi.gr

Abstract An exponential test with emphasis on the tail of the distribution is proposed as a tool for analyzing extreme values and developing hypothesis tests. A statistic based on the theory of U-statistics is introduced, and its approximation and its theoretical properties are examined. The theoretical results are illustrated with applications in risk theory and finance, where accurate tail characterization is of vital importance.

Keywords: Exponential tail test, U-statistics, risk applications.

***Acknowledgement*:** This work has been partly supported by the University of Piraeus Research Center.

On joint modelling of discrete and continuous time series

Anna Nalpantidi[1], **Dimitris Karlis**[1]

[1]Athens University of Economics and Business, Greece.

analpa@aueb.gr; karlis@aueb.gr

Abstract. We propose a general framework for jointly modeling time series of diverse types, including continuous, count, ordinal, and binary data. To define the joint distribution of mixed type time series we are based on copulas. To overcome computational complexity as dimension increases, while allowing for different dependence structure between each pair of time series, we propose approximating the full likelihood of the multivariate model by conditional pairwise likelihood. As estimation is concerned, we propose maximizing the log-likelihood of each bivariate model separately and then combine the estimates to obtain unique estimates. The synthesis can be just a simple averaging or a weighted mean, where weights come from the hessian matrix of each bivariate model. Simulation study has shown that asymptotically the model behaves well, while no essential differences are detected between the two methods of synthesis. The methodology is presented through a real data example, where issues related to model's assessment and forecasting are also discussed.

References

[1] Fieuws, S., & Verbeke, G. (2006). Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics*, *62*(2), 424-431.

[2] Song, P. X. K., Li, M., & Yuan, Y. (2009). Joint regression analysis of correlated data using Gaussian copulas. *Biometrics*, *65*(1), 60-68.

[3] Hui, F. K., Müller, S., & Welsh, A. H. (2018). Sparse pairwise likelihood estimation for multivariate longitudinal mixed models. *Journal of the American Statistical Association*, *113*(524), 1759-1769.

Some results on random distortions and hazard-operator families

Antonio Di Crescenzo[1]

[1]University of Salerno, Italy.

adicrescenzo@unisa.it

Abstract. A deterministic distortion function d : [0,1] → [0,1] is a non- decreasing function satisfying d(0) = 0 and d(1) = 1. Distortion functions play an important role in several applications, as they allow probability distributions to be transformed in predefined ways. Since in some cases deterministic distortion functions are too rigid to describe the behavior of phenomena in ambiguous or fluctuating contexts, extensions based on randomness have recently been proposed.

Along this line, we propose a class of random distortions obtained by acting on the baseline cumulative hazard through a family of random hazard operators. We first focus on some properties for such operator family. We then consider the associated random quantiles. In conclusion, we also address order-preserving results for distorted lifetimes under various comparison schemes.

These results are based on a recent joint paper with Alessandra Meoli and Alfonso Suárez-Llorens.

References

[1] Balbás, A., Garrido, J. & Mayoral, S. (2009). Properties of distortion risk measures, *Methodology and Computing in Applied Probability*, **11**, 385-399.

[2] Di Crescenzo, A., Meoli, A. & Suárez-Llorens, A. (2026). On a class of random distortions based on hazard-operator families, preprint.

[3] Lin, F., Peng, L., Xie, J. & Yang, J. (2018). Stochastic distortion and its transformed copula, *Insurance: Mathematics and Economics*, **79**, 148-166.

[4] Zang, X., Jiang, F., Xia, C. & Yang, J. (2024). Random distortion risk measures, *Insurance: Mathematics and Economics*, **116**, 51-73.

Bounding Posterior Evidence for Variable Screening

Argyro Damoulaki[1], Ioannis Ntzoufras[1]

[1]Athens University of Economics and Business, Greece.

argyrodam@aueb.gr; ntzoufras@aueb.gr

Abstract. Bayesian model evaluation and selection is a fundamental challenge in statistical analysis, particularly in higher-dimensional settings where the number of possible models grows exponentially with the number of covariates under consideration (p). Posterior model probabilities and inclusion probabilities are typically computed using Bayes factors, which require evaluating 2^p marginal likelihoods. This process often relies on Markov Chain Monte Carlo (MCMC) methods, which can be also computationally demanding. In this work, we examine an alternative approach for variable selection without relying on MCMC or other computationally intensive sampling methods. Instead, we propose an efficient screening approach that reduces model space based on lower and maximum values of posterior inclusion probabilities. Specifically, we explore cases where a conjugate prior structure is assumed while the suggested scalable method accounts for complicated designs such as strong collinearity among predictors.

References

[1] Liang, F., Paulo, R., Molina, G., Clyde, M. A., & Berger, J. O. (2008). Mixtures of g Priors for Bayesian Variable Selection. *Journal of the American Statistical Association*, 103(481), 410–423. https://doi.org/10.1198/016214507000001337

[2] Paroli, R., Fouskakis, D. & Ntzoufras, I. (2025). Fast bayesian variable screening using correlation thresholds. *Stat Comput* 35, 74. https://doi.org/10.1007/s11222-025-10608-8

Kernel Smoothing Techniques for the Transition function of semi-Markov systems

Ayhar Chafiâa[1]

University Center of Nour Bachir Of El Bayadh, Algeria.

ayharchafiaa@yahoo.com

Abstract. This study focuses on kernel-based estimation of the semi-Markov transition function for homogeneous continuous-time semi-Markov processes. The primary objective is to investigate the asymptotic properties of the proposed estimators. To illustrate and validate the theoretical findings, a numerical example is presented

References

[1] C. Ayhar, V.S. Barbu, F. Mokhtaria, & S. Rahmani. (2022). On the asymptotic properties of some kernel estimators for continuous-time semi-Markov processes. *Journal of Nonparametric Statistics*, **34**(1), 1–21.

[2] V.S. Barbu, N. Limnios. (2008). *Semi-Markov chains and hidden semi-Markov models toward applications-their use in reliability and DNA analysis*. Lecture Notes in Statistics. Springer, New York, 191.

[3] F. Mokhtari, C. Ayhar, V. S. Barbu, & S. Rahmani. (2025). Kernel estimators of Markov renewal and semi-Markov transition functions of semi-Markov systems. *Journal of Nonparametric Statistics*, 1-24.

[4] B. Ouhbi and N. Limnios. (1996). Nonparametric estimation for semi-Markov kernels with application to reliability analysis. *Appl. Stoch. Models Data Anal*., **12**, 209-220.

[5] B. Ouhbi and N. Limnios. (1997). Comportement asymptotique de la matrice de renouvellement markoview, C. R. Acad. Sci. Paris. **325** : 921-924

[6] B. Ouhbi and N. Limnios. (1999). Nonparametric estimation for semi-Markov processes based on its hazard rate functions. *Stat. Inference Stoch. Process*., **2**, 151-173.

A Characteristic Function-Based Test for the Multi-Sample Problem under Random Right Censoring

James S. Allison[1], Marija Cuparić[2], Bojana Milošević[2], Anke Steyn[3,4] & I.J.H. Visagie[4]

[1] Department of Statistics, College of Science, Engineering and Technology, University of South Africa, South Africa.

[2] University of Belgrade, Faculty of Mathematics, Serbia

[3] Business Mathematics and Informatics, North-West University, South Africa.

[4] Pure and Applied Analytics, North-West University, South Africa.

allisjs@unisa.ac.za; marija.cuparic@matf.bg.ac.rs; bojana@matf.bg.ac.rs; anke.steyn@nwu.ac.za; jaco.visagie@nwu.ac.za

Abstract. We consider the multi-sample problem in the presence of random right censoring and propose a new class of nonparametric tests based on weighted L^2-distances between empirical characteristic functions constructed from Kaplan–Meier estimators. The methodology is applicable to both two-sample and general multi-sample settings and accommodates several weighting schemes, including Gaussian, Laplace, uniform, and Dirac-type kernels. Asymptotic properties of the proposed statistics are established under the null hypothesis, and bootstrap procedures are used for finite-sample calibration. A simulation study compares the new tests with classical competitors adapted to this framework. The results demonstrate competitive performance and increased sensitivity under complex alternatives, including non-proportional hazards and crossing survival curves.

A nonparametric approach for noisy compositional data

Chiara Passamonti[1], Marco Di Marzio[2], Stefania Fensore[3] & Agnese Panzera[2]

[1]University of Luxembourg, Luxembourg, [2]University of Chieti-Pescara, Italy, [3]University of Florence, Italy.

chiara.passamonti@uni.lu; marco.dimarzio@unich.it; stefania.fensore@unich.it; agnese.panzera@unifi.it

Abstract. Compositional data describe the relative contributions of parts of a whole and are naturally defined on the simplex. In practice, such data, which arise in several scientific fields, are often affected by measurement error, due to instrumental imprecision, survey bias, sensor noise, or inaccuracies hidden by the closure operation. Although the so-called errors-in-variables problem is well studied in Euclidean spaces, it remains unexplored for compositional data.

We propose deconvolution kernel estimators for density and regression functions on the simplex when compositional variables are observed with error. For density estimation, we develop a Mellin-based deconvolution approach and show its equivalence with Fourier deconvolution applied to ILR coordinates. In the regression setting, a nonparametric estimator for a compositional response with either contaminated real-valued or compositional predictors is provided.

We illustrate the proposed methods through numerical experiments on simulated data and real soil-composition observations.

References

[1] Aitchison, J., & Shen, S. M. (1984). Measurement error in compositional data. *Journal of the International Association for Mathematical Geology*, **16**(6), 637-650.

[2] Di Marzio, M., Panzera, A., & Venieri, C. (2015). Non-parametric regression for compositional data. *Statistical Modelling*, **15**(2), 113-133.

[3] Greenacre, M. (2021). Compositional data analysis. *Annual Review of Statistics and Its Application*, **8**(1), 271-299.

Diameter Thresholds of Random Cayley Graphs - I

Demetres Christofides[1], Klas Markström[2] & Christina Savvidou[1]

[1]INSPIRE Research Center and UCLan Cyprus, Cyprus, [2] Umeå University, Sweden

dchristofides@uclan.ac.uk; klas.markstrom@umu.se; csavvidou@uclan.ac.uk

Abstract. For any fixed integer $d \geqslant 2$, the asymptotic threshold for the binomial random graph $\mathcal{G}(n,p)$ to achieve diameter at most d is a classical result in probabilistic combinatorics, exhibiting a sharp phase transition driven by independent edge exposures. While random regular graphs share identical thresholds due to asymptotic universality, this universality breaks down when structural dependencies are introduced. In this first of a two-part talk, we review the classical diameter thresholds for the binomial random graph model along with the core probabilistic techniques behind their proofs. We then introduce the model $\mathcal{G}(G,p)$ of random Cayley graphs, where edges are governed by an underlying group structure. Finally, we highlight the unique probabilistic challenges that arise when dealing with algebraically dependent edge distributions and present our main results establishing new upper and lower bounds for the diameter.

An overview of Principal Component Analysis methods for Directional Data

Christophe Ley[1], Anahita Nodehi[2], & Meisam Moghimbeygi[3]

[1]University of Luxembourg, Luxembourg, [2]University of Bristol, UK, [3] Kharazmi University, Iran.

Christophe.ley@uni.lu; anahita.nodehi@bristol.ac.uk

Abstract. The high dimensionality of the input data can pose multiple problems when implementing statistical techniques. The presence of many dimensions in the data can lead to challenges in visualizing the data, higher computational demands, and a higher probability of over-fitting or under-fitting in modeling. Furthermore, the curse of dimensionality contributes to these issues by stating that the necessary number of observations for accurate modeling increases exponentially as the number of dimensions increases. Dimension reduction tools help overcome this challenge. Principal Component Analysis (PCA) is the most widely used technique, intensively studied in classical linear spaces. However, in applied sciences such as biology, bioinformatics, astronomy and geology, there are many instances in which the data's support are non-Euclidean spaces. In fact, the available data often include elements of Riemannian manifolds such as the unit circle, torus, sphere, and their extensions. Therefore, the terms "manifold-valued" or "directional" data are used in the literature for these situations. When dealing with directional data, the linear nature of PCA might pose a challenge to achieve accurate data reduction. This paper talks presents the methodological aspects of PCA on directional data and their practical applications.

A Refined Approach to Approximating the Mean Number of System Repairs

Christos G. Papadopoulos[1] and Konstandinos Politis[2]

[1] Department of Statistics and Insurance Science, University of Piraeus

cpapadopoulos@unipi.gr & kpolitis@unipi.gr

Abstract. In reliability modeling, an accurate calculation of the renewal function is critical in implementing system maintenance policies and determining the mean numbers of repairs. Asadi (2023) introduced a parametric model for the average number of repairs in a system, that accommodates a broad range of classes for repairs, from better-than-minimal to worse-than-minimal, utilizing a function g. The quality of this approximation heavily depends on the choice of g which remains analytically challenging for most lifetime distributions. In this research, a new approach is presented on the choice of this function g which yields improved numerical precision for some lifetime distributions (e.g. Weibull, Pareto etc.). Numerical experimentation verifies that the new strategy significantly reduces deviation between the theoretical and approximate renewal function, particularly in the intermediate time interval where most of existing methods degrade in accuracy.

Keywords: renewal function, renewal theory, reliability, maintenance

References

[1] Asadi, M. (2023). On a parametric model for the mean number of system repairs with applications. *Reliability Engineering & System Safety, 234*, 109137.

[2] Jiang, R., & Chen, Z. (2020). A practical approximation of Weibull renewal function for solving relevant optimization problems. *Quality Engineering, 32*(3), 333–341.

[3] Lorden, G. (1970). On excess over the boundary. *Annals of Mathematical Statistics, 41*, 520–527.

[4] Papadopoulos, C. G., & Politis, K. (2026). New approximations for the renewal function of the Weibull distribution. *Reliability Engineering & System Safety, 272*, 112425.

[5] Politis, K., & Koutras, M. (2006). Some new bounds for the renewal function. *Probability in the Engineering and Informational Sciences, 20*, 231–250.

[6] Xie, M. (1989). On the solution of renewal-type integral equations. *Communications in Statistics - Simulation and Computation, 18*, 281–293.

Perpetual American Options in a jump-diffusion model under Poisson observation times

Boutsikas M.V.[1], Economides D-J[1]

[1]Dept. of Statistics & Insurance Science, Univ. of Piraeus, Greece

mbouts@unipi.gr; davoik@unipi.gr

Abstract. We investigate the problem of pricing Perpetual American put and call options, considering the exercise opportunity occurring under Poissonian inspection times. Employing an exponential Levy model, we specifically examine a jump diffusion with two-sided jumps. We present key identities to express the options payoff and exercise probability in terms of undershoot/overshoot for put/call scenarios. Additionally, we derive explicit formulas for options payoff, exercise probability, and optimal boundary under double exponential jump diffusion process. Valuation under the risk-neutral probability measure is provided, along with asymptotic formulas yielding well-established results from continuous time model. Our numerical examples illustrate the impact of inspection intensity sensitivity on pricing outcomes.

References

[1] Dupuis P., and Wang H. (2002) Optimal stopping with random intervention times. *Advances in Applied Probability*, **34**(1), 141–157.

[2] Gerber H.U. and Shiu E.S.W. (1994) Option pricing by Esscher transforms. *Trans. Soc. Actuaries* **46**, 99–191.

[3] Kou S.G. and Wang H. (2004) Option pricing under a double exponential jump diffusion model. *Management Science*, **50**, 1178–1192.

[4] Mordecki E. (2002) Optimal stopping and perpetual options for Lévy processes. *Finance Stochastics* **6**, 273–293.

[5] Pérez J.L. and Yamazaki K. (2018) American options under periodic exercise opportunities. *Statistics and Probability Letters*, **135**, 92–101.

Modelling Structural Changes in Random-Effect Variability for Short Longitudinal Panels

Dário Ferreira[1,2] & **Sandra S. Ferreira**[1,2]

[1]University of Beira Interior, Portugal, University Mathematics Department and [2]Instituto de Telecomunicações.

dario@ubi.pt; sandra@ubi.pt;

Abstract. Standard mixed-effects models assume constant covariance in individual heterogeneity, which can be restrictive in short panels affected by structural events. We propose a mixed-effects framework that allows random-effect variability to change after a known or unknown break point by introducing post-event latent components. Estimation relies on a quasi-REML approach, and identifiability conditions are established to distinguish baseline from post-break heterogeneity. A screening procedure identifies candidate break dates, and inference is based on a bootstrap test for covariance stability. Simulations show that modeling covariance changes improves variance estimation and helps separate true heterogeneity shifts from residual noise. The approach is particularly useful in longitudinal settings where events alter the dispersion and correlation of individual responses.

Funding.This work is funded by national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., and, when eligible, co-funded by EU funds under project/support UID/50008/2025 - Instituto de Telecomunicações, https://doi.org/10.54499/UID/50008/2025.

References

[1] Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes, *Econometrica*, **66**, 47–78.

[2] Ferreira, D. & Ferreira, S. (2025). CMMSE: A time-varying approach to linear mixed models, Mathematical Methods in the Applied Sciences, **48**, 9562–9568.

[3] Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems, *Journal of the American Statistical Association*, 72, 320–338.

Stopping on the last success with unknown odds

Davy Paindaveine[1]

[1]Université libre de Bruxelles, Belgium.

Davy.Paindaveine@ulb.be

Abstract. Optimal stopping problems lie at the interface of probability, statistics, and sequential decision theory. They model situations in which decisions must be made online, using only the information revealed so far and without access to future outcomes. We will focus on the classical last-success problem: one observes a sequence of Bernoulli trials and seeks to stop exactly on the final success. When the success probabilities are known, the problem admits an elegant and complete solution via Bruss' sum-the-odds theorem (Bruss, 2000, Annals of Probability), which yields a simple threshold rule and an explicit formula for the optimal win probability. In most applications, however, these probabilities are unknown, so the oracle rule cannot be implemented. We therefore study the last-success problem under a minimal information structure, where the decision maker observes only the sequential Bernoulli outcomes. Our contribution provides a quantitative theory of oracle-freeness: we analyze the performance of the natural plug-in odds rule and, more importantly, characterize what is possible---and what is fundamentally impossible---when the success probabilities are unknown.

References

[1] Bruss, F. T. (2000). Sum the odds to one and stop, *The Annals of Probability*, **28**, 1384--1391.

Diameter Thresholds of Random Cayley Graphs - II

Demetres Christofides[1], Klas Markström[2] & Christina Savvidou[1]

[1]INSPIRE Research Center and UCLan Cyprus, Cyprus, [2] Umeå University, Sweden

dchristofides@uclan.ac.uk; klas.markstrom@umu.se; csavvidou@uclan.ac.uk

Abstract. In the random Cayley graph model $\mathcal{G}(G,p)$, edges are determined by choosing elements from an underlying group G independently at random with probability p. Unlike classical independent models like the Erdős-Rényi random graph, the geometry of $\mathcal{G}(G,p)$ is heavily constrained by the group's algebraic structure. This creates complex structural dependencies and forced path intersections, causing a breakdown of the standard universality constants typically observed in diameter phase transitions. In this second of a two-part talk, we focus on the probabilistic machinery and core proof strategies developed to determine these diameter thresholds for both constant and growing diameters d. We will first provide the combinatorial intuition behind why different group families shift the threshold constants by factors related to commutativity or the density of involutions. We then present our unified framework designed to overcome these algebraic dependencies, highlighting an optimized dependency graph construction and a refined application of Janson's inequality over intersecting, non-independent paths. While complementing the first of a two-part talk, this talk is self-contained and emphasizes modern probabilistic tools for constrained random structures.

Some spatio-temporal models for count data

Dimitrios Karlis

Athens University of Economics and Business, Greece.

karlis@aueb.gr

Abstract. Consider spatio-temporal data that are counts. Such data are common in epidemiology, where, for example disease occurrences are observed across regions over time, criminology, where crime counts are analyzed within specific jurisdictions across successive periods, and many other disciplines. This talk introduces a class of models designed to account for the discrete nature of these observations explicitly. Such models offer better forecasting accuracy and they can provide better insights compared to existing continuous approximations. We define a general spatio-temporal framework based on the idea of thinning operation that preserves the discrete nature of the data. We also derive its fundamental statistical properties. For parameter estimation, we employ a conditional maximum likelihood approach facilitated by an Expectation-Maximization (EM) algorithm. Furthermore, we discuss the development of forecasting procedures and validate the models performance through extensive Monte Carlo simulation experiments. The practical utility of the proposed framework is demonstrated using a real-world application: the spatio-temporal progression of the COVID-19 epidemic in the Netherlands. Possible extensions are also discussed.

References

[1] Karlis, D., Chutoo, A., Mamode Khan, N., & Jowaheer, V. (2024). The multilateral spatial integer valued process of order 1. *Statistica Neerlandica*, **78**(1), 4–24

Detecting Bilinear Dependence in Financial Panel Data: A Rank-Based Approach

Dounia Bourzik[1], Aziz Lmakri[2], Hassan Oukhouya[3], Amal Mellouk[4] & Abdelhadi Akharif[5]

[1]LMA, FSTT, Abdelmalek Essaadi University, Tetouan, Morocco, [2]AIMCE Laboratory, ENSAM, Hassan II University, Casablanca, Morocco, [3]LaMSD, MSASE, FSJES, Mohammed First University of Oujda, BV Mohammed VI, B.P. 724, Oujda, 60000, Morocco, [4]Regional Center for Education and Training Trades, Tangier, Morocco, [5]LMA, FSTT, Abdelmalek Essaadi University, Tetouan, Morocco.

dounia.bourzik@etu.uae.ac.ma; lmakri.aziz@ensam-casa.ma; oukhouya.hassan@ump.ac.ma; prmelloukamal@gmail.com; a.akharif@uae.ac.ma

Abstract. This paper examines superdiagonal bilinear dependence in short panel data structures, characterized by a large cross-sectional dimension and a limited time dimension. Our main objective is to develop efficient and robust testing procedures for detecting this form of nonlinear dependence. We first establish a local asymptotic normality (LAN) property for the log-likelihood ratio within a broad class of bilinear models. Based on this asymptotic framework, we derive optimal parametric tests and rank-based nonparametric procedures. We investigate their asymptotic properties, including local power and asymptotic relative efficiency with respect to pseudo-Gaussian benchmarks. Extensive Monte Carlo simulations are conducted to evaluate finite-sample size, power, and robustness under various distributional settings. The results show that rank-based methods maintain accurate size control and strong power, particularly in the presence of non-Gaussian innovations. An empirical application to financial panel data further supports the theoretical findings and highlights the practical relevance of the proposed methodology.

References

[1] Granger, C.W.J. & Andersen, A.P. (1978). *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Göttingen.

[2] Lmakri, A., Akharif, A., Mellouk, A. & Fihri, M. (2021). Pseudo-Gaussian and Rank-Based Optimal Tests for First-Order Superdiagonal Bilinear Models in Panel Data, *REVSTAT-Statistical Journal*, **19(3)**, 443–462.

Measuring Population Complexity through entropy measures

Cipu Elena Corina[1,2,4], Preda Vasile[3,4,5], Vasile Nicolae Stănciulescu[4]

[1]Centre for Research and Training in Innovative Techniques of Applied Mathematics in Engineering "Traian Lalescu" (CiTi), UNSTPB, Romania
[2]Department of Applied Mathematics, Faculty of Applied Sciences, UNSTPB
[3]"Gheorghe Mihoc – Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy (ISMMA)
[4]V. Trebici" Center for Demographic Research, "Costin C. Kiritescu" National Institute of Economic Research of Romanian Academy, Romania
[5]Faculty of Mathematics, University of Bucharest, Romania
corina.cipu@upb.ro; vasilepreda0@gmail.com; vasile.nicolae.stanciulescu@gmail.com

Abstract. Entropy measures provide powerful quantitative tools to assess the diversity, divergence, or inequality inherent in the demographic and economic implications of a given population. The purpose of this paper is to analyses tensions and imbalances between generations by a measuring the complexity of a population based on social cohesion and demographic resilience, building composite indicators such as Social Cohesion Index and Generational Distress index. Distances between different economic, spatial, and social data are defined by calculating Shannon, Rao, and Kullback -Leibler entropy.
Comparisons and analyses are performed based on these indicators and metrics, using demographic data from Romania and across Europe, obtaining useful classifications related to human capital assessments or regional development strategies.

Keywords: *population dynamics, regional demography, Rao entropy, Kullback-Leibler entropy*

References

[1] Z. Qun, Y. Menghui, (2023) An efficient entropy of sum approach for measuring diversity and interdisciplinarity, *Journal of Informetrics* 17 101425, Doi: 10.1016/j.joi.2023.101425

[2] C. Ricotta, L. Szeidl (2006). Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index, *Theoretical Population Biology,* **70** 237–243, doi: 10.1016/j.tpb.2006.06.003

[3] N. Kumar, A. Dixit, V. Vijay (2025). Entropy measures and their applications: A comprehensive review, DOI:10.48550/arXiv.2503.15660

Mixture Kernels for Non-Stationary Time Series

Emmanouil-Nektarios Kalligeris[1], Alex Karagrigoriou[1] & Kostas Triantafyllopoulos[2]

[1]Department of Statistics and Insurance Science, University of Piraeus, Piraeus, Greece,
[2]School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, UK
ekalligeres@gmail.com; alex.karagrigoriou@unipi.gr; k.triantafyllopoulos@sheffield.ac.uk

Abstract. In this work we seek to bridge the gap between traditional stationary GP models and the dynamic needs of complex time-series data. By incorporating a novel non-stationary kernel the proposed unified framework addresses the limitations of traditional GPs and offers improved predictions and deeper insights into dynamic processes, supporting adaptive management and informed decision-making.

References

[1] Conti, S., Gosling, J. P., Oakley, J. E. & O'Hagan, A. (2009). Gaussian process emulation of dynamic computer codes, *Biometrika* **96**, 663-676.

Sharp Stein factors for integrated Pearson distributions through kernel representations

Ferdinand Rapin[1], Yvik Swan[1,2]

[1]Université Libre de Bruxelles, Belgium, [2]Vrij Universiteit Brussel, Belgium,

Ferdinand.rapin@ulb.be; yvik.swan@ulb.be

Abstract. Stein's method is a powerful technique to assess distances between probability distributions. One of the main tools of this method is so-called Stein factors, that is uniform bounds on the solution to a distribution-characterizing differential equation called the Stein equation. In this work, we derive new representations of the solution to the Stein equation and use it to provide explicit conditions on the distribution to obtain tractable Stein factors for the solution and its derivatives of any order. The conditions naturally apply to integrated Pearson distributions and lead to optimal uniform Stein factors.

Applications includes new theoretical guarantees for Satterthwaite's approximation, rates of convergence to the gamma-distribution for a sequence of r.v. whose *k*-th first moments match the gamma and improved bound on the Beta approximation of the number of white ball drawn in a Pólya-Eggenberger urn.

References

[1] Bailly, G., Rapin, F., Swan, Y., Von Sachs, R. (2026). Explicit Gamma-Stein bounds for Satterthwaite's approximation, *to appear in ESAIM:P&S*

[2] Gaunt, R. E., Pickett, A. M., and Reinert, G. (2017). Chi-square approximation by Stein's method with application to Pearson's statistic. *The Annals of Applied Probability 27.*

[3] Afendras, G., Papadatos, N. (2015). Integrated Pearson family and orthogonality of the Rodrigues polynomials: A review including new results and an alternative classification of the Pearson system, *Applicationes Mathematicae 42*

Stochastic Reliability Indicators for Wind Energy Systems: A Markov and Semi-Markov Perspective

Guglielmo D'Amico, Filippo Petroni

University "G. d'Annunzio", Italy

g.damico@unich.it; filippo.petroni@unich.it

Abstract. We propose a stochastic framework for the analysis of reliability in wind energy systems based on finite-state Markov and semi-Markov processes. Within this setting, the system dynamics are described through transition mechanisms between working and failure states, allowing the definition of time-dependent reliability indicators. Beyond classical measures such as the Rate of Occurrence of Failures (ROCOF), we consider additional rate-based functionals capturing repair dynamics and state persistence, leading to a more complete characterization of system behavior. The approach provides a unified probabilistic structure for deriving and interpreting these indicators, with potential applications to performance assessment and decision-making in wind farm management.

Information theory and spatial risk analysis

Guglielmo D'Amico

[1]Department of Economics, University G. d'Annunzio of Chieti-Pescara, Viale Pindaro, 42, Pescara, 65127, Italy,

g.damico@unich.it

Abstract. In this paper, we extend Batty's entropy by introducing a coherent spatially weighted entropy indicator. We study its analytical properties, optimization behavior, and links to established entropy and divergence measures. The effects of affine transformations of damages are analyzed through a dynamical systems perspective. Numerical examples illustrate the proposed framework, highlighting its relevance for spatial risk assessment and comparison.

References

[1] M. Batty, Spatial entropy, Geographical analysis 6 (1974) 1–31.

[2] G. D'Amico, S. Scocchera, L. Storchi, Randentropy: a software to measure inequality in random systems, Informatica 33, (2022) 279–298.

[3] T. Gkelsinis, A. Karagrigoriou, Theoretical aspects on measures of directed information with simulations, Mathematics 8, (2020) 587.

Classification of Climate Change Adaptation Plans (CCAP/PACC) for New Brunswick Municipalities using Machine Learning, Deep Learning, and Generative AI Models

Hassan Oukhouya[1], Mohamed Amazigh[2], Cherifa Fatima Choukhan[3], Salah Eddine El Adlouni[2], Martin Simon[4] & Solym Manou-abi[5]

[1]University of Mohammed First, Morocco, [2]University of Moncton, Canada, [3]University of Mohammed V, Morocco, [4]University of Applied Sciences, Germany, [5]University of Poitiers, France.

Oukhouya.hassan@ump.ac.ma; ema1687@umoncton.ca; cherifafatima_choukhan@um5.ac.ma; salah-eddine.el.adlouni@umoncton.ca; martin.simon@fra-uas.de; solym.manou.abi@univ-poitiers.fr

Abstract. This study investigates the supervised classification of climate change adaptation planning among 77 municipalities in New Brunswick, Canada. The objective is to predict municipalities' adaptation status—No Plan, Partial Plan, or Whole Territory—using geographic, administrative, and temporal characteristics. Following the CRISP-DM framework, the analysis combines exploratory data analysis, unsupervised clustering (K-Means, HAC, DBSCAN), autoencoder-based dimensionality reduction, and a benchmark of twelve machine learning classifiers. Model interpretability is assessed using SHAP values. Results show that Logistic Regression and k-Nearest Neighbours achieve the best predictive performance, reaching 81.25% accuracy, while k-NN attains the highest macro-AUC (0.8986). Ensemble tree-based methods exhibit substantial overfitting due to the limited sample size. Autoencoder-based compression slightly reduces classification performance, suggesting that information loss outweighs potential denoising benefits in small tabular datasets. SHAP analysis identifies geographic area, municipality type, and local governance status as the most influential predictors of adaptation plan adoption. These findings demonstrate the potential of interpretable machine learning methods for assessing municipal climate adaptation readiness and supporting evidence-based policy development.

References

[1] Ladi, T., Jabalameli, S., & Sharifi, A. (2022). Applications of machine learning and deep learning methods for climate change mitigation and adaptation. Environment and Planning B: Urban Analytics and City Science, 49(4), 1314-1330.

Sequential Model Selection for Nonparametric Autoregressive Processes

Ouerdia Arkoun[1], J.-Yves Brua[2] & Serge Pergamenchtchikov[2]

[1]Laboratoire BIRL, Sup'Biotech and Laboratoire de Mathématiques Raphaël Salem, Université de Rouen Normandie

[2]Laboratoire de Mathématiques Raphaël Salem, Université de Rouen Normandie

ouerdia.arkoun@supbiotech.fr; Jean-Yves.brua@univ-rouen.fr;
serge.pergamenchtchikov@univ-rouen.fr

Abstract. In this talk, we study the problem of nonparametric estimation of a function in an autoregressive model under quadratic risk. To this end, we develop an adaptive sequential model selection method based on efficient sequential kernel estimators proposed by Arkoun and Pergamenshchikov (2016). Through the sequential kernel estimators, one passes from an autoregressive model to a regression model, in which the noises are conditionally independent random variables. Then, in the framework of this regression model, the unknown function is estimated on the basis of the model selection procedure based on the family of the weighted least square estimators with the weights defined through the Pinsker efficient method. Then for this procedure, one derives sharp non-asymptotic oracle inequalities for the robust quadratic risk. This inequality means that the proposed estimation procedure is quasi-optimal in the non-asymptotic sense, i.e. the quadratic risk for the constructed procedure is minimal up to some coefficient close to one among all weighted least square estimators with the coefficients of Pinsker form, which under known regularity of the unknown function provide the efficient estimation in the asymptotic setting.

References

[1] Arkoun, O., Brua, J.-Y. and Pergamenchtchikov, S. (2024): Adaptive efficient robust sequential analysis for autoregressive big data models. Sequential Anal. 43 (4), 432-460.

[2] Arkoun, O. and Pergamenchtchikov, S. (2016): Sequential robust estimation for nonparametric autoregressive models. Sequential Anal. 35 (4), 489-515.

How to Analyse Survival Data with AI-informed Non-linear Cox Model

Katarzyna Szczerba[1,2], Laurent Malisoux[3], Christophe Ley[1]

[1]University of Luxembourg, Luxembourg, [2] Information Technology for Translational Medicine (ITTM), Luxembourg [3]Physical Activity, Sport and Health, Luxembourg Institute of Health, Luxembourg

katarzyna.szczerba@uni.lu; laurent.malisoux@lih.lu; christophe.ley@uni.lu

Abstract. The Cox proportional hazards model is the most commonly used method for multivariate survival analysis. Despite its many advantages, such as simplicity and interpretability, it has a serious drawback: it fails to capture non-linear relationships. In this study, we propose AI-informed Non-linear Cox Model, a method that uses insights from a highly predictive machine learning model, extracted with an interpretable machine learning tool, to integrate non-linear relationships into the traditional Cox model via means of splines. On simulated data with a deliberately introduced non-monotonic relationship between the predictor and the outcome variable, the AI-informed Cox model outperformed the traditional proportional hazards (PH) Cox model. Its concordance index (C-index) was also comparable to that of the best-performing machine learning model - gradient boosted Cox model. Similar results were observed when the models were applied to a prospective dataset in running.

References

[1] Chen, Y., Jia, Z., Mercola, D. & Xie, X. (2013). A gradient boosting algorithm for survival analysis via direct optimization of concordance index, Computational and Mathematical Methods in Medicine, Article ID 873595.

[2] Moncada-Torres, A., van Maaren, M. C., Hendriks, M. P., et al. (2021). Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival, Scientific Reports, 11, 6968.

[3] Felice, F., Ley, C., Bordas, S. P. & Groll, A. (2025). Boosting any learning algorithm with Statistically Enhanced Learning, Scientific Reports, 15, 1605.

Inference and Applications of Multivariate Integer-Valued Time Series

Konstantinos Fokianos

University of Cyprus, Cyprus

fokianos@ucy.ac.cy

Abstract. We study inference and modelling techniques for multivariate count time series. In particular, we focus on autoregressive, network, and space–time models. The development of both the underlying theory and its applications depends critically on the choice of an appropriate response distribution for multivariate count data, as well as on the framework of observation-driven models. Our aim is to highlight recent methodological advances in this area and to outline several directions for future research.

Penalized sequential estimation for recurrent events with high-dimensional covariates

Laura Dumitrescu[1] & **Ioana Schiopu-Kratina**[2]

[1]Fairfield University, United States, [2]University of Ottawa, Canada.

ldumitrescu@fairfield.edu; ioanakratina@gmail.com

Abstract. The study of recurrent events, such as hospital readmissions or stock market volatility spikes, is important in biostatistics and finance but poses statistical challenges due to complex dependencies and censoring. Under a semiparametric framework for modeling gap times between events, we employ a penalized sequential estimating equations approach to handle high-dimensional covariates. The procedure yields estimators with attractive large sample properties and we illustrate how the rate of divergence in the number of parameters affects the convergence rate of penalized estimators when the number of covariates grows to infinity as the number of clusters increases. The proposed method has the advantage of simultaneously estimating parameters while selecting important variables and accounting for the within-cluster correlation.

References

[1] Dumitrescu, L. and Schiopu-Kratina, I. (2026). Asymptotics for arrays of martingale differences in recurrent event analysis. *Sequential Analysis*, **45**(2), 304–320.

Fast Matrix-Convolution Methods for Semi-Markov Modelling

Leonidas Kordalis[1], Samis Trevezas[1,2]

[1]Department of Mathematics, National and Kapodistrian University of Athens, Greece [2]MICS Laboratory, CentraleSupélec - Université Paris-Saclay, France,

lkordali@math.uoa.gr; strevezas@math.uoa.gr

Abstract. Discrete-time matrix convolution products play a fundamental role in semi-Markov modelling, since the associated distributions as well as the Markov renewal equations can be transformed into convolution-type equations. This work concerns the algebraic structure of matrix convolutions, obtaining explicit representations for the convolutional inverse. Furthermore, the computational complexity is significantly reduced by adapting FFT-accelerated convolution techniques and Newton-type inversion schemes, combined with Gauss–Jordan methods in truncated power-series rings. This framework is naturally applied to discretization approximations of matrix Stieltjes convolutions arising in continuous-time semi-Markov models and is particularly efficient in the presence of heavy-tailed distributions.

References

[1] Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. Mathematics of Computation, 19(90), 297–301..

[2] V. S. Barbu and N. Limnios. Semi-Markov chains and hidden semi-Markov models toward applications: their use in reliability and DNA analysis, volume 191. Springer Science & Business Media, 2009.

[3] Xie, M., Preuss, W., & Cui, L.-R. (2003). Error analysis of some integration procedures for renewal equation and convolution integrals. Journal of Statistical Computation and Simulation, 73(1), 59-70.

Optimal stopping valuation of modular advertising strategies under copula-based markov dependence

Lucianna Cananà[1], Guglielmo D'Amico[2], Salvatore Vergine[3]

[1] Ionic Department in Legal and Economic Systems of the Mediterranean: Society, Environment, Culture, University of Bari Aldo Moro, Piazza Umberto I, Taranto, 70121, Italy,

[2] Department of Economics, University G. d'Annunzio of Chieti-Pescara, Viale Pindaro, 42, Pescara, 65127, Italy,

[3]Department of Management, Marche Polytechnic University, Piazzale R. Martelli, 8, Ancona, 60121, Italy.

lucianna.canana@uniba.it; g.damico@unich.it; s.vergine@staff.univpm.it;

Abstract. In this paper, we estimate the value of a multi-channel television advertising contract between a firm and a broadcasting network. We propose a real-option framework for allocating advertisements across channels under copula-based audience dependence. Viewer evolution is modeled via a Markov reward process, which is subsequently transformed into revenue streams by means of a suitable payoff function. The analysis is extended beyond the standard two-channel setting by introducing a hierarchical macro/micro-class structure for modeling audience heterogeneity and cross-channel interactions. We compute moments of discounted payoffs to recover the payoff distribution and determine the option value. We also incorporate modular advertising strategies based on channel and audience segmentation. The decision problem is formulated as an optimal stopping problem for determining when to launch the advertising campaign. Finally, the methodology is illustrated using real television audience data.

References

[1] L. Cananà et al., Valuing advertising investment with modular strategies under copula-based dependence, Quality & Quantity, 2025. Springer. van de Geer, S. (2000). *Empirical Processes in M-estimation*, Cambridge University Press, London, UK.

[2] G D'Amico, R De Blasis, V Vigna Advertising investments on television: real option estimation through Markov chains, Quality & Quantity, 2024 , Springer.

When Is an Observation Truly an Outlier? A Probability-Coordinate Perspective

Manuela-Simona Cojocea[1]

[1]University of Bucharest, Romania

simona.cojocea@fmi.unibuc.ro

Abstract. Outlier detection is one of the oldest problems in statistics, yet what counts as an outlier often depends on how the data are represented. An observation that looks extreme under one model may seem perfectly ordinary under another. This raises a fundamental question: when is an observation truly an outlier? We approach this question through the lens of probability coordinates. Given a probability chart G, observations are mapped to their probability positions $U = G(X)$, shifting the study of outliers from value space to probability space. In this setting, unusual observations become associated with proximity to the boundaries of the unit interval, giving outlyingness a natural geometric interpretation. Rather than relying on a single model, we study observations across multiple probability charts. This leads to notions of outlier stability and chart agreement, measuring how robust an observation's classification is to changes in probability representation. The resulting perspective distinguishes observations that are merely model-dependent anomalies from those that occupy persistently exceptional positions across different probability geometries. Examples involving skewed, heavy-tailed, and contaminated distributions illustrate how this perspective reveals structural features that traditional approaches can miss. More broadly, the framework suggests that outliers may be viewed as geometric objects rather than merely value-space anomalies.

A Divergence-Based Framework for Statistical Modeling and Reliability Analysis

Maria Kateri

RWTH Aachen University, Germany

maria.kateri@rwth-aachen.de

Abstract. Shannon entropy and relative entropy, commonly known as the Kullback–Leibler (KL) divergence, are fundamental tools in statistical inference, with established applications in estimation and goodness-of-fit testing. Beyond these classical roles, divergences offer a powerful framework for constructing and analyzing statistical models. In particular, models such as association models for contingency tables and the Mallows model for ranking data can be characterized through divergence-based structural properties, yielding deeper insight and alternative interpretations.The KL divergence belongs to the class of φ-divergences, which naturally gives rise to minimum φ-divergence estimation as an alternative to maximum likelihood estimation. This yields a family of estimators that may improve robustness while preserving desirable statistical properties. Furthermore, many standard statistical models can be reformulated within an information-theoretic framework based on divergences, enabling both new structural insights and extensions to flexible model classes. Specifically, attention is given to the Cressie–Read (CR) divergence family, indexed by a single parameter, which allows the construction of flexible parametric families of statistical models that encompass classical models as special cases. This parameter may be fixed or estimated from the data, providing additional modeling flexibility. Divergence-based ideas also extend to Bayesian computation. In this context, coordinate ascent variational inference (CAVI) for logistic regression, traditionally based on the reverse KL divergence, is generalized using the CR-divergence.In reliability analysis, Jensen-φ divergences are introduced for comparing coherent and mixed systems with identically distributed components, and their fundamental properties are established. The proposed divergences quantify system predictability, where smaller divergence values correspond to higher predictability.

References

[1] Balakrishnan, N,, Kateri, M. & Szymkowiak, M. (2026). Divergence of Lifetimes of Reliability Systems and Some Associated Properties, Characteristics, and Comparison results. *(submitted)*.

[2] Kateri, M. & Nikolov, N.I. (2022). A generalized Mallows model based on ϕ–divergence measures. *Journal of Multivariate Analysis*, **190**, 104958 (1 – 14).

[3] Theocharous, D. & Kateri, M. (2026). Power-divergence coordinate ascent variational inference for logistic regression *(submitted)*.

Testing Independence under a Monotone Missingness Surrogate Model

Marija Cuparić[1], Bojana Milošević[1], Jelena Radojević[2,1]

[1]University of Belgrade, Faculty of Mathematics, Serbia [2]University of Belgrade, Faculty of Civil Engineering, Serbia.

marija.cuparic@matf.bg.ac.rs; bojana@matf.bg.ac.rs; jradojevic@grf.bg.ac.rs

Abstract. We consider a monotone missingness surrogate model in which the variables of interest are observed through a surrogate observation mechanism governed by an auxiliary variable. This setting arises naturally when the availability of one variable implies the availability of another, leading to a monotone pattern of missingness. We introduce a novel modification of Kendall's coefficient for testing the independence of two random variables under this framework and investigate its performance through an extensive simulation study. The results highlight both the strengths and limitations of the proposed approach and suggest several promising directions for future research.

Maximum-likelihood estimation for censored semi-Markov chains

Maroua Souiah[1], Vlad Stefan Barbu[2] & Saâdia Rahmani[1]

[1]University of Saida, Algeria, [2]University of Rouen–Normandy, Saint Etienne du Rouvray, France

maroua.souiah@univ-saida.dz ; vladstefan.barbu@univ-rouen.fr;

saadia.rahmani@univ-saida.dz

Abstract. The present contribution deals with maximum-likelihood estimation in discrete-time homogeneous nonparametric semi-Markov models with a finite state space. We provide the exact maximum-likelihood estimator of the conditional sojourn time distribution and analyze its asymptotic properties. Two scenarios are considered: a single observed trajectory as its length tends to infinity, and multiple independent trajectories censored at a fixed time as the number of copies increases. In both frameworks, we establish strong consistency and asymptotic normality for the proposed estimators.

References

[1] V. S. Barbu and N. Limnios. *Semi-Markov chains and hidden semi-Markov models toward applications: their use in reliability and DNA analysis*, volume 191. Springer Science & Business Media, 2009.

[2] O. Chryssaphinou, M. Karaliopoulou, and N. Limnios. On discrete time semi-markov chains and applications in words occurrences. *Communications in Statistics—Theory and Methods*, 37(8):1306–1322, 2008.

[3] N. Limnios and G. Oprisan. *Semi-Markov processes and reliability*. Springer Science & Business Media, 2001.

[4] S.Trevezas and N.Limnios. Exact MLE and Asymptotic Properties for Nonparametric Semi-Markov Models, *Journal of Nonparametric Statistics*, 23(3), 719–739, 2011.

A Factor-Driven Shared Hidden Markov Model for High Frequency Assets Returns

Guglielmo D'Amico[1], Michele Palma[2], Filippo Petroni[3]

[1]University "G. d'Annunzio" of Chieti-Pescara, Italia, [2]IDSIA- Università della Svizzera Italiana, Switzerland, [3]University "G. d'Annunzio" of Chieti-Pescara.

g.damico@unich.it; michele.palma@usi.ch; filippo.petroni@unich.it;

Abstract. Markov processes are widely used in finance to model endogenous volatility regimes, where states correspond to distinct dynamics in asset returns. A central challenge in such models is the identification of the optimal number of states and the related discretization of state space. While change-point methods can infer regime structures, extending them to multivariate settings is difficult due to strong cross-sectional dependence among financial assets. This work introduces a novel methodology that captures cross-sectional dependence through latent market factors rather than explicit asset-level dependencies. Common factors are identified via Principal Component Analysis, which informs the selection of the number of hidden states in a consistent and data-driven manner across all assets. Asset-specific dynamics are modeled at the micro level, while being governed by shared macro-level hidden states. Empirical results on high-frequency financial data show that the proposed model reproduces key stylized facts of asset returns.

References

[1] Yacine Ait-Sahalia and Dacheng Xiu. (2017) Using principal component analysis to estimate a high dimensional factor model with high-frequency data. *Journal of Econometrics*, **201**(2):384–399

[2] Guglielmo D'Amico, Ada Lika, and Filippo Petroni. (2019) Change point dynamics for financial data: an indexed markov chain approach. *Annals of Finance*, **15**(2):247–266

Transition to a green economy in Romania - Population Perspectives: Perceptions, Barriers and Expectations

Mihaela Irma Vlădescu[1, 2], Mihaela Georgiana Oprea[2]

[1] Romanian Academy, School of Advanced Studies of the Romanian Academy, Doctoral School of Economic Sciences, National Institute for Economic Research "Costin C. Kirițescu", Institute of National Economy, Romania, [2] Center for Demographic Research, National Institute for Economic Research "Costin C. Kiritescu", Romania.

mihaelavladescu22@gmail.com; oprea@ince.ro

Abstract. One of the most complex challenges facing modern society today is the transition to a green economy which involves the transformation of energy systems, labor market dynamics, consumption patterns, and the legislative and institutional framework. Even though researchers' interest in this subject has increased in recent years, the way in which the population becomes aware of, perceives and feels the effects of this process remains insufficiently researched in Romania. Given this aspect, the study aims to partially cover this deficit through a two-stage methodological approach, combining bibliometric analysis with original empirical research based on data collected through a questionnaire. The bibliographical data used in the first stage were extracted from the Web of Science Core Collection (WoS) database, published in the last 49 years. In order to process and visualize the data were used two complementary tools: Bibliometrix (R package) and VOSviewer (Van Eck & Waltman, 2010). In the second stage, the data collected through a questionnaire distributed in the first part of this year at national level were interpreted. This questionnaire was organized in several modules and tried to cover numerous dimensions. The results obtained highlight the fact that the transition to the green economy represents a complex process that, in order to produce real and sustainable results, must be understood, accepted and assumed by all the factors involved.

References

[1] Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.

Stochastic Ordering of Dependent Systems under Transformation Models and Archimedean Copulas

Idir Arab[1], Milto Hadjikyriakou[2], Paulo Eduardo Oliveira[1]

[1]University of Coimbra, Portugal, [2]UCLan Cyprus

idir.bhh@gmail.com; mhadjikyriakou@uclan.ac.cy; paulo@mat.uc.pt

Abstract. We study stochastic ordering of system lifetimes with dependent and heterogeneous components whose marginal distributions are obtained through transformations of a common baseline. The dependence structure is modeled via Archimedean copulas, allowing for a unified treatment of several transformation-based models, including proportional hazard, proportional reversed hazard rate and proportional odds families. For parallel, series and (n−k)-out-of-n systems, we derive conditions for stochastic dominance based on monotonicity of the transformation and structural properties of the copula generators, formulated through super-additivity and Schur-type arguments. The results provide tractable criteria that extend existing comparisons beyond independence and illustrate the combined effect of dependence and parameter heterogeneity on system reliability.

References

[1] Arab, I., Hadjikyriakou, M., Oliveira, P.E. (2025). Stochastic orders and shape properties for a new distorted proportional odds model. *Statist. Papers* **66** (5), 107-122

[2] Marshall, A.W., Olkin, I., Arnold, B.C. (2011). *Inequalities: Theory of Majorization and Its Applications*. Elsevier

[3] Navarro, J., ÅL Aguila, Y. (2017). Stochastic comparisons of distorted distributions, coherent systems and mixtures with ordered components. *Metrika* **80** (6-8), 627-648

Modelling New Group Chain Acceptance Sampling Plans via the Minimum Angle Method

Mohd Azri Pawan Teh[1, 2, 3], Aiman Fikri Jamaludin[1] & Nazrina Aziz[1]

[1]Universiti Utara Malaysia, Malaysia, [2] Institute of Strategic Industrial Decision Modelling (ISIDM), Malaysia, [3] Institute of Zakat Research and Innovation (IPIZ), Malaysia.

mohd.azri.pawan@uum.edu.my; aiman_fik@ahsgs.uum.edu.my; nazrina@uum.edu.my

Abstract. Acceptance sampling is a fundamental statistical quality control tool used to determine whether product lots meet specified quality standards. Although new group chain acceptance sampling plans (NGChSP) were designed to balance the strict selection criteria of existing plans, the current NGChSP are developed by minimizing the consumer's risk only, neglecting the producer's risk, which is unfair to the producer. To resolve this imbalance, this study develops an NGChSP using the minimum angle method, which simultaneously considers both producer's and consumer's risks by minimizing the angle between the actual and ideal operating characteristic (OC) curves. The research is systematically executed across six distinct phases: defining design parameters, establishing operating procedures, deriving probability of lot acceptance, determining the probability of defectives, calculating the Weibull distribution operating fraction defective, and evaluating performance. The findings demonstrate that the optimal number of groups obtained through this method successfully satisfies the quality interests of both parties. The minimized angle ensures that the resulting OC curve closely resembles the ideal OC curve, proving that the proposed plan minimizes risks for both producers and consumers, making it highly suitable for practical quality control applications.

References

[1] Teh, M. A. P., Aziz, N. & Zain, Z. (2021). New method in designing group chain acceptance sampling plans (GChSP) for generalized exponential distribution, *International Journal of Quality & Reliability Management*, **38**, 1116-1129.

Modelling New Two-Sided Group Chain Acceptance Sampling Plans for Gamma Distribution

Mohd Azri Pawan Teh[1, 2, 3], Syed Alhanif Syed Alhakim[1], Aiman Fikri Jamaludin[1] & Nazrina Aziz[1]

[1]Universiti Utara Malaysia, Malaysia, [2] Institute of Strategic Industrial Decision Modelling (ISIDM), Malaysia, [3] Institute of Zakat Research and Innovation (IPIZ), Malaysia.

mohd.azri.pawan@uum.edu.my; syed_alhanif_syed@sqs.uum.edu.my; aiman_fik@ahsgs.uum.edu.my; nazrina@uum.edu.my

Abstract. Acceptance sampling is a fundamental statistical quality control method for determining lot conformity. The new two-sided group chain acceptance sampling plans (NTSGChSP) were developed as an improvement over the new two-sided complete group chain acceptance sampling plans (NTSCGChSP). However, no research has yet applied the NTSGChSP to products with lifetimes following a gamma distribution via the Poisson distribution, despite many modern products exhibiting these characteristics. To address this limitation, this study aims to develop the NTSGChSP specifically for products whose lifetime follows a gamma distribution modeled via the Poisson distribution. The methodology comprises six phases: defining design parameters, establishing operating procedures, deriving probability of lot acceptance, determining probability of defectives, calculating the gamma-Poisson operating fraction defective, and evaluating performances. Evaluated based on the minimized number of groups and probability of lot acceptance, the findings demonstrate that the NTSGChSP exhibits better performance compared to the NTSCGChSP, as it successfully requires smaller number of groups while yielding a higher probability of lot acceptance. Consequently, these findings offer industry practitioners a highly cost-effective and dependable quality assurance acceptance sampling plans that curtails inspection overheads without compromising rigorous quality standards.

References

[1] Teh, M. A. P., Aziz, N. & Zain, Z. (2021). New group chain acceptance sampling plans (NGChSP-1) using minimum angle method for generalized exponential distribution, *Sains Malaysiana*, **50**, 1121-1129.

Extensions of Grey Forecasting Models for Insurance Loss Prediction

Nikolaos G. Galanopoulos[1], Alex Karagrigoriou[2] & Andreas Makrides[1]

[1]University of the Aegean, Greece, [2]University of Piraeus, Greece

ngalan@aegean.gr; alex.karagrigoriou@unipi.gr; amakridis@aegean.gr

Abstract. Grey models GM(M,N) have been widely applied in systems characterized by incomplete information, uncertainty, and relatively small sample sizes. In actuarial science, insurance loss data often exhibits structural heterogeneity and volatility, making Grey system methodologies a promising framework for forecasting applications. This study investigates extensions of classical Grey forecasting models for automobile insurance loss prediction using state-level data from the United States. Building upon previous work on GM(1,1) and Markov Chain Grey Model MCGM(1,1), we explore multivariate Grey models GM(1,N) incorporating actuarial decomposition components. Additional extensions involving adaptive model specifications and alternative background-value formulations are also considered. The proposed framework aims to improve forecasting flexibility and predictive performance while preserving the suitability of Grey models for small-sample environments. Empirical results are evaluated using standard forecast accuracy measures and compared with classical Grey forecasting approaches and traditional statistical benchmark models. The study contributes to the growing literature on Grey systems and actuarial forecasting by examining generalized Grey model structures for insurance loss prediction under uncertainty.

References

[1] D'Amico, G., Karagrigoriou, A. and Vigna, V. (2024). Forecasting the power generation mix in Italy based on Grey Markov models, *Energies*, **17**, 2184.

[2] Deng, J. (1982). Control problems of grey systems, *Systems & Control Letters*, **1**, 288-294.

Robust Adaptive Efficient Estimation for Autoregressive Big Data Models

Ouerdia Arkoun[1], J.-Yves Brua[2] & Serge Pergamenchtchikov[2]

[1] Laboratoire BIRL, Sup'Biotech and Laboratoire de Mathématiques Raphaël Salem, Université de Rouen Normandie, [2] Laboratoire de Mathematiques Raphaël Salem, Université de Rouen Normandie

ouerdia.arkoun@supbiotech.fr ; Jean-Yves.brua@univ-rouen.fr
serge.pergamenchtchikov@univ-rouen.fr

Abstract. In this talk, we propose an efficient adaptive estimation method using robust sequential model selection procedures for big data models represented through the non-parametric autoregressive processes. To this end, by applying the Van Trees inequality to this model, we obtain a lower bound for the robust risk expressed in the form of a Pinsker-type constant. Then, using the weighted least squares method together with non-asymptotic oracle inequalities, we develop an analytical framework that allows us to establish the minimax efficiency property of the proposed estimation procedure, that is, we show that the upper bound of its risk coincides with the obtained lower bound. It is worth emphasizing that this efficiency is achieved in an adaptive setting where the regularity of the autoregressive function is unknown.

References

[1] Arkoun, O., Brua, J.-Y. and Pergamenchtchikov, S. (2024): Adaptive efficient robust sequential analysis for autoregressive big data models. *Sequential Anal.* **43** (4), 432-460.

[2] Arkoun, O., Brua J.-Y. and Pergamenchtchikov, S. (2019): Sequential model selection method for nonparametric autoregression. *Sequential Anal*. **38** (4), 437-460.

[3] Arkoun, O. and Pergamenchtchikov, S. (2016): Sequential robust estimation for nonparametric autoregressive models. *Sequential Anal*. **35** (4), 489-515.

[4] Arkoun, O. (2011): Sequential adaptive estimators in nonparametric autoregressive models. *Sequential Anal*. **30** (2), 228-246.

The Statistical Engine of Constraint-Based Causal Discovery: Conditional Independence Tests in Practice

Pavel Averin[1], Theodoros Moysiadis[1], Ioannis Katakis[1]

[1] Department of Computer Science, School of Sciences and Engineering, University of Nicosia, 2417, Nicosia, Cyprus

averin.p1@live.unic.ac.cy; moysiadis.t@unic.ac.cy; katakis.i@unic.ac.cy.

Abstract. Constraint-based algorithms such as Peter-Clark and Fast Causal Inference learn a causal graph from observational data by performing a large number of conditional independence tests, each asking whether two variables are independent given a set of other variables. Every edge in the graph is kept or removed on the basis of such a test, so the calibration and power of the test directly govern the quality of the recovered graph. This talk surveys the landscape of conditional independence tests used in constraint-based causal discovery and groups them into six families: partial-correlation, contingency-table, regression, nearest-neighbor, kernel, and machine-learning based. For each family we discuss the underlying assumptions, calibration under the null, power against relevant alternatives, behavior with mixed continuous and categorical data, and computational cost. We then turn to challenges that are particularly sharp in biomedical data, where samples are high-dimensional, small, and a mixture of continuous and discrete variables. Three stand out: unified testing for mixed-type data without ad hoc discretization, error control in small samples, and scalability as the conditioning set grows. The talk closes with early results from ongoing work benchmarking representative tests within the Peter-Clark algorithm, and a discussion of open problems in the field.

References

[1] Shah, R. D. & Peters, J. (2020). The hardness of conditional independence testing and the generalised covariance measure, *The Annals of Statistics*, **48**, 1514-1538.

[2] Spirtes, P., Glymour, C. & Scheines, R. (2000). *Causation, Prediction, and Search*, 2nd ed., MIT Press, Cambridge, MA.

[3] Zhang, K., Peters, J., Janzing, D. & Schölkopf, B. (2011). Kernel-based Conditional Independence Test and Application in Causal Discovery, In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence (UAI)*, 804-813.

Tsallis and Kaniadakis divergences between probability measures

Răzvan-Cornel Sfetcu[1] & **Vasile Preda**[2,3,1]

[1]University of Bucharest, Romania, [2]"Gheorghe Mihoc-Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics, Romania, [3]"Costin C. Kiriţescu" National Institute of Economic Research, Romania

razvancornelsfetcu@gmail.com; vasilepreda0@gmail.com

Abstract. We investigate several fundamental properties of Tsallis and Kaniadakis divergences associated with two probability measures. In particular, we analyze the behavior of these generalized divergence measures and establish a number of relevant theoretical results. More precisely, we study their pseudo-additivity, non-negativity and monotonicity and derive several bounds for both Tsallis and Kaniadakis divergences. The obtained results provide further insight into the structure and applicability of these generalized information-theoretic quantities.

References

[1] Barbu, V.S., Karagrigoriou, A. & Preda, V. (2017). Entropy, divergence rates and weighted divergence rates for Markov chains. I: The alpha-gamma and beta-gamma case, *Proc. Rom. Acad., Ser. A, Math. Phys. Tech. Sci. Inf. Sci.* **18**, 293-301.

[2] Barbu, V.S., Karagrigoriou, A. & Preda, V. (2018). Entropy and divergence rates for Markov chains. II: The weighted case, *Proc. Rom. Acad., Ser. A, Math. Phys. Tech. Sci. Inf. Sci.* **19**, 3-10.

[3] Barbu, V.S., Karagrigoriou, A. & Preda, V. (2018). Entropy and divergence rates for Markov chains. III: The Cressie and Read case and applications, *Proc. Rom. Acad., Ser. A, Math. Phys. Tech. Sci. Inf. Sci.* **19**, 413-421.

[4] Sfetcu, R.-C., Sfetcu, S.-C. & Preda, V. (2022). On Tsallis and Kaniadakis divergences, *Math. Phys. Anal. Geom.* **25**, 7.

[5] Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics*, Springer Science Business Media LLC.

Stein's method for the matrix normal distribution

Robert E. Gaunt[1], Frederic Ouimet[2] & Donald Richards[3]

[1]The University of Manchester, UK, [2]Université du Québec à Trois-Rivières, Canada, [3]Penn State University, USA.

Robert.gaunt@manchester.ac.uk

Abstract. Stein's method is a powerful technique for bounding the distance between two probability distributions with respect to a probability metric. Stein's method has proven to be a powerful tool for deriving quantitative probabilistic limit theorems, and has recently found exciting statistical applications beyond the traditional application to distributional approximation. In this talk, we present a basic framework for Stein's method for matrix normal approximation, which represents the first systematic treatment of Stein's method for matrix variate distributions.

References

[1] Gaunt, R. E., Ouimet, F. and Richards, D. (2026). Stein's method for the matrix normal distribution. arXiv:2601.11422,.

Indexed Markov Chains for Intraday Trading with Technical Indicators

Guglielmo D'Amico[1], Filippo Petroni[1], Salvatore Vergine[2]

[1]Department of Economics, University G. d'Annunzio of Chieti-Pescara, Viale Pindaro, 42, Pescara, 65127, Italy, [2]Department of Management, Marche Polytechnic University, Piazzale R. Martelli, 8, Ancona, 60121, Italy

g.damico@unich.it; filippo.petroni@unich.it; s.vergine@staff.univpm.it.

Abstract. This paper proposes a technical approach in which financial indicators improve intraday trading decisions when they are used to drive regime-dependent transition dynamics in discretized return processes. To this end, we propose an Indexed Markov Chain framework in which the one-minute return process is mapped into a finite state space and the obtained transition probabilities depend on the level of a technical indicator. Regime changes are identified through a change-point procedure, which partitions the indicator space into homogeneous regions and allows us to estimate a different transition matrix for each regime. Statistical significance of regime dependence is assessed by a bootstrap likelihood-ratio test against the homogeneous Markov benchmark. The empirical analysis is conducted on one-minute data for eight U.S. equities and considers eleven technical indicators. The results show that, for most indicators and stocks, the indexed model outperforms the corresponding homogeneous Markov model in terms of final cumulative gain. Finally, we aggregate the information contained in the different indexed models through a constrained least squares combination and construct a super index, which further improves the overall trading performance.

References

[1] G. D'Amico, A. Lika, F. Petroni, (2019). Change point dynamics for financial data: an indexed markov chain approach, *Annals of Finance* **15** (2) 247–266.

[2] G. D'Amico, F. Petroni, (2012). Weighted-indexed semi-markov models for modeling financial returns, *Journal of Statistical Mechanics: theory and experiment* (07) P07015.

A Variance-Stabilizing Framework for Higher-Order Cumulant Estimation

Sandra S. Ferreira[1,2], Dário Ferreira[1,2]

[1]University of Beira Interior, Mathematics Department and [2]Instituto de Telecomunicações

sandraf@ubi.pt; dario@ubi.pt

Abstract. Higher order cumulants play an important role in the analysis of stochastic models, particularly in settings where deviations from Gaussian, dependence structures, asymmetry, and tail behavior are of interest. However, the estimation of fourth-order cumulants is often affected by pronounced heteroscedasticity, since the variance of the corresponding estimators may depend strongly on eighth-order quantities. This dependence can lead to unstable inference, especially in finite samples or in models with heavy-tailed or non-standard distributions. In this talk, we propose and discuss a variance-stabilizing approach for cumulant estimation. The proposed perspective provides a principled way to address variability inflation in cumulant-based procedures and contributes to the development of more reliable methods for inference in complex stochastic models.

References

[1] Zhang, A.R. and Xia, D. (2023). Tensor SVD: Statistical and computational limits. *IEEE Transactions on Information Theory*, **69** (2), 953-972.

[2] Withers, C.S. and Nadarajah, S. (2021). Formulae for moments and cumulants. *Communications in Statistics - Theory and Methods*, **50** (24), 5889-5901.

[3] Speed, T.P. and Silcock, H.L. (2020). Cumulants and partition lattices. Australian & New Zealand *Journal of Statistics*, **62** (1), 1-22.

Risk models in actuarial science using entropy measures

Silvia Dedu [1,2], Vasile Preda[2,3,4] & Muhammad Sheraz[5]

[1]Bucharest University of Economic Studies, Romania, [2]Centre for Demographic Research "Vladimir Trebici", "Costin C. Kiritescu" National Institute of Economic Research of Romanian Academy, Romania, [3]University of Bucharest, Romania [4]"Gheorghe Mihoc-Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of Romanian Academy, Romania, [5]Institute of Business Administration Karachi, University Rd, University of Karachi Sindh, Pakistan.

silvia.dedu@csie.ase.ro; vasilepreda0@gmail.com; msheraz@iba.edu.pk

Abstract. Effective management of financial and actuarial risks represents a major challenge within a volatile economic environment. Traditional loss models often rely on parametric assumptions that can underestimate tail risk and the frequency of extreme events. This paper explores and extends the application of entropy measures, including Shannon, Tsallis, Rényi and Kaniadakis entropies, as fundamental mathematical tools for constructing and calibrating loss models in insurance. The proposed methodology utilizes entropic techniques to estimate claim severity from an informational theory perspective and establishes their main mathematical properties. Empirical results, tested on real-world insurance portfolios, prove the improved accuracy of entropy-based loss models in capturing the asymptotic behavior of heavy-tailed losses. Integrating entropy into actuarial loss models provides a more flexible and reliable framework for underwriting risk management.

Keywords: Risk models, Entropy measures, Loss models, Actuarial science.

References

[1] Elbatal, I., Diab, L.S., Ghorbal, A.B., Yousof, H. M., Elgarhy, M., & Ali, E.I.A. (2024). A new losses (revenues) probability model with entropy analysis, applications and case studies for value-at-risk modeling and mean of order-P analysis, *AIMS Mathematics*, **9**, 3, 7169-7211.

[2] Preda, V., Dedu, S., Iatan, I., Cernat, I. D. & Sheraz, M. (2022). Tsallis entropy for loss models and survival models involving truncated and censored random variables, *Entropy*, **24**, 11, 1654

[3] Xu, M., & Angulo, J. M. (2019). Divergence-based risk measures: A discussion on sensitivities and extensions, *Entropy*, **21**, 7, 634.

Multiple change-point detection on the circle via isolation using permutation testing

Sophia Loizidou[1], Andreas Anastasiou[2], Christophe Ley[1]

[1]University of Luxembourg, Luxembourg, [2]University of Cyprus, Cyprus

Sophia.loizidou@uni.lu; Anastasiou.andreas@ucy.ac.cy; Christophe.ley@uni.lu

Abstract. In this talk we propose a new method for multiple change-point detection for piecewise-constant circular signals, a setting that, despite its importance in many scientific domains, remains comparatively under-explored. The proposed method, Permutation-based Circular Isolate-Detect, denoted PCID, uses an appropriately chosen contrast function and permutation testing to detect change-points in an offline manner, for the data sequence under consideration. Prior to detection, PCID isolates the change-points. The contrast function used is derived under the assumption of von Mises distribution for the noise, but we show that the method is robust and performs well for other distributions as well. Simulations are used to showcase the usability of the method in different signal and noise structures, including serially correlated noise. In order to exhibit the practical relevance of the method in real-world applications, PCID is applied to real-world datasets.

A new family of generalized r-modified reliability systems

Stylianos-Marinos Charalampidis[1], Spiros D. Dafnis[1] & Markos V. Koutras[2]

[1]University of the Aegean, Greece, [2]University of Piraeus, Greece

sasm23014@aegean.gr; sdafnis@aegean.gr; mkoutras@unipi.gr

Abstract. Motivated by recent works on protection blocks in the reliability literature, in this article we introduce a new family of reliability systems that generalizes the *r*-modified family. The new systems exhibit a fundamental structural difference compared to their predecessors: instead of locking after the occurrence of a single working component, they now lock only after a run of *d* consecutive working components. This run of working components acts as the analogue of a protection block, serving as the local condition that activates protection; the subsequent *r*-1 components form the protected block, which becomes locked and is treated as non-failed. We derive closed-form expressions for the reliability of the fundamental members of this new family, and we show how these formulas recover well-known reliability results as special cases. Numerical results and simulations further demonstrate the applicability of the proposed framework.

References

[1] Balakrishnan, N., Dafnis, S.D., & Makri, F.S. (2025). On a new family of *r*-modified reliability systems. *Probability in the Engineering and Informational Sciences*, **39**(4), 605–618, https://doi.org/10.1017/S0269964825100053.

[2] Dafnis, S.D., Koutras, M.V., & Makri, F.S. (2022). Binomial Distribution of Order *k* in a Modified Binary Sequence. *Journal of Statistical Theory and Practice*, **16**(42), https://doi.org/10.1007/s42519-022-00267-7.

Data-Driven Modelling of Agro-Environmental Systems

Steliana Rodino[1,2], Ruxandra Pop[1], Carmen Adriana Gheorghe[3]

[1]Research Institute for Agriculture Economy and Rural Development (ICEADR), Bucharest, Romania

[2]National Institute of Research and Development for Biological Sciences, Bucharest, Romania,

[3]National Institute for Economic Research "Costin C. Kirițescu", Romanian Academy, 050711 Bucharest, Romania,

steliana@rodino.ro; Ruxandra.pop@iceadr.ro; d_c_adriana@yahoo.com

Abstract. The increasing complexity of agro-environmental systems under the pressure of climate change and resource constraints calls for advanced modelling approaches to support sustainable development. Understanding the interactions between environmental processes, agricultural practices and resource use is essential for designing efficient and resilient systems within a circular bioeconomy framework. This study investigates the role of applied and data-driven modelling in analysing complex agro-environmental systems, with a focus on integrating environmental indicators, agricultural productivity and sustainability metrics. By combining statistical and on-farm data sources, including climate variables, vegetation indices and biomass-related information, the research highlights how modelling techniques can be used to capture system dynamics and support informed decision-making. Particular attention is given to the use of quantitative indicators for monitoring system responses and evaluating the impact of environmental changes on agricultural systems. The results emphasize the importance of interdisciplinary approaches that connect environmental data analysis with economic and sustainability-oriented perspectives.

Acknowledgment: This paper was partly supported by Ministry of Agriculture and Rural Development Romania, Project ADER 22.1.4/2023 and the National Authority of Research through the Core Program of the National Research, Development and Innovation Plan 2022-2027, project no. PN 23-02-0101-Contract No. 7N/2023

On the Weighted (Φ, a) Power Divergence Family

Vlad Stefan Barbu [1,2], Andreas Makrides[3], Alex Karagrigoriou[4], & Vasile Preda[2,5,6]

[1]University of Rouen Normandy, France, [2]Centre for Demographic Research "Vladimir Trebici", "Costin C. Kiritescu" National Institute of Economic Research of Romanian Academy, Romania, [3]University of the Aegean, Greece, [4]University of Piraeus, Greece, [5]University of Bucharest, Romania [6]"Gheorghe Mihoc-Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of Romanian Academy, Romania.

barbu@univ-rouen.fr; andreasmaths@hotmail.com; alex.karagrigoriou@unipi.gr; vasilepreda0@gmail.com

Abstract. This paper introduces the weighted (Φ, a) power divergence family, a general framework that unifies several classical divergence measures, including the Kullback–Leibler, Csiszár, Cressie–Read, and density power divergences. We develop weighted versions for continuous and discrete probability distributions and establish their main mathematical properties. The framework includes important special cases obtained through specific choices of the convex function (Phi) and weighting schemes. Quadratic convergence and asymptotic behaviour are studied through discretization and limit theorems. The proposed family provides a flexible tool for statistical inference and goodness-of-fit testing.

References

[1] Basu, A., Harris, I.R., Hjort, N.L. & Jones, M.C. (1998). Robust and efficient estimation by minimising a density power divergence, *Biometrika*, 85(3), 549–559.

[2] Mattheou, K., Karagrigoriu, A. (2010). A new family of divergence measures for tests of fit. *Aust. N. Z. J. Stat.* 52(2),187–200.

[3] Vonta, F., Mattheou, K. & Karagrigoriou, A. (2012). On Properties of the (Φ, a)-Power Divergence Family with Applications in Goodness of Fit Tests, *Methodology and Computing in Applied Probability*, 14, 335–356.

Convergence of PSO and the non-Gaussian limit

Alexandra Andriciuc[1], Iulian Cîmpean[1], Ionel Popescu[1]

[1]University of Bucharest and Institute of Mathematics "Simion Stoilow" of the Romanian Academy

alexandra.andriciuc@unibuc.ro; iulian.cimpean@unibuc.ro; ionel.popescu@fmi.unibuc.ro

Abstract. I will discuss three things related to the Particle Swarm Optimization (PSO) which is an optimization tools of zero order, namley the function to be optimized does not require smoothness, only continuity. The first thing I will discuss is convergence in distribution of the stagnant version of the PSO. We will present some kind of universality result which is rather surprising. The second part is the convergence of the full PSO and the conditions under which we have convergence. The third part is the resolution of an open problem in the community related to the non-Gaussianity of the limiting distribution.

Asymptotic properties of semiparametric M-estimators with multiple change points

Salim Bouzebda[1], Anouar Abdeldjaoued Ferfache[2]

[1]University of Technology of Compiègne, France, [2]Aix-Marseille University, France.

Email1 salim.bouzebda@utc.fr

Abstract. Statistical models with multiple change points are widely used in many fields. However, the theoretical properties of semiparametric M-estimators for such models have received comparatively limited attention. The main objective of this work is to study the asymptotic properties of semiparametric M-estimators based on non-smooth criterion functions in multiple change-point models. We consider a general class of models in which the distributional form may vary from one segment to another, while some parameters may remain common across all segments. We establish the consistency of the semiparametric M-estimators of the change points and determine their rate of convergence. Under fairly general conditions, we also prove the asymptotic normality of the semiparametric M-estimators of the within-segment distributional parameters. These results, together with a general framework for analyzing semiparametric M-estimation in multiple change-point settings, extend previous work on related parametric and semiparametric maximum-likelihood estimators. As an illustration, we examine in detail a classification model with missing data and provide a brief simulation study.

N

O

P

R

S

T

V