BOOK OF ABSTRACTS

ICAS CONFERENCE

INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

12th INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

10-12 APR 2026

ICAS CONFERENCE

INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

April 10-12, 2026

https://www.icasconference.com

ICAS'2026

12th International Conference on Advances in Statistics

Published by the ICAS Secretariat

Editors:
Prof. Dr. Ismihan BAYRAMOGLU

ICAS Secretariat
Büyükdere Cad. Ecza sok. Pol Center 4/1 Levent-İstanbul
E-mail: icasconference.academic@gmail.com
http://www.icasconference.com

ISBN: TBA

INTERNATIONAL SCIENTIFIC COMMITTEE

LOCAL SCIENTIFIC COMMITTEE

Prof. Dr. Aydin ERAR
Mimar Sinan Fine Arts University – Turkey

Prof. Dr. Baris ASIKGIL
Mimar Sinan Fine Arts University – Turkey

Prof. Dr. Coskun KUS
Selcuk University – Turkey

Prof. Dr. I. Esen YILDIRIM
Marmara University – Turkey

Prof. Dr. Esra AKDENIZ
Marmara University – Turkey

Prof. Dr. Gulay BASARIR
Mimar Sinan Fine Arts University – Turkey

Prof. Dr. Ismail KINACI
Selcuk University – Turkey

Prof. Dr. Mehmet Fedai KAYA
Selcuk University – Turkey

Prof. Dr. Mujgan TEZ
Marmara University – Turkey

Prof. Dr. Sahamet BULBUL
Fenerbahce University – Turkey

Prof. Dr. Yazgi TUTUNCU
Izmir University of Economics, Turkey

Assoc. Prof. Dr. Konul KAVLAK
Wageningen University – Holland

ORGANISING COMMITTEE

Prof. Dr. Ismihan BAYRAMOGLU
Izmir University of Economics – Turkey
Conference Chair

Prof. Dr. Fatma NOYAN TEKELI
Yıldız Technical University – Turkey

Prof. Dr. Gulhayat GOLBASI SIMSEK
Yıldız Technical University – Turkey

Prof. Dr. Seher Nur SULKU
Ankara Hacı Bayram Veli University – Turkey

Assoc. Prof. Dr. Aysegul EREM
Cyprus International University – TRNC

Phd. Ayse BELER
Izmir University of Economics, Turkey

Gokce OZALTUN
Izmir University of Economics, Turkey

Dear Colleagues,

Dear Colleagues,

On behalf of the Organizing Committee, I am pleased to invite you to participate in 8th International Conference on Advances in Statistics which will be organised in Turkish Republic of Northern Cyprus on dates between 10-12 April 2026.

All informations are available in conference web site. For more information please do not hesitate to contact us. info@icascconference.com

We cordially invite prospective authors to submit their original papers to ICAS-2026,

. Applied Statistics
· Bayesian Statistics
· Big Data Analytics
· Bioinformatics
· Biostatistics
· Computational Statistics
· Data Analysis and Modeling
· Data Envelopment Analysis
· Data Management and Decision Support Systems
· Data Mining
· Energy and Statistics
· Entrepreneurship
· Mathematical Statistics
· Multivariate Statistics
· Neural Networks and Statistics
· Non-parametric Statistics
. Operations Research
· Optimization Methods in Statistics
· Order Statistics
· Panel Data Modelling and Analysis
· Performance Analysis in Administrative Process
· Philosophy of Statistics
· Public Opinion and Market Research
· Reliability Theory
· Sampling Theory
· Simulation Techniques
· Spatial Analysis
· Statistical Software
· Statistical Training
· Statistics Education
· Statistics in Social Sciences
· Stochastic Processes
· Supply Chain
· Survey Research Methodology
· Survival Analysis
· Time Series
· Water and Statistics
· Other Statistical Methods

Selected papers will be published in Journal of the Turkish Statistical Association. https://dergipark.org.tr/en/pub/ijtsa

We hope that the conference will provide opportunities for participants to exchange and discuss new ideas and establish research relations for future scientific collaborations.

Conference Website : https://icasconference.com
E Mail: icasconference.academic@gmail.com

On behalf of Organizing Committee:
Conference Chair
Prof. Dr. İsmihan BAYRAMOĞLU
Izmir University of Economics

ICAS CONFERENCE
INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

CONFERENCE VENUE

TEMBO HOTEL BARCELONA

10 APRIL 2026 FRIDAY

Welcome Speech (09:45 - 10:00)

Prof. Dr. Ismihan BAYRAMOĞLU, Izmir University of Economics
Conference Chair

KEYNOTE SPEECH #1 (10:00 – 10:40)

Prof. Dr Maria LONGOBARDI
Speech Title: ON A UNIFIED VERSION OF EXTROPY AND ITS WEIGHTED DEFINITIONS

TEA & COFFEE BREAK 10:40 – 11:00

SESSION A (11:00 – 12:40)

SESSION CHAIR: Prof. Dr Maria LONGOBARDI

11:00 – 11:20
PAPER TITLE : ON THE NUMBERS OF COMPONENTS IN EACH STATE IN MULTI-STATE DISCRETE K-OUT-OF-N SYSTEMS
AUTHOR(S) : Jakub SADOWY

11:20 – 11:40
PAPER TITLE : FINITE-STEP BOUNDS FOR ITERATED CORRELATION DYNAMICS
AUTHOR(S) : Ishrak ALHAJJ HASSAN

11:40 – 12:00
PAPER TITLE : A NOVEL IMPUTATION-BASED APPROACH FOR SYNTHETIC MINORITY OVERSAMPLING IN IMBALANCED CLASSIFICATION
AUTHOR(S) : **Pornthera AIMROD**, Anamai NA-UDOM, Jaratsri RUNGRATTANAUBOL

12:00 – 12:20
PAPER TITLE : MODELING OCCUPATIONAL ACCIDENT FREQUENCIES IN AN INDUSTRIAL SECTOR USING NEGATIVE BINOMIAL REGRESSION MODEL
AUTHOR(S) : Pelin TOKTAŞ

12:20 – 12:40 : QUESTIONS & CERTIFICATE AWARDING

LUNCH BREAK 12:40 – 13:40

SESSION B (13:40 – 15:20)

SESSION CHAIR: Prof. Dr. Yeliz MERT KANTAR

13:40 – 14:00
PAPER TITLE : ROBUST PARAMETER ESTIMATION APPROACH FOR UNIVERSAL KRIGING UNDER CONTAMINATED SPATIAL DATA
AUTHOR(S) : Vural YILDIRIM, Ezio TODINI, **Yeliz MERT KANTAR**

14:00 – 14:20
PAPER TITLE : A MULTI-LAYER STATISTICAL FRAMEWORK FOR INTEGRATED AND INTERPRETABLE STRUCTURES IN LARGE-SCALE TEXTUAL CORPORA, WITH APPLICATION TO PERMACULTURE DISCOURSE
AUTHOR(S) : **Alessandra MERIANI**, Ida D'ATTOMA

14:20 – 14:40
PAPER TITLE : MULTI-STEP TEMPERATURE FORECASTING WITH AN ATTENTION-BASED GATED RECURRENT UNIT MODEL
AUTHOR(S) : **Vildan KARA**, Gülesen ÜSTÜNDAĞ ŞİRAY

14:40 – 15:00
PAPER TITLE : IMPACT OF IMPUTATION ON LOGISTIC CLASSIFICATION WITH MCAR AND MAR MISSING DATA IN CARDIOVASCULAR DATASET
AUTHOR(S) : **Ratchaneewan PAISANWARAKIAT**, Anamai NA-UDOM, Jaratsri RUNGRATTANAUBOL

15:00 – 15:20 : QUESTIONS & CERTIFICATE AWARDING

TEA & COFFEE BREAK 15:20 – 16:00

SESSION C (16:00 – 17:10)

SESSION CHAIR: Prof. Dr. Dora PRATA GOMES

16:00 – 16:20
PAPER TITLE : COMPUTATIONAL PROCEDURES FOR THE ESTIMATION OF THE EXTREMAL INDEX
AUTHOR(S) : **Dora PRATA GOMES**, M. Manuela NEVES

16:20 – 16:40
PAPER TITLE : REVISITING THE EXTREME VALUE INDEX: THE PROBABILITY WEIGHTED MOMENTS FRAMEWORK FOR HEAVY-TAILED DISTRIBUTIONS
AUTHOR(S) : **Ayana MATEUS**, Frederico CAEIRO

16:40 – 17:00
PAPER TITLE : A REFINED BIAS-CORRECTED ESTIMATOR FOR THE SHAPE PARAMETER OF THE LOG-LOGISTIC MODEL
AUTHOR(S) : **Frederico CAEIRO**, Ayana MATEUS

17:00 – 17:10 : QUESTIONS & CERTIFICATE AWARDING

KEYNOTE SPEECH #2 (17:20 – 18:00)

Prof. Dr. Ömer ÖZTÜRK

Speech Title: MIXED-MODEL INFERENCE FOR PAIRWISE TREATMENT CONTRASTS WITH JUDGEMENT COVARIATES IN RANDOMIZED BLOCK DESIGNS

SOCIAL PROGRAM

HOTEL DEPARTURE 19:00

Will be organised as part of conference dinner program.
Please have your tickets with you

ICAS CONFERENCE
INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

11 April 2026 Saturday

KEYNOTE SPEECH #3 (10:15 – 11:00)

Prof. Dr Serkan ERYILMAZ

Speech Title: ANALOGIES BETWEEN ENGINEERING AND BIOLOGICAL SYSTEMS BASED ON RELIABILITY CONCEPTS AND MODELS

TEA & COFFEE BREAK 11:00 – 11:20

SESSION D (11:20 – 12:40)

SESSION CHAIR: Prof. Dr. Yeliz MERT KANTAR

11:20 – 11:40
PAPER TITLE : EFFICIENCY ANALYSIS OF TURKISH AIRPORTS: A STOCHASTIC FRONTIER ANALYSIS APPROACH
AUTHOR(S) : **Erdem Korhan AKÇAY**, Yeliz MERT KANTAR

11:40 – 12:00
PAPER TITLE : IIMAGINED GEOGRAPHIES OF SUSTAINABILITY: HOW SOCIAL MEDIA, COGNITIVE HEURISTICS, AND ATTITUDES SHAPE GENERATION Z' DESTINATION IMAGINATION—A PLS-SEM APPROACH
AUTHOR(S) : Serkan OLGAÇ, **Semra GÜNAY**, Almira SEISINBINOVA

12:00 – 12:20
PAPER TITLE : ROBUST DENSITY POWER DIVERGENCE BASED BAYESIAN ANALYSIS OF NON-DESTRUCTIVE ONE-SHOT DEVICE TESTING DATA
AUTHOR(S) : Shuvashree MONDAL

12:20 – 12:40 : QUESTIONS & CERTIFICATE AWARDING

LUNCH BREAK 12:40 – 13:40

SESSION E (13:40 – 15:00)

SESSION CHAIR: Prof. Dr. Ismihan BAYRAMOGLU

13:40 – 14:00
PAPER TITLE : WEIGHTED NON-STOCHASTIC FORECASTING RANGE OF TIME SERIES
AUTHOR(S) : Richard HINDLS, **Stanislava HRONOVÁ**, Luboš MAREK

14:00 – 14:20
PAPER TITLE : BIVARIATE BINOMIAL DISTRIBUTION AND BIVARIATE BERNSTEIN POLYNOMIALS
AUTHOR(S) : Ayşe BELER, **İsmihan BAYRAMOĞLU**

14:20 – 14:40
PAPER TITLE : ENTROPY-BASED BAYESIAN SEM FOR STUDENT SATISFACTION MODELING
AUTHOR(S) : Aslı CANAL

14:40 – 15:00 : QUESTIONS & CERTIFICATE AWARDING

TEA & COFFEE BREAK 15:00 – 15:20

SESSION F (15:20 – 16:40)

SESSION CHAIR: Prof. Dr. Coşkun KUS

15:20 – 15:40
PAPER TITLE : SIN-DUS CLASS OF DISTRIBUTIONS: THEORY AND APPLICATIONS ON TIMES TO INFECTION OF KIDNEY DIALYSIS PATIENTS
AUTHOR(S) : **Shakila BASHIR** , Ishmal SHAHZADI

15:40 – 16:00
PAPER TITLE: THE UNIT SUJIT DISTRIBUTION: A ONE-PARAMETER MODEL FOR PROPORTIONS WITH CLOSED-FORM RELIABILITY AND AFT REGRESSION
AUTHOR(S) : **Coşkun KUS** , Nihan Kübra KABAKBAS OKSUZ

16:00 – 16:20
PAPER TITLE : AGEING PATTERN IN INDIA: AN APPLICATION OF STABLE POPULATION THEORY
AUTHOR(S) : **Acharya BALKRISHNA** , K.N.S. YADAVA

16:20 – 16:40 : QUESTIONS & CERTIFICATE AWARDING

ICAS CONFERENCE
INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

Contents

On the numbers of components in each state in multi-state discrete k-out-of-n systems

Jakub Sadowy[1]
[1]Słowackiego 108/12, 82-300, Toruń, Polska, jsadowy@mat.umk.pl

Abstract

We consider multi-state discrete **k**-out-of-*n* systems, based on the definition introduced by Huang et al. [1] composed of components, which lifetimes are modeled by independent and identically distributed discrete random variables. We put a special emphasis on four-state **k**-out-of-*n* systems - that is systems with two intermediate states between perfect functioning and failure. We provide the joint distribution of the random vector representing the numbers of components in each state during system break down, as well the system's lifetime and expectation. We illustrate the theoretical results for four-state **k**-out-of-*n* systems with numerical examples concerning components, for which time spent in each state is geometrically distributed. The main results are a generalisation of the work on three-state systems by Goroncy and Jasniński [2] and have been sent for publication, undergoing review [3].

Key Words: reliability theory, coherent systems, multi-state systems, k-out-of-n systems, order statistics,

References

[1] Huang, J., Zuo, M. J., Wu, Y. (2000). Generalized multi-state k-out-of-n:G systems. *IEEE Transactions on Reliability*, 49(1), 105–111.
[2] Goroncy A., Jasiński K. (2025). Discrete time three-state k-out-of-n system's failure and numbers of components in each state. *J. Comput. Appl. Math.* 457, 116255.
[3] Goroncy A., Jasiński K., Sadowy J. (2025), Joint component-state distributions at system failure In discrete-time multi-state k-out-of-n systems, in review.

Finite-Step Bounds for Iterated Correlation Dynamics

Ishrak ALHAJJ HASSAN[1]

[1]University of Ostrava, Chittussiho 10, 710 00 Ostrava-Slezská Ostrava, Czech Republic;
Ishrak.Alhajj.s01@osu.cz

Abstract

Iterated Pearson correlation matrices arise in classical clustering constructions and association-plot methods [2]. Although the iteration converges empirically, asymptotic convergence does not describe finite-step behavior: ratios between successive update magnitudes may fluctuate and may exceed one along a convergent trajectory.

We investigate this finite-step behavior in the stabilized regime obtained after discarding the initial transient part of the iteration. Update magnitudes are measured in the Frobenius norm, and recent empirical work establishes universal, dimension-independent laws for the dynamics, including a conditional law governing the dependence of successive update ratios on the current update magnitude [1]. Treating this empirical law as given, we construct upper bounds for the next update ratio as a function of the current update magnitude, with a fixed probability under random initialization.

The bounds are defined as an ordered family. A baseline bound is obtained from empirical conditional quantiles of logarithmic update ratios, using logarithmic binning of the normalized update magnitude. This bound is enlarged by a uniform multiplicative factor to account for finite-sample quantile effects [3], and a further tolerance dilation [4] preserves the same dependence on update magnitude. The resulting bounds are dimension-uniform and are validated using independent trials, providing finite-step probabilistic bounds not captured by asymptotic convergence results.

***Key Words*:** *Iterated correlation matrices; Empirical laws; Local contraction analysis; Finite-step dynamics; Probabilistic bounds*

References

[1] AlHajj Hassan, I. (2025). *Empirical laws for iterated correlation matrices*. https://arxiv.org/abs/2512.15421
[2] Chen, C.-H. (2002). *Generalized association plots: Information visualization via iteratively generated correlation matrices*. Statistica Sinica, 12, 7–29.
[3] Taleb, N. N., & Cirillo, P. (2025). The Regress of Uncertainty and the Forecasting Paradox. Risks, 13(12), 247. https://doi.org/10.3390/risks13120247.
[4] Zadeh, L. A. (1965). *Fuzzy sets*. Information and Control, 8, 338–353.

A Novel Imputation-Based Approach for Synthetic Minority Oversampling in Imbalanced Classification

Pornthera Aimrod[1], Anamai Na-udom[2], Jaratsri Rungrattanaubol[3]

[1]Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, Thailand, 65000
porntheraa67@nu.ac.th
[2]Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, Thailand, 65000
anamain@nu.ac.th
[3]Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, Thailand, 65000
jaratsrir@nu.ac.th

Abstract

Class imbalance challenges supervised classification, often biasing models against minority classes. This issue spans domains like medical diagnosis, financial risk, and fault detection. Simple oversampling and Synthetic Minority Over-sampling Technique (SMOTE), which interpolates features to generate artificial data, are common remedies, but SMOTE can introduce noise, overlap, or unrealistic samples.

In this paper, we introduce a novel imputation-based synthetic minority oversampling approach that reframes synthetic sample generation as a structured missing-data imputation problem, rather than relying on traditional interpolation. The proposed method strategically masks feature values of minority class instances under a Missing Completely at Random (MCAR) mechanism, then imputes the masked values to generate synthetic samples. This reframing enables the existence of more diverse and statistically plausible minority class samples. We investigate three imputation techniques: Regression Imputation (RI), Expectation–Maximization Imputation (EMI), and Random Forest Imputation (RFI).

The proposed oversampling approach is evaluated using average balanced accuracy within a stratified five-fold cross-validation framework, employing Decision Trees (DT), Logistic Regression (LR), and Support Vector Machines (SVM) as base classifiers. Experiments are conducted on three benchmark datasets with varying imbalance ratios to assess the robustness of the model across different class distributions. Experimental results demonstrate that RFI consistently outperforms simple oversampling, SMOTE, RI, and EMI, based on Kendall's ranking across DT, LR, and SVM classification models on three datasets. Overall, these findings suggest that the proposed imputation-based synthetic minority oversampling approach offers an effective and robust solution for enhancing classification performance on imbalanced datasets.

Key Words: *Expectation–Maximization Imputation; Regression Imputation; Random Forest Imputation; Decision Trees; Logistic Regression; Support Vector Machines*

MODELING OCCUPATIONAL ACCIDENT FREQUENCIES IN AN INDUSTRIAL SECTOR USING NEGATIVE BINOMIAL REGRESSION MODEL

Pelin TOKTAŞ

Başkent University, Faculty of Engineering, Industrial Engineering Department, Ankara, 06790, Türkiye
e-mail: ptoktas@baskent.edu.tr

Abstract

This study examines occupational accidents that occurred in an industrial sector between 2019 and 2023. The main objective of the study is to determine how the frequency of occupational accidents changes over time and across different categories, and the aim is for these findings to contribute to the development of occupational health and safety strategies.

The dataset was compiled from multi-year statistical records belonging to the industry sector in question and converted into a long format suitable for statistical modeling. Each observation year includes information on the day of the week, time range, and type of accident (dangerous / very dangerous). The number of accidents observed as the dependent variable was used, while the independent variables consisted of temporal and categorical indicators.

The analysis results show that work accidents are not randomly distributed over time but rather exhibit systematic patterns during specific times. On weekdays, the number of work accidents was generally found to be higher compared to weekends, and a significant decrease in accident rates was observed on days when work intensity decreased. When examining the time intervals throughout the day, it was found that the accident frequency increased during peak production times in the morning and afternoon. Additionally, it has been determined that dangerous accidents occur more frequently than dangerous accidents. Between the years, an increasing trend in accident frequency was observed.

These findings indicate that occupational accidents are related to temporal factors and that the risk increases during specific periods. The results obtained emphasize the importance of considering time-based risk analyses in the development of occupational health and safety policies. Recognizing heightened risks, particularly on certain days and times, facilitates the more precise planning of preventive strategies.

In conclusion, negative binomial regression has proven to be an appropriate technique for the analysis of occupational accident data characterized by overdispersion. The model provides an evidence-based contribution applicable to industrial occupational safety management by revealing the time- and category-based changes in accidents.

Key Words: *Work accidents; Negative Binomial Regression; Occupational Health and Safety; Count Data.*

References

[1] Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
[2] Yang, S., & Berdine, G. (2015). The negative binomial regression. *The Southwest respiratory and critical care chronicles, 3*(10), 50-54.
[3] El-Basyouny, K., & Sayed, T. (2006). Comparison of two negative binomial regression techniques in developing accident prediction models. *Transportation research record*, *1950*(1), 9-16.
[4] Cameron, A. C. & Trivedi, P. K. 1998. *Regression Analysis of Count Data*. New York: Cambridge Press.
[5] Introduction to SAS. UCLA: Statistical Consulting Group. from https://stats.oarc.ucla.edu/spss/dae/negative-binomial-regression/ (accessed November 10, 2025).

ROBUST PARAMETER ESTIMATION APPROACH FOR UNIVERSAL KRIGING UNDER CONTAMINATED SPATIAL DATA

Vural YILDIRIM[1], Ezio TODINI[2], Yeliz MERT KANTAR[3]

[1]Eskisehir Technical University, Institute of Earth and Space sciences, Eskisehir, Türkiye, vural.yildirim@gmail.com
[2]Italian Hydrological Society, 40127 Bologna, Italy, presidente-onorario@sii-ihs.it
[3]Eskisehir Technical University, Faculty of Science, Department of Statistics, Eskisehir, Türkiye, ymert@eskisehir.edu.tr

Abstract

Spatial estimation problems are frequently addressed using the universal kriging framework, which extends classical kriging by combining a deterministic trend component with a stochastic residual process. In practice, the unknown parameters associated with the trend and spatial covariance structure are typically estimated using traditional techniques such as maximum likelihood estimation or weighted least squares. Despite their widespread use, it is well known that these classical methods are highly sensitive to the presence of outliers. Even a small number of outlying observations can significantly impact the estimation of both trend parameters and the variogram model, resulting in unreliable spatial predictions.

To address this issue, this study presents a robust parameter estimation approach specifically designed for universal kriging models. The proposed method integrates the principles of M-estimation into a likelihood-based estimation framework, thereby enabling the control of the effect of outliers via a bounded influence function. Within this framework, the likelihood function has been reformulated to derive robust estimation equations for both regression coefficients and variogram parameters.

An iterative estimation algorithm is employed to estimate the trend component and variogram parameters. The algorithm operates alternately between trend estimation and variogram parameter estimation until convergence is achieved. The performance of the proposed estimator has been evaluated through comprehensive Monte Carlo simulation experiments, taking into account different trend structures, variogram characteristics, and levels of data contamination. The results demonstrate that the robust approach provides more stable parameter estimates than traditional estimation methods in the presence of outliers and consistently produces lower mean squared prediction errors.

Keywords: *Universal kriging; spatial trend; robust estimation, outlier resistance.*

References

[1] Cressie N (1993) Statistics for spatial data, Revised edn. Wiley, New York.

[2] Yildirim V, Kantar YM (2020) Robust estimation approach for spatial error model. J Stat Comput Simul 90(9):1618–1638.

[3] Yildirim V (2026) Robust likelihood–based estimation of spatial and spatio–temporal variograms in ordinary kriging. PhD Thesis, Eskişehir Technical University, Eskişehir, Türkiye.

A MULTI-LAYER STATISTICAL FRAMEWORK FOR INTEGRATED AND INTERPRETABLE STRUCTURES IN LARGE-SCALE TEXTUAL CORPORA, WITH APPLICATION TO PERMACULTURE DISCOURSE

Alessandra MERIANI[1], Ida D'ATTOMA,
[1] Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy
[1]E-mail: alessandra.meriani2@unibo.it
[2] Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy
[2] E-mail: ida.dattoma2@unibo.it

Abstract

The increasing availability of large-scale unstructured textual data requires statistical approaches capable of integrating multiple sources of information into coherent analytical outputs. While statistical modelling of textual data enables the extraction of lexical distributions, relational structures, affective signals, and latent thematic patterns, these representations are typically analysed separately, leading to fragmented interpretations and limited interpretability in high-dimensional settings.

This study proposes a multi-layer statistical framework for the integrated analysis of large-scale textual corpora. The approach combines complementary statistical techniques within a unified analytical structure. After corpus construction through automated web scraping, standard preprocessing procedures — including tokenisation, lemmatisation, and part-of-speech filtering — are applied following established protocols in computational text analysis [1, 2] while incorporating implementation choices adapted to the specific analytical requirements of the study. The resulting corpus is validated through lexical distribution diagnostics, including adherence to Zipf's law [3], and dimensionality reduction techniques are used to mitigate sparsity [4].

The analytical framework integrates multiple statistical layers. Lexical distributions are examined through normalized term frequencies, while affective signals are captured using multi-lexicon sentiment scoring. Relational structures are identified through correlation measures and co-occurrence statistics, including Log-Likelihood Ratio, Mutual Information, and Dice coefficients. Latent thematic structures are estimated via probabilistic topic modelling, and hierarchical clustering is employed to capture higher-level structural organization, with cluster validity assessed through stability and cohesion diagnostics [5, 6].

Rather than treating these components as independent outputs, the proposed framework introduces cross-metric alignment and consistency diagnostics to integrate heterogeneous statistical representations and identify patterns that remain stable across analytical layers. Interpretability is operationalized as a structured validation process, in which results are assessed through the convergence of statistical signals across lexical, relational, and latent dimensions.
The framework is applied to a corpus of 496 documents related to European permaculture discourse, conceptualized as a transformative grassroots niche within agricultural systems. Results indicate the emergence of coherent thematic structures centred on ecological practices, regeneration, and community-oriented dynamics, supported by predominantly positive affective signals (e.g., trust, anticipation). At the same time, variations in relational and thematic patterns reveal heterogeneous representations of sustainability-related concepts, illustrating how integrated statistical modelling can uncover structured and interpretable patterns in complex textual data.

Key Words: *Statistical text analysis; multi-layer modelling; permaculture; high-dimensional data; topic modelling*

References

[1] Antons D, Grünwald E, Cichy P, Salge O (2020) The application of text mining methods in innovation research: Current state, evolution patterns, and development priorities. R&D Management 50:329–351.

[2] Farrell MJ, Le Guillarme N, Brierley L, Hunter B, Scheepens D, Willoughby A, Yates A, Mideo N (2024) The changing landscape of text mining: A review of approaches for ecology and evolution. Proceedings of the Royal Society B 291:20240423.

[3] Zipf GK (1949) Human behavior and the principle of least effort: An introduction to human ecology. Addison-Wesley Press.

[4] Bolasco S (2022) Analisi multidimensionale dei dati: Metodi, strategie e criteri d'interpretazione. Carocci Editore.

[5] Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet Allocation. Journal of Machine Learning Research 3:993–1022.

[6] Deveaud R, SanJuan E, Bellot P (2014) Accurate and effective latent concept modeling for ad hoc information retrieval. Document Numerique 17:61–84.

MULTI-STEP TEMPERATURE FORECASTING WITH AN ATTENTION-BASED GATED RECURRENT UNIT MODEL

Vildan KARA[1], Gülesen ÜSTÜNDAĞ ŞİRAY[2],

[1]Cukurova University, Adana, Türkiye, vildankara@cu.edu.tr
[2] Cukurova University, Adana, Türkiye, gustundag@cu.edu.tr

Abstract

Temperature forecasting is a crucial problem for climate analysis, energy planning, and environmental decision support systems due to the nonlinear structure and long-term dependencies inherent in time series data [1]. In recent years, deep learning–based models have been increasingly adopted in time series forecasting to capture complex temporal relationships. Within this context, attention mechanisms have become increasingly important in allowing models to focus on the most informative temporal patterns, thereby enhancing forecasting performance [2][3]. Accordingly, this study proposes a deep learning framework based on a gated recurrent unit architecture enhanced with an attention mechanism for temperature forecasting using multivariate meteorological data [4].

The data set consists of daily temperature, wind speed, humidity, and surface pressure variables obtained from the NASA POWER database for a specific geographical location over the period from 2020 to 2024 [5]. To assess the effectiveness of the proposed approach, a comparative analysis is conducted with commonly utilized deep learning models reported in the literature, namely long short-term memory and gated recurrent unit architectures.

The comparative analysis is conducted under a multi-step forecasting framework considering prediction horizons of 7, 14, and 30 days. Model accuracy is evaluated using two metrics, which are Root Mean Squared Error and Mean Absolute Error. The empirical results demonstrate that the attention-based gated recurrent unit model achieves lower error rates across all forecasting horizons, exhibiting superior predictive performance compared to the baseline models.

In conclusion, this study indicates that the attention-based gated recurrent unit model provides an effective and reliable approach for temperature forecasting using meteorological time series, outperforming commonly used basic deep learning models.

Key Words: *Attention mechanism; Deep learning; Gated Recurrent Unit; Meteorological time series; Temperature forecasting*

References

[1] Shen J, Wu W, Xu Q (2024) Accurate prediction of temperature indicators in eastern China using a multi-scale CNN–LSTM–attention model. arXiv preprint arXiv:2412.07997.

[2] Wen, X., & Li, W. (2023). Time series prediction based on LSTM-attention-LSTM model. *IEEE access*, *11*, 48322-48331.

[3] Wu, S. S., Du, Z. H., Zhang, F., Zhou, Y., & Liu, R. Y. (2023). Time-Series Forecasting of Chlorophyll-a in Coastal Areas Using LSTM, GRU and Attention-Based RNN Models. *Journal of Environmental Informatics*, *41*(2).

[4] Munir, H. S., Ren, S., Mustafa, M., Siddique, C. N., & Qayyum, S. (2021). Attention based GRU-LSTM for software defect prediction. *Plos one*, *16*(3), e0247444.

[5] NASA POWER Project (2024) NASA Prediction of Worldwide Energy Resources (POWER). Available at: https://power.larc.nasa.gov

Impact of Imputation on Logistic Classification with MCAR and MAR Missing Data in Cardiovascular Dataset

Ratchaneewan Paisanwarakiat[1], Anamai Na-udom[2], Jaratsri Rungrattanaubol[3]

[1]Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, Thailand, 65000
ratchaneewanpa64@nu.ac.th
[2]Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, Thailand, 65000
anamain@nu.ac.th
[3]Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, Thailand, 65000
jaratsrir@nu.ac.th

Abstract

Missing data is a persistent challenge in medical datasets, particularly in cardiovascular research, where incomplete records can reduce model accuracy and affect clinical decision-making. Although different missing data mechanisms may alter the performance of imputation techniques, their behavior across varying levels of missingness remains insufficiently explored. This challenge affects the ability to identify the most effective imputation strategy, which is essential for ensuring reliable classification performance.

This study aims to identify the most effective imputation technique based on the performance of a Logistic Regression classification model. It explores two types of missing data mechanisms: Missing Completely at Random (MCAR) and Missing at Random (MAR). The analysis examines various levels of missing data, ranging from 10% to 30%, using a cardiovascular dataset. We evaluated four primary imputation techniques: Deterministic Linear Regression (DLR), Expectation-Maximization (EM), Random Forest (RF), and K-Nearest Neighbors (KNN). Additionally, we developed six hybrid approaches, resulting in a total of ten techniques. The accuracy of the Logistic Regression classification models, assessed through 10-fold cross-validation, is used to evaluate the effectiveness of the imputation techniques.

The results show that under MCAR, the hybrid techniques DLRRF and RFEM achieved the highest accuracy, followed by RFKNN. This indicates that hybrid imputation methods improve model learning and classification performance when data is missing at random. In contrast, under MAR, the EM and DLREM techniques achieved the highest accuracy, with DLR close behind. This suggests that statistical estimation methods like EM, especially when paired with DLR, are more effective for recovering data structure and enhancing model performance when missingness is related to variable relationships. The findings offer insights into the effectiveness of imputation strategies for missing data, particularly for MCAR and MAR, enhancing decision-making reliability in clinical data analysis.

Key Words: *Missing Completely at Random; Missing at Random; Hybrid Imputation Techniques; Logistic Regression Model*

Computational Procedures for the Estimation of the Extremal Index

Dora PRATA GOMES[1], M. Manuela NEVES[2],

[1]Center for Mathematics and Applications (NOVA Math) and Department of Mathematics, NOVA School of Science and Technology (NOVA FCT), Portugal and dsrp@fct.unl.pt
[2]CEAUL–Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, Portugal and manela@isa.ulisboa.pt

Abstract

Extreme observations often arise in clusters over time, reflecting persistence in the underlying stochastic mechanism and inducing serial dependence among exceedances. This behavior has important implications for risk assessment. Within the framework of Extreme Value Theory, the extremal index is a key parameter that quantifies the degree of clustering of extreme events. Although a substantial body of literature addresses its estimation, practical challenges remain, particularly regarding the selection of the number of upper order statistics used in the estimation procedure. In this work, we investigate bias-reduced estimators based on the Generalized Jackknife methodology. A stability-based algorithm for extreme value parameter estimation, originally proposed in [1] and further developed in [2], has shown strong empirical performance. We explore how this procedure can be extended to guide the choice of a tuning parameter in the Generalized Jackknife scheme, once an optimal threshold—determined according to an appropriate criterion—for estimating the extremal index has been identified, as discussed in [3]. The finite-sample behavior of the estimators is evaluated through an extensive Monte Carlo simulation study, including scenarios with temporal dependence. The results suggest that the Generalized Jackknife approach provides an effective bias reduction and that, with a suitable selection of the tuning parameter, the proposed procedure yields more stable and reliable estimates of the extremal index.

Key Words: *Adaptive algorithm; Bias reduction; Extreme value theory; Extremal index; Generalized Jackknife*

References

[1] Gomes MI, Henriques-Rodrigues L, Fraga Alves MI, Manjunath, BG (2013) Adaptive PORT–MVRB estimation: an empirical comparison of two heuristic algorithms. Journal of Statistical Computation and Simulation 83(6): 1129–1144.
[2] Neves MM, Gomes MI, Figueiredo F, Gomes, DP (2015) Modelling Extreme Events: Sample Fraction Adaptive Choice in Parameter Estimation. Journal of Statistical Theory and Practice, 9(1):184–199.
[3] Prata Gomes D, Neves MM (2020) Resampling Procedures for a More Reliable Extremal Index Estimation. In Data Analysis and Applications 4 (eds A. Makrides, A. Karagrigoriou and C.H. Skiadas).

REVISITING THE EXTREME VALUE INDEX: THE PROBABILITY WEIGHTED MOMENTS FRAMEWORK FOR HEAVY-TAILED DISTRIBUTIONS

Ayana MATEUS[1], Frederico CAEIRO[2],

[1]Center for Mathematics and Applications (NOVA Math), Portugal, and NOVA School of Science and Technology (NOVA FCT), Campus de Caparica, Portugal, amf@fct.unl.pt
[2]Center for Mathematics and Applications (NOVA Math), Portugal, and NOVA School of Science and Technology (NOVA FCT), Campus de Caparica, Portugal, fac@fct.unl.pt

Abstract

The statistics of extreme values seeks to model and characterize the tail behavior of probability distributions, where rare but impactful events occur. This field has gained significant prominence in recent decades due to its critical applications in risk assessment, reliability analysis, and climate science. Accurate estimation of the Extreme Value Index (EVI) is essential for understanding the tail heaviness of stochastic processes, particularly when modeling the frequency and magnitude of extreme events. This research revisits the Probability Weighted Moment framework for EVI estimation. We examine two classes of estimators for a positive extreme value index: the classical Probability Weighted Moments and the Generalized Probability Weighted Moments (GPWM) estimators. Notably, the GPWM framework is shown to be highly versatile, encompassing the well-known Hill estimator as a special case. We establish the non-degenerate limiting distributions for both classes of estimators, demonstrating that their tuning parameters can be strategically adjusted to manage the fundamental trade-off between bias and variance. A method for selecting the optimal threshold is presented, based on the minimization of the Asymptotic Mean Squared Error. To evaluate the finite-sample performance of these classes of EVI estimators, a Monte-Carlo simulation study was conducted across various sample sizes and several distinct distribution models. Furthermore, the study explores the practical utility of these estimators in calculating high quantiles and exceedance probabilities. By providing a more stable foundation for tail inference, this research contributes to the development of robust methodologies for extreme value analysis, with significant implications across a wide range of applied scientific fields.

Key Words: *Extreme Value Index, Hill Estimator, Monte-Carlo Simulation, Order Statistics Probability Weighted Moments.*

References

[1] Greenwood, J. A., J. M.Landwehr, N. C.Matalas, and J. R.Wallis (1979), Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form, Water Resour. Res., 15(5), 1049–1054.
doi:10.1029/WR015i005p01049.

[2] Chen, H., Cheng, W., Zhao, J., & Zhao, X. (2017). Parameter estimation for generalized Pareto distribution by generalized probability weighted moment-equations. Communications in Statistics-Simulation and Computation, 46(10), 7761–7776.
https://doi.org/10.1080/03610918.2016.1249884

A REFINED BIAS-CORRECTED ESTIMATOR FOR THE SHAPE PARAMETER OF THE LOG-LOGISTIC MODEL

Frederico CAEIRO[1], Ayana MATEUS[2],

[1]Center for Mathematics and Applications (NOVA Math), Portugal, and NOVA School of Science and Technology (NOVA FCT), Campus de Caparica, Portugal, fac@fct.unl.pt
[2]Center for Mathematics and Applications (NOVA Math), Portugal, and NOVA School of Science and Technology (NOVA FCT), Campus de Caparica, Portugal, amf@fct.unl.pt

Abstract

In the literature, the log-logistic distribution is often presented with two parameters: one governing the shape of the model, the other governing its scale. To make this model more suitable for the data, an additional location parameter can be added, yielding the three parameters or shifted log-logistic model. Regarding parameter estimation, Balakrishnan et al. [1] derived the best linear unbiased estimators for the location and scale parameters. This paper deals with a new method for estimating the shape parameter of a log-logistic distribution. We propose an estimator based on the largest order statistics. The theoretical and simulated results confirm the significantly better performance of the new estimation method, relatively to other methods from the literature.

Key Words: *Bias correction; Log-logistic model; Largest Order Statistics, Monte-Carlo simulation; Parameter estimation.*

References

[1] Balakrishnan, N., Malik, H., Puthenpura, S. (1987). Best linear unbiased estimation of location and scale parameters of the log-logistic distribution. Communications in Statistics - Theory and Methods 16(12), 3477-3495.

Mixed-Model Inference for Pairwise Treatment Contrasts with Judgement Covariates in Randomized Block Designs

Omer Ozturk

Department of Statistics, The Ohio State University, USA

And

Olena Kravchuk

The Biometry Hub, School of Agriculture, Food and Wine, Adelaide University, Australia

Abstract

This study introduces an innovative statistical inference procedure for estimating pairwise differences in treatment means within a generalized randomized complete block design (RCBD) featuring both fixed and random block effects. Unlike traditional methods, our approach leverages post-experimental information to enhance precision. Following data collection, experimental units (EUs) are ranked based on the predictive interaction effects and error terms of paired EUs within each block.

These ranked units are then partitioned into two distinct strata, which are integrated into the model as auxiliary blocking factors. This stratification effectively captures the relative magnitudes of interaction and error terms that are often lost in standard Least Squares (LS) estimations. Based on this post-stratified framework, we construct new estimators for pairwise treatment differences and develop a robust multiple comparison procedure for all pairwise contrasts.

Our results demonstrate that this procedure significantly outperforms the conventional least squares method, yielding higher efficiency and narrower confidence intervals. The practical utility of the proposed method is validated through an application to field experimental data, showcasing its potential for more sensitive hypothesis testing in agricultural and biological research.

Key words: Contrast estimates; randomized complete block design; judgment post-stratification; Tukey multiple comparison; random effect model, mixed effect model.

Efficiency Analysis of Turkish Airports: A Stochastic Frontier Analysis Approach

Erdem Korhan AKÇAY[1], Yeliz MERT KANTAR[2],

[1]erdemkorhan.akcay@ogu.edu.tr
[2]ymert@eskisehir.edu.tr

Abstract

This study examines the technical efficiency of 21 airports operating in Türkiye using Stochastic Frontier Analysis (SFA), a parametric efficiency measurement method. Airport efficiency has traditionally been assessed using both parametric and non-parametric approaches; this study prefers SFA due to its critical methodological advantage in terms of capacity across decision-making units in the context of developing economies.

The analysis was conducted using data obtained from the GDSAA (General Directorate of State Airports Authority) statistical yearbook. The dataset covers both operational and financial dimensions. Airport output is represented by three variables: passenger traffic, total cargo volume, and non-aviation revenues, capturing the multidimensional production structure of airports and addressing the limitations of single output specifications based solely on passenger movements. On the input side, the model includes capital investment, labor costs, airport terminal area, total number of personnel, number of ticket counters, and average monthly landings. The number of ticket counters is included as an indicator of service capacity at the passenger interface level, reflecting the airport's ability to manage transaction volume beyond just physical infrastructure. Regional population is also included as an environmental variable representing demand-side conditions beyond managerial control; this allows the observed efficiency differences to reflect operational performance rather than external market conditions.

The results reveal heterogeneity in the efficiency of Turkish airports. The findings identify both structural and operational factors associated with efficiency differences among airports and have direct implications for prioritizing infrastructure investments and rationalizing resource allocation in the civil aviation sector.

Key Words: *Stochastic Frontier Analysis; Aiport; Efficiency*

Imaginative Geographies of Sustainability: A Structural Equation Modeling Analysis of the Role of Social Media and Attitudes in Generation Z's Destination Imaginations

Serkan OLGAÇ[1], Semra GÜNAY[1], Almira SEISINBINOVA [2]

[1] Anadolu University Tourism Faculty of Tourism Management Department, Eskişehir, Türkiye

[2] L.N. Gumilyov Eurasian National University, Management Department, Astana, Kazakhstan

Abstract

Sustainable tourism has become an increasingly important concept in tourism research in recent years [1]. However, individuals' perceptions of sustainable tourism are often based not only on objective information but also on mental images and media representations [2]. Therefore, the perception of sustainable tourism is a complex structure constructed by mental images and media representations beyond objective criteria. In this context, the aim of this study is to examine the perceptions of Generation Z towards sustainable tourism destinations using a quantitative and multivariate approach within the framework of the Imaginative Geography concept [3], and to reveal the key factors shaping these perceptions. The research statistically analyzes the predictive effects of social media influence, attitudes towards sustainable tourism, and non-informative assessments on Generation Z's perception of sustainable destinations, which is the dependent variable.

The research data was collected from university students representing Generation Z through a survey form. In the analysis of the data, the Partial Least Squares Structural Equation Modeling (PLS-SEM) method was preferred due to its advantages in modeling multidimensional and complex causal relationships between variables, and the analyses were performed using SmartPLS software [4]. Within the framework of the research questions, participants were investigated in terms of which destinations they associated with sustainable tourism and whether these perceptions were based on factual information or mental imaginations. In the statistical evaluation process, factor loadings, Cronbach's Alpha, Composite Reliability (CR), and Explained Mean Variance (AVE) values were examined for the validity of the measurement model; and path coefficients were used to test the structural model. Data used included coefficients , significance levels (p-values), and explanatory power (R^2).

Research findings are expected to reveal that the influence of social media (H1), attitudes towards sustainable tourism (H2), and intuitive evaluations not based on rational information (H3) positively and significantly affect Generation Z's perceptions of destinations. The findings are also expected to show that Generation Z's perceptions of sustainable tourism destinations are influenced not only by environmental awareness but also by media representations and mental constructs. In this respect, the study aims to provide a methodological contribution to the literature by explaining perceptions of sustainable tourism through statistical modeling techniques and the "imaginative geographies" theory; and to provide strategic implications for understanding the perceptual dynamics of younger generations in terms of sustainable tourism policies and destination marketing.

Key Words: *Structural Equation Modeling, Sustainable tourism, Imaginative Geography; Social science, SmartPLS software*

References

[1] Loureiro, I. S., Costa, V., & Malheiro, A. (2024). Sustainable tourism: A systematic literature review and bibliometric analysis. European Journal of Tourism, Hospitality and Recreation. https://doi.org/10.2478/ejthr-2024-0009

[2] Salazar, N. B. (2012). Tourism imaginaries: A conceptual approach. *Annals of Tourism Research*, 39(2), 863-882. https://doi.org/10.1016/j.annals.2011.10.004

[3] Günay, S.; Ateş Akkaya, D.; Akgiş İlhan, Ö. Imagined Geographies of Sustainability: Rethinking Responsible Tourism Consumption Through the Utopias of Generation Z. *Sustainability* 2025, *17*, 10280. https://doi.org/10.3390/su172210280

[4] Joseph F. Hair, G. Tomas M. Hult, Christian M. Ringle, & Marko Sarstedt (2021). A primer on partial least squares structural equation modeling (PLS-SEM) (3rd ed.). Sage Publications.

Robust density power divergence based Bayesian analysis of non-destructive one-shot device testing data

Shuvashree Mondal[1]

[1] Department of Mathematics and Computing, IIT (ISM) Dhanbad, Dhanbad, India, shuvasri29@iitism.ac.in

Abstract

The prevalence of one-shot devices is quite prolific in engineering and medical domains. Unlike typical one-shot devices, non-destructive one-shot devices (NOSD) may survive multiple tests and offer additional data for reliability estimation. The present study aims to implement Bayesian approach in analysis of highly durable non-destructive one-shot device testing data under a step-stress accelerated life testing (SSALT) experiment applying a cumulative risk model (CRM). In a step-stress accelerated life testing experiment, cumulative risk model retains the continuity of hazard function by allowing the lag period before the effects of stress change emerge. However, with small deviations from the assumed model conditions, conventional likelihood based Bayesian estimation may result in misleading statistical inference, raising the need for a robust Bayesian method. This work exploits a robust Bayesian estimation method where the density power divergence measure substitutes the likelihood in posterior density function. Further, testing of hypotheses is performed through the robust Bayes factor. Moreover, we assess the extent of resistance of the proposed methods to small deviations from the assumed model conditions by applying the influence function approach. In testing of hypothesis, influence function reflects how outliers impact the decision-making through Bayes factor under null hypothesis. Finally, the theoretical developments are validated through the extensive simulation study.

Key Words: Cumulative risk model; density power divergence; influence function; one-shot device; robust Bayes estimation

WEIGHTED NON-STOCHASTIC FORECASTING RANGE OF TIME SERIES

Richard HINDLS[1], Stanislava HRONOVÁ[2], Luboš MAREK[3]

[1]Prague University of Economics and Business, Czech Republic, hindls@vse.cz
[2]Prague University of Economics and Business, Czech Republic, hronova@vse.cz
[3]Prague University of Economics and Business, Czech Republic, marek@vse.cz

Abstract

The article examines the possibilities of using non-stochastic statistical forecasting to predict the short-term development of macroeconomic indicators within the system of national accounts, particularly GDP. Forecast construction is traditionally approached probabilistically, i.e., by building confidence intervals. However, this conventional procedure usually results in excessively wide prediction intervals, which either lack practical applicability (precisely because of their large width) or struggle to meet probabilistic assumptions, especially for time series of economic indicators. Constructing non-stochastic forecasts using a non-probabilistically derived forecast range offers an alternative technique that, among other benefits, more effectively mitigates the influence of the *ceteris paribus* principle—a limitation on which many commonly used stochastic approaches fail. At the same time, it significantly narrows the width of the interval forecast, thereby increasing its informative value.

The concept of a non-stochastic range is based, in its first phase, on constructing a large number of different models for the same time series, from which an average point forecast of short-term development is then derived. An interval forecast is subsequently created from these averaged point forecasts (which, naturally, produce different forecast values individually). The novelty introduced in our paper lies in applying a stricter evaluation of the individual primary point forecasts that arise from the summary of different time series models and are then averaged. This stricter approach consists of "weighting" the results of interpolation models according to their success in describing historical data. As is well known, models can describe the past to varying degrees, but due to the *ceteris paribus* principle, this does not guarantee successful prediction. Nevertheless, there is a certain connection between a model's interpolation quality and its extrapolation performance. Therefore, when averaging point forecasts, these values are weighted using the mean squared error, which reflects the model's ability to describe past developments in the time series. This improves the quality of the averaged forecast, from which the non-stochastic forecast range is then derived.

To illustrate this method, the paper uses annual GDP development indices for the Czech Republic as a time series with a sufficiently large number of observations and relatively low variability. The data are seasonally and calendar-adjusted and capture the economic crisis periods of the past 20 years (the 2008–2009 global financial crisis, the 2020–2022 Covid crisis, and the 2022–2023 energy crisis), which increases the demands on the robustness of the proposed procedure.

***Keywords*:** *point forecast; weighted non-stochastic forecast range; interval forecast; statistical forecastin;, time serie..*

References

[1] Eppelbaum LV (2013) Non-stochastic long-term prediction model for US tornado level. *Natural Hazards* 69: 2269-2278. DOI: 10.1007/s11069-013-0787-7.

[2] Hui L (2020) *Smart Cities: Big Data Prediction Methods and Applications*. Springer Nature Singapore Pte Ltd. and Science Press.

[3] Marek L, Hindls R, Hronová S (2021) Non-Stochastic Argumentation in Predicting Economic Indices *Statistika*: *Statistics and Economy Journal* 101(1): 5-23.

[4] Ryan O, Haslbeck JMB, Waldorp LJ (2025) Non-Stationarity in Time-Series Analysis: Modeling Stochastic and Deterministic Trends. *Multivariate Behavioral Research* 60(3): 556-588. https://doi.org/10.1080/00273171.2024.2436413

[5] Yachao Z, Kaipei L, Liang Q, Xueli A (2016) Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods. *Energy Conversion and Management* 112: 208-219. DOI:10.1016/j.enconman.2016.01.023.

BIVARIATE BINOMIAL DISTRIBUTION AND BIVARIATE BERNSTEIN POLYNOMIALS

Ayşe BELER[1], İsmihan BAYRAMOĞLU[2],

[1]Izmir University of Economics, e-mail: ayse.beler@ieu.edu.tr
[2] Izmir University of Economics, e-mail: ismihan.bayramoglu@ieu.edu.tr

We consider a bivariate binomial distribution in a fourfold sampling scheme and introduce a bivariate Bernstein polynomial (BBP) [1]. For any continuously bivariate real function, we show that the bivariate Bernstein polynomial [2] defined on the basis of BBP uniformly converges to this function. Some illustrative examples and graphs are provided.

***Key Words*:** Bivariate binomial distribution; Bivariate Bernstein polynomials; Fourfold sampling scheme; Uniform convergence; Approximation theory

References

[1] Bairamov I, Gultekin OE (2010) Hacettepe Journal of Mathematics and Statistics 39(1):109-120

[2] Akhiezer NI (1956) Theory of Approximation. Ungar, New York

ENTROPY-BASED BAYESIAN SEM FOR STUDENT SATISFACTION MODELING

Aslı CANAL[1]

[1]1Kapadokya University, Department of Data Science and Analytics, Mustafapaşa, Ürgüp, Nevşehir, Türkiye
Email: asli.canal@kapadokya.edu.tr

Abstract

Student satisfaction has become a key performance indicator for higher education institutions, particularly in competitive private university environments. Understanding the multidimensional structure of satisfaction requires robust analytical frameworks capable of modeling complex relationships among institutional service factors.

This study proposes an entropy-based Bayesian Structural Equation Modeling (SEM) approach for analyzing student satisfaction. By integrating entropy measures into Bayesian prior specification and model evaluation processes, the proposed framework aims to enhance parameter stability and improve explanatory power under uncertainty.

The empirical application is based on student satisfaction survey data collected from undergraduate programs. Latent constructs such as teaching quality, academic support, infrastructure, administrative services, and career alignment are modeled to evaluate their direct and indirect effects on overall satisfaction.

Comparative analyses are conducted between classical SEM, Bayesian SEM, and the entropy-based Bayesian SEM framework. Findings are expected to demonstrate that the proposed model provides more robust parameter estimation and improved decision-support insights for institutional quality management.

The study contributes methodologically to Bayesian SEM literature while offering practical implications for higher education administrators seeking data-driven strategies to enhance student satisfaction.

Key Words: *Bayesian SEM, Entropy, Student Satisfaction, Higher Education, Decision Support.*

References

[1] Lee SY (2007) Structural Equation Modeling: A Bayesian Approach, Wiley.

[2] Muthén B, Asparouhov T (2012) Bayesian structural equation modeling: A more flexible representation of substantive theory, Psychological Methods.

[3] Shannon CE (1948) A mathematical theory of communication, Bell System Technical Journal.

[4] Zeleny M (1982) Multiple Criteria Decision Making, McGraw-Hill.

[5] Elliott KM, Shin D (2002) Student satisfaction: An alternative approach to assessing higher education quality. Journal of Higher Education Policy and Management.