

BOOK OF ABSTRACTS



9th International Conference on Advances in Statistics

ICAS CONFERENCE
INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

June 20-22 2023

<http://www.icasconference.com/>

ICAS'2023

9th 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: aioc.statistics@gmail.com

<http://www.icasconference.com>

ISBN:

Copyright @ 2023 AIOC and Authors

All Rights Reserved

No part of the material protected by this copyright may be reproduced or utilized in any form or by any means electronic or mechanical, including photocopying , recording or by any storage or retrieval system, without written permission from the copyrights owners.

INTERNATIONAL SCIENTIFIC COMMITTEE

Prof. Dr. Alexei Stepanov
Immanuel Kant Baltic Federal University – Russia

Prof. Dr. Ayman BAKLEEZI
Qatar University – Qatar

Prof. Dr. Barry C. ARNOLD
University of California, Riverside – USA

Prof. Dr. George YANEV
The University of Texas Rio Grande Valley – USA

Prof. Dr. Hamparsum BOZDOGAN
The University of Tennessee – USA

Prof. Dr. Hamzeh TORABI
Yazd University – IRAN

Prof. Dr. Ilham AKHUNDOV
University of Waterloo – Canada

Prof. Dr. Jorge NAVARRO
Universidad de Murcia – Spain

Prof. Dr. Jose Maria SARABIA
University of Cantabria – Spain

Prof. Dr. Leda MINKOVA
University of Sofia – Bulgaria

Prof. Dr. Narayanaswamy BALAKRISHNAN
McMaster University – Canada

Prof. Dr. Nikolai KOLEV
University of Sao Paulo – Brasil

Prof. Dr. Sarjinder SINGH
Texas A&M University-Kingsville – USA

Prof. Dr. Stelios PSARAKIS
Athens University of Economics & Finance – Greece

Prof. Valerii B. Nevzorov
St. Petersburg State University – Russia

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

Bogazici University – Turkey

ORGANIZATION 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

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

Assoc. Prof. Dr. Gulder KEMALBAY
Yıldız Technical University – Turkey

Assoc. Prof. Dr. Murat OZKUT
Izmir University of Economics, Turkey

Instructor PhD Ozlem BERAK KORKMAZOGLU
Yıldız Technical University – Turkey

Res. Asst. Ayse BELER
Izmir University of Economics, Turkey

Res. Asst. 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 9th International Conference on Advances in Statistics which will be organised online on dates between 20-22 June 2023.

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

We cordially invite prospective authors to submit their original papers to ICAS-2023,

- 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

20 JUNE 2023 TUESDAY

11:45 – 12:00

OPENING CEREMONY

Professor Ismihan BAYRAMOGLU / Izmir University of

Economics - Turkey

12:00 – 12:30	Keynote Speech: Professor José Fernando LOPEZ-BLAZQUEZ / University of Sevilla – Spain <i>Discrete q-Beta Distributions and Analogs of Order Statistics in Geometric Distributions</i>
---------------	--

12:30- 12:45	B R E A K
--------------	------------------

SESSION A

SESSION CHAIR	Krzysztof JASIŃSKI	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
12:45 - 13:00	Optimal Plan for Ordered Step-stress Stage Life Testing	Shuvashree MONDAL / Debashis SAMANTA / Debasis KUNDU
13:00 - 13:15	Sequences of random variables predicted with conditional expectations	Ismihan BAYRAMOGLU
13:15 - 13:30	On the status of component failures in a working coherent system under inspections	Krzysztof JASIŃSKI

13:30 - 13:35	B R E A K
---------------	------------------

13:35 - 13:55	Invited Speech: Professor Anna DEMBIŃSKA / Warsaw University of Technology – Poland <i>Likelihood Inference for Geometric Lifetimes of Components of k-out-of-n Systems</i>
---------------	--

13:55 - 14:00	B R E A K
---------------	------------------

SESSION B

SESSION CHAIR	Ozge KURAN	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
14:00- 14:15	Efficient Sampling from the PKBD Distribution	Lukas SABLICA / Kurt HORNIK / Josef LEYDOLD
14:15 - 14:30	Robust Estimation of Correlation Coefficient via Ranked Set Sampling: Application to Bivariate Normal Distribution	Yusuf Can SEVIL / Tugba OZKAL YILDIZ
14:30 - 14:45	A New Two-Parameter Biased Prediction Class in Linear Mixed Models	Ozge KURAN

14:45 – 14:50	B R E A K
---------------	------------------

14:50 - 15:20	<p style="text-align: center;">Keynote Speech: Professor Leda MINKOVA / Sofia University – Bulgaria <i>Compound Discrete Time Geometric Process</i></p>
---------------	---

15:20 - 15:30	B R E A K
---------------	------------------

SESSION C

SESSION CHAIR	Jamila KALANTAROVA	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
15:30 - 15:45	Benford's Law and the Digits of Powers of Two in Ternary Numeral System	Yagub N. ALIYEV
15:45 - 16:00	Gegenbauer Wavelet Solutions of Lane-Emden Equations	Demet OZDEK / Sevin GUMGUM
16:00 - 16:15	Global non-existence of the initial value problem for nonlinear thermoelasticity-type equations	Jamila KALANTAROVA

16:15 - 16:20	B R E A K
---------------	------------------

16:20 – 16:40	<p style="text-align: center;">Invited Speech: Professor Maria LONGOBARDI / Università di Napoli Federico II – Italy <i>Cumulative entropies in terms of moments of order statistics</i></p>
---------------	--

16:40 - 16:45	B R E A K
---------------	------------------

SESSION D

SESSION CHAIR	Gulder KEMALBAY	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
16:45 – 17:00	Skewness Correction Ewma for Non-Normal Processes	Aminou MOUSTAPHA / Derya KARAGÖZ
17:00 – 17:15	Bayesian Inversion into Soil Types with Kernel-Likelihood Models	Selamawit SERKA MOJA / Henning MORE
17:15 – 17:30	Computations of Conditional Expectations for General Models of Ordered Random Variables	Ege CAGATAY / Ismihan BAYRAMOGLU
17:30 - 17:45	Distribution of conditional bivariate order statistics via modifications of the bivariate binomial distribution	Gulder KEMALBAY

17:45 – 18:00	B R E A K
---------------	------------------

18:00 – 18:30	<p>Keynote Speech: Professor Narayanaswamy BALAKRISHNAN / McMaster University – Canada <i>Relationships between cumulative entropy/extropy, Gini mean difference and probability weighted moments</i></p>
---------------	---

18:30 – 18:45	B R E A K
---------------	------------------

18:45 – 19:15	<p>Keynote Speech: Professor Barry C. ARNOLD / University of California- USA <i>Half-normal curiosities</i></p>
---------------	---

21 JUNE 2023 WEDNESDAY

11:45 – 12:15	<p>Keynote Speech: Professor Eugenia STOIMENOVA / Bulgarian Academy of Sciences – Bulgaria <i>Multi-Sample Rank Tests for Location against Ordered Alternatives</i></p>
---------------	---

12:15 - 12:30	B R E A K
---------------	------------------

SESSION E

SESSION CHAIR	Fatma NOYAN TEKELI	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
12:30 – 12:45	Drought Area Prediction with Remote Sensing Data Based on Machine Learning Methods	Gokce GOK/ Burcu HUDAVERDI
12:45 – 13:00	Online Statistics Education Using Web-based Software, eStat	Jung Jin LEE
13:00 – 13:15	Implementation of Recommender Systems Using Genetic Algorithm for E-Commerce	Merve POSLU / Guvenc ARSLAN / Sevil BACANLI
13:15 – 13:30	Daily Natural Gas Price Prediction with Long Short-Term Memory	Coskun PARIM / Batuhan OZKAN / Fatma NOYAN TEKELI

13:30- 13:40	B R E A K
--------------	------------------

SESSION F

SESSION CHAIR	Gulhayat GOLBASI SIMSEK	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
13:40 - 13:55	Analysis of Maximum Precipitation in Thailand Using Model Averaging of Non-Stationary Extreme Value Models	Thanawan PRAHADCHAI / Jeong-Soo PARK
13:55 - 14:10	Machine Learning Applications in Detecting Obfuscated Malware	Opetunde IBITOYE / Mary AGOYI / Aysegul EREM

14:10 – 14:25	Analyzing Centrality and Connectivity in Airport Networks: A Focus on Turkey and Europe	Cem ERSOZ / Filiz KARAMAN
14:25 – 14:40	A Hybrid Method Based on Wavelet Transform and Neural Network Approach for Meteorological Data Prediction	Gokce Nur TASAGIL ARSLAN / Serpil KILIC DEPREN
14:40 - 14:55	Stochastic Processes in Insurance Science	Andreas MAKRIDES

14:55 - 15:15	B R E A K	
----------------------	------------------	--

15:15–15:45	Keynote Speech: Professor Haikady NAGARAJA/ The Ohio State University – USA <i>Odds ratios from logistic, geometric, Poisson, and negative binomial regression models</i>	
--------------------	---	--

15:45 - 16:00	B R E A K	
----------------------	------------------	--

SESSION G

SESSION CHAIR	Arzu BAYGUL EDEN	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
16:00 – 16:15	Unleashing Diversity Potential: The What-If Group Voting Recommendation System	Uguray DURDU / Atilla Recep BASARAN / Alperen BALIK / Enes KACAR / Alper ONER
16:15 - 16:30	A New Perspective on the Evaluation of Comorbidity Indices on Survival	Alev BAKIR / Benan MUSELLIM / Mustafa S. SENOCAK
16:30 – 16:45	Turkish Validation of Meta-analysis Observational Studies in Epidemiology (MOOSE) Check List	Arzu BAYGUL EDEN / Neslihan GÖKMEN INAN / Alev BAKIR KAYI

16:45 - 17:00	B R E A K	
----------------------	------------------	--

17:00 - 17:30	Keynote Speech: Professor Omer OZTURK / The Ohio State University – USA <i>Order Restricted Randomized Block Designs</i>	
----------------------	--	--

17:30 - 18:00	Keynote Speech: Professor Ilham AKHUNDOV / University of Waterloo – Canada <i>Characterizations of distributions based on regression and related statistics</i>	
----------------------	---	--

18:00 - 18:15	B R E A K	
----------------------	------------------	--

SESSION H

SESSION CHAIR	Yazgi TUTUNCU	
TIME	PAPER TITLE	PRESENTER / CO-AUTHOR
18:15 – 18:30	Statistical Analysis of Altitude Effect on Daily Selective Attention Task Performance	Yazgi TUTUNCU / Sermin TUKEK / Hasan KAZDAGLI
18:30 – 18:45	A Numerical Approach to Investigate the Spread of COVID-19	Ayse BELER / Sevin GUMGUM

22 JUNE 2023 THURSDAY

09:00 – 09:30	Keynote Speech: Professor Alex KARAGRIGORIOU / University of the Aegean – Greece <i>Power Divergence Statistics and the Modified Family for the Zero Count Case</i>
---------------	---

09:30 - 09:40	B R E A K
---------------	------------------

09:40 - 10:00	Invited Speech: Professor Coskun KUS / Selcuk University – Turkey <i>A Class of Non-parametric Tests for the Two-Sample Problem Based on Order Statistics</i>
---------------	---

10:00 - 10:10	B R E A K
---------------	------------------

10:10 - 10:30	Invited Speech: Professor Debasis KUNDU / Indian Institute of Technology Kanpur – India <i>Bivariate Distributions with Singular Components</i>
---------------	---

10:30 – 10:45 CLOSING CEREMONY

Professor Ismihan BAYRAMOGLU / Izmir University of Economics - Turkey

Discrete q-beta distributions and analogs of order statistics in geometric distributions.....	
<i>Fernando L'opez-Bl'azquez</i>	16
Optimal plan for ordered step-stress stage life testing	
<i>Debashis Samanta, Shuvashree Mondal and Debasis Kundu</i>	17
On the status of component failures in a working coherent system under inspections.....	
<i>Krzysztof Jasiński,</i>	18
LIKELIHOOD INFERENCE FOR GEOMETRIC LIFETIMES OF COMPONENTS OF K-OUT-OF-N SYSTEMS.....	
<i>Anna DEMBIŃSKA, Krzysztof JASIŃSKI</i>	19
Efficient Sampling from the PKBD Distribution.....	
<i>Lukas SABLICA, Kurt HORNIK, Josef LEYDOLD</i>	20
Robust Estimation of Correlation Coefficient via Ranked Set Sampling: Application to Bivariate Normal Distribution	
<i>Yusuf Can SEVİL, Tuğba ÖZKAL YILDIZ,</i>	21
A NEW TWO-PARAMETER BIASED PREDICTION CLASS IN LINEAR MIXED MODELS	
<i>Özge KURAN,</i>	23
Compound Discrete Time Geometric Process	
<i>Leda Minkova</i>	24
BENFORD'S LAW AND THE DIGITS OF POWERS OF TWO IN TERNARY NUMERAL SYSTEM.....	
<i>Yagub N. ALIYEV</i>	25
GEGENBAUER WAVELET SOLUTIONS OF LANE-EMDEN EQUATIONS	
<i>Demet ÖZDEK¹, Sevin GÜMGÜM²,</i>	26
Global non-existence of the initial value problem for nonlinear thermoelasticity type equations	
<i>Jamila Kalantarova</i>	27
Cumulative entropies in terms of moments of order statistics.....	
<i>Maria Longobardi</i>	28
SKEWNESS CORRECTION EWMA CONTROL CHART FOR NON-NORMAL PROCESSES	
<i>Aminou Tukur Moustapha, Derya Karagöz,</i>	29
Bayesian Inversion into Soil Types with Kernel-Likelihood Models.....	
<i>Selamawit SERKA MOJA, Henning OMRE,</i>	30
Half-normal curiosities.....	
<i>Barry C. Arnold</i>	31
Multi-Sample Rank Tests for Location Against Ordered Alternatives.....	
<i>Eugenia Stoimenova</i>	32
DROUGHT AREA PREDICTION WITH REMOTE SENSING DATA BASED OF MACHINE LEARNING METHODS	
<i>Gokce Gok, Burcu Hudaverdi</i>	33

Online Statistics Education Using Web-based Software, eStat	
<i>Jung Jin LEE</i>	35
IMPLEMENTATION OF RECOMMENDER SYSTEMS USING GENETIC ALGORITHM FOR E COMMERCE	
<i>Merve POSLU, Güvenç ARSLAN, Sevil BACANLI</i>	36
Daily Natural Gas Price Prediction with Long Short-Term Memory	
<i>Batuhan ÖZKAN¹, Coşkun PARİM², Fatma NOYAN TEKELİ³</i>	37
Analysis of Maximum Precipitation in Thailand Using Model Averaging of Non-Stationary Extreme Value Models	
<i>Thanawan PRAHADCHAI, Jeong-Soo PARK</i>	38
MACHINE LEARNING APPLICATIONS IN DETECTING OBFUSCATED MALWARE.....	
<i>Opetunde IBITOYE¹, Mary AGOYI, Aysegul EREM,</i>	39
Analyzing Centrality and Connectivity in Airport Networks: A Focus on Turkey and Europe.....	
<i>Cem ERSÖZ¹, Filiz KARAMAN,</i>	40
A Hybrid Method Based on Wavelet Transform and Neural Network Approach for Meteorological Data Prediction.....	
<i>Gökçe Nur TAŞAĞIL ARSLAN, Serpil KILIÇ DEPREN,</i>	41
Stochastic Processes in Insurance Science.....	
<i>Andreas Makrides</i>	43
Odds ratios from logistic, geometric, Poisson, and negative binomial regression models*	
<i>H. N. Nagaraja,</i>	44
A NEW PERSPECTIVE ON THE EVALUATION OF COMORBIDITY INDICES ON SURVIVAL <i>Alev BAKIR¹, Benan MÜSELLİM², Mustafa Ş. ŞENOCAK²</i>	45
Turkish Validation of Meta-analysis Observational Studies in Epidemiology.....	
<i>Arzu Baygül EDEN, Neslihan GÖKMEN İNAN, Alev BAKIR KAYI</i>	47
Order Restricted Randomized Block Designs.....	
<i>Omer Ozturk</i>	49
Statistical Analysis of Altitude Effect on Daily Selective Attention Task Performance	
<i>Şermin TÜKEL, Hasan KAZDAĞLI, G.Yazgı TÛTÛNCÛ</i>	50

Discrete q -beta distributions and analogs of order statistics in geometric distributions

Fernando Lopez-Blazquez
Universidad de Sevilla, Spain

Abstract

We present the discrete q -beta distribution and study some of its properties. We show that the order statistics of non-decreasing sequences of geometric random variables follow a discrete q -beta distribution. We study properties of these analogs of order statistics and show that despite the presence of ties, these order statistics exhibit properties similar to those of continuous distributions, so, for instance the dependence structure is markovian.

Optimal plan for ordered step-stress stage life testing

Debashis Samanta¹, Shuvashree Mondal² and Debasis Kundu³

¹ Department of Mathematics and Statistics, Aliah University, Kolkata, India

² Department of Mathematics and Computing, IIT (ISM) Dhanbad, Dhanbad, India, shuvasri29@iitism.ac.in

³ Department of Mathematics and Statistics, IIT Kanpur, Kanpur, India, kundu@iitk.ac.in

Abstract

In a progressively censored life testing experiment, the censored units can be used for some other life testing experiments. In a step-stress stage life testing experiment, we run parallel experiments while placing the progressively censored units at the higher stress level along with the main experiment. In this work, multiple step-stress stage life testing experiments are performed under the assumption that the experimental units at different stress levels follow Weibull distributions with the same shape parameter, but different scale parameters. The main objectives of this paper are to provide an order-restricted classical inference of the model parameters and also to find out the optimal life testing scheme. We provide the order-restricted maximum likelihood estimators of the unknown parameters and the associated confidence intervals based on the observed Fisher information matrix. An extensive simulation study and a data analysis are provided for illustrative purposes. In search of the optimal life testing plan under this set ups, a meta-heuristic approach based on a variable neighbourhood search algorithm is employed using D-optimality criterion. Finally, we have indicated how the method can be extended when the shape parameter of the Weibull distributions are also different at different stress levels.

Key Words: *Optimal life testing; order-restricted inference; progressive censoring; stage life testing; variable neighbourhood search*

On the status of component failures in a working coherent system under inspections

Krzysztof Jasiński¹,

¹ Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopina 12/18, 87-100 Toruń, Poland, e-mail: krzys@mat.umk.pl

Abstract

We investigate the number of failed components in a coherent system with possibly dependent and not necessarily identically distributed component lifetimes, see [1]. Based on two different partial information about the status of component failures, obtained either with a single inspection or under double monitoring of the system, we compute the probability of the number of failures in a working coherent system. In the computations we use the concept of minimal paths sets of the system. These results enable us to establish explicit expressions for survival functions of two types of residual lifetimes of a used system.

Key Words: *Coherent system; discrete lifetime distribution; reliability theory; residual lifetime; order statistics*

References

[1] Jasiński K (2023) On the status of component failures in a working coherent system when the lifetimes are DNID random variables, *Statistics* 57:175-194.

LIKELIHOOD INFERENCE FOR GEOMETRIC LIFETIMES OF COMPONENTS OF k -OUT-OF- n SYSTEMS

Anna DEMBIŃSKA¹, Krzysztof JASIŃSKI²

¹ Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warsaw, Poland, e-mail: dembinsk@mini.pw.edu.pl

² Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland, e-mail: krzys@mat.umk.pl

Abstract

So called k -out-of- n systems form an important class of systems studied in reliability theory. These systems have n components and functions as long as at least k of the components work. Their examples include series (when $k = n$) and parallel (when $k = 1$) systems.

The talk will be concerned with the likelihood inference based on failure times of components of a k -out-of- n system registered up to and including the moment of the breakdown of the whole system. We will assume that the component lifetimes are independent and have the same geometric distribution with an unknown parameter and our aim will be to estimate this parameter. First, we will find the corresponding maximum likelihood estimator (MLE). Then, restricting the attention to series systems, we will show some exact and asymptotic properties of the MLE of interest. Next, we will generalize the obtained results to the situation when we observe N independently working k -out-of- n systems rather than one such system. In particular, we will describe the asymptotic behavior of the MLE as N grows to infinity.

The talk will be based on the paper [1].

Key Words: *MLE; inference; k -out-of- n system; limit theorem; reliability theory*

References

[1] Dembińska A, Jasiński K. (2022), Likelihood inference for geometric lifetimes of components of k -out-of- n systems, Manuscript.

Efficient Sampling from the PKBD Distribution

Lukas SABLICA¹, Kurt HORNIK², Josef LEYDOLD³

¹E-mail: lukas.sablica@wu.ac.at

²E-mail: kurt.hornik@wu.ac.at

³E-mail: josef.leydold@wu.ac.at

Abstract

We present and analyze random number generators for Poisson Kernel-Based Distribution (PKBD) on the sphere. We show that the currently only available sampling scheme presented in [1], can be improved by a better selection of hyper-parameters and that this sampling method yields an unbounded rejection constant as the concentration parameter approaches 1. Furthermore, we introduce two additional and superior sampling methods for which boundedness in the above mentioned case can be obtained.

The first method proposes initial draws from angular central Gaussian distribution and offers uniformly bounded rejection constant for a significant part of the PKBD parameter space. The second method uses adaptive rejection sampling and the results of [2] to sample from the projected Saw distribution (Saw, 1978). Finally, both new methods are compared in a simulation study.

Key Words: *Sampling; Rejection sampling; Simulation, Poisson kernel; PKBD distribution; Spherical distributions*

References (Times New Roman, 12pt, Bold)

- [1] Golzy, M. and Markatou, M. (2020). Poisson Kernel-Based Clustering on the Sphere: Convergence Properties, Identifiability, and a Method of Sampling. *Journal of Computational and Graphical Statistics* 0 1–13.
- [2] Ulrich, G. (1984). Computer Generation of Distributions on the m-Sphere. *Journal of the Royal Statistical Society Series C (Applied Statistics)* 33 158–163.
- [3] Saw, J. G. (1978). A Family of Distributions on the m-Sphere and Some Hypothesis Tests. *Biometrika* 65 69–73.

Robust Estimation of Correlation Coefficient via Ranked Set Sampling: Application to Bivariate Normal Distribution

Yusuf Can SEVİL¹, Tuğba ÖZKAL YILDIZ²,

¹ Dokuz Eylul University, The Graduate School of Natural and Applied Science, Tinaztepe Campus, 35390, Buca, Izmir, Turkey E-mail: yusufcansevil92@gmail.com

² Dokuz Eylul University, Department of Statistics, Faculty of Science, Tinaztepe Campus, 35390, Buca, Izmir, Turkey E-mail: tugba.ozkal@deu.edu.tr

Abstract

Ranked set sampling (RSS) is a popular data collection scheme that has been used in many areas such as environmental, ecological, agricultural, biological, medical, etc. where exact measurements of the sampling unit are either difficult and/or expensive. First, McIntyre [1] used the RSS for seeking to estimate the yield of pasture. in Australia. Using RSS and its modifications, some authors dealt with the estimation of the dependence parameter, see Stokes [2], Zheng and Modarres [3], Hui et al. [4], and Sevil and Yildiz [5].

To give our motivation, it is assumed that (X, Y) is an absolutely continuous random vector with the joint cumulative distribution function $F(x, y)$ and probability density function $f(x, y)$. Under assumption of perfect ranking, it is supposed that $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{k:k}$ are the order statistics and Y associated with $X_{r:k}$ is called the concomitant of the r th order statistic. Thus, the concomitant variable is denoted by $Y_{[r:k]}$. There may be an error in the ordering of Y due to low correlation coefficient between X and Y . Another potential ranking error is brought on by ordering of X . The present work focus on the second potential ranking error.

In the present work, we examine the maximum pseudo likelihood (MPL) estimate based on RSS for the correlation coefficient of BVN distribution. The new correlation coefficient estimator is compared with its counterparts in SRS and RSS. According to the results, the developed correlation coefficient estimator is biased but the bias values decrease as the sample size increases. In most cases, the developed estimator is more efficient than the other correlation coefficient estimates based on SRS and RSS under the assumptions of perfect and imperfect ranking. Also, it has been observed that the developed correlation coefficient estimate is robust under the assumption of imperfect ranking.

Key Words: *Correlation coefficient; ranked set sampling; pseudo maximum likelihood; maximum pseudo likelihood; bivariate normal distribution*

References

- [1] McIntyre (1952) A method for unbiased selective sampling, using ranked set sampling. Australian Journal of Agricultural Research 3:385-390.
- [2] Stokes SL (1980) Inference on the correlation coefficient in bivariate normal populations from ranked set samples. Journal of the American Statistical Association 75:989-995.

- [3] Zheng G, Modarres R (2006) A robust estimate of the correlation coefficient for bivariate normal distribution using ranked set sampling. *Journal of Statistical Planning and Inference* 136:298-309.
- [4] Hui TP, Modarres R, Zheng G (2009) Pseudo maximum likelihood estimates using ranked set sampling with applications to estimating correlation. *Test* 18:365-380.
- [5] Sevil YC, Yildiz TO (2022) Gumbel's bivariate exponential distribution: estimation of the association parameter using ranked set sampling. *Computational Statistics* 37:1695-1726.
- [6] Genest C, Favre A-C (2007) Everything you always wanted to know about copula modelling but were afraid to ask. *Journal of Hydrological Engineering* 12:347-368.

A NEW TWO-PARAMETER BIASED PREDICTION CLASS IN LINEAR MIXED MODELS

Özge KURAN¹,

¹Dicle University, Faculty of Science, Department of Statistics, Diyarbakır/TURKEY

ozge.kuran@dicle.edu.tr;ozge-kuran@hotmail.com

Abstract

In this paper, we propose a new two-parameter biased prediction class, referred to as the modified Liu prediction, specifically designed for linear mixed models with multicollinearity. By employing the mean square error criterion, we establish the necessary and sufficient conditions for the superior performance of the modified Liu prediction compared to the existing ridge and Liu predictions in the literature of linear mixed models. Additionally, we address the selection of biasing parameters for the modified Liu predictors. To assess the theoretical performance of our newly proposed biased prediction class, we conduct a real data analysis, which concludes the paper.

Key Words: *The best linear unbiased prediction; Ridge prediction; Liu prediction; The modified Liu prediction; Multicollinearity*

Compound Discrete Time Geometric Process

Leda Minkova

Faculty of Mathematics and Informatics,
So_a University "St. Kl.Ohridski",
e-mail: leda@fmi.uni-so_a.bg

Abstract

In this talk we introduce a new discrete time stochastic process with geometric underlying distribution (DTGP). The process is a discrete version of the well-known geometric process (GP) introduced by Lam (2007). In order to analyze the DTGP, we propose the so called hypogeometric distribution, which is a sum of independent geometrically distributed variables with different parameters. Then, as one special case, we obtain the geometric process. In a similar way, in the case of modified geometric underlying distribution we introduce the compound discrete time geometric process. The talk is in collaboration with Stefanka Chukova and Silvana Paralloi.

BENFORD'S LAW AND THE DIGITS OF POWERS OF TWO IN TERNARY NUMERAL SYSTEM

Yagub N. ALIYEV

School of IT and Engineering, ADA University, Azerbaijan, yaliyev@ada.edu.az

Abstract

We study the digits of the powers of 2 in the ternary number system. We propose an easy algorithm for doubling numbers in ternary number system. Using this algorithm, we explain the appearance of “stairs” of 2’s and 0’s when the number $2^{(n+1)}$ is written on top of 2^n ($n=0,1,2,\dots$) in a natural way so that for example the last digits are forming one column, the pre-last digits are forming another column, etc. We also look at the patterns formed by the first digits, the patterns formed by the last digits and use this to prove that the sizes of these “stairs” as blocks of 0’s and 2’s are unbounded. We used Benford’s law [1], well known to Statisticians to prove the results.

Key Words: Benford’s law; ternary; digits; zeros; twos; powers of two.

References

[1] A. Klenke, (2020) Probability Theory, A Comprehensive Course, Universitext, 716 p.
<https://doi.org/10.1007/978-3-030-56402-5>

GEGENBAUER WAVELET SOLUTIONS OF LANE-EMDEN EQUATIONS

Demet ÖZDEK¹, Sevin GÜMGÜM²,

¹ Izmir University of Economics, Faculty of Arts and Science, Department of Mathematics, 35330, Izmir, Turkey

e-mail: demet.ersoy@ieu.edu.tr

² Izmir University of Economics, Faculty of Arts and Science, Department of Mathematics, 35330, Izmir, Turkey

e-mail: sevin.gumgum@ieu.edu.tr

Abstract

The goal of this study is to use the Gegenbauer wavelet method to solve Lane-Emden equations numerically. Gegenbauer wavelets are formed from Gegenbauer polynomials which are orthonormal and this is an advantage of the method as it reduces the computational cost. In literature nonlinearity in these equations are handled by linearizing the nonlinear terms. The application of the current method does not require this. It basically transforms the differential equations to a system of nonlinear algebraic equations. The accuracy and efficiency of the Gegenbauer wavelet method is demonstrated on three problems modelling different physical phenomena. The results show that either analytical solution or accurate numerical solutions can be obtained using this method.

Key Words: *Lane-Emden equations; Gegenbauer wavelets; Initial value problems; Boundary value problems; nonlinear ordinary differential equations*

References

- [1] Wazwaz AM (2001) A new algorithm for solving differential equations of Lane-Emden type. *Applied Mathematics and Computation* 118:287-310. doi:10.1016/S0096-3003(99)00223-4.
- [2] Mohsenyazadeh M, Maleknejad K, Ezzati R (2015) A numerical approach for the solution of a class of singular boundary value problems arising in physiology. *Advances in Difference Equations* 231. doi: 10.1186/s13662-015-0572-x.
- [3] Lima PM, Morgado L (2010) Numerical modeling of oxygen diffusion in cells with Michaelis–Menten uptake kinetics. *Journal of Mathematical Chemistry* 48:145-158. doi: 10.1007/s10910-009-9646-x.
- [4] Kumar S, Pandey P, Das S (2019) Gegenbauer wavelet operational matrix method for solving variable-order non-linear reaction–diffusion and Galilei invariant advection–diffusion equations. *Computational and Applied Mathematics* 162:1-22. doi: 10.1007/s40314-019-0952-z.
- [5] Mall S, Chakraverty S (2015) Numerical solution of nonlinear singular initial value problems of Emden-Fowler type using Chebyshev Neural Network method. *Neurocomputing*, 149: 975-982. doi: 10.1016/j.neucom.2014.07.036.

Global non-existence of the initial value problem for nonlinear thermoelasticity type equations

Jamila Kalantarova¹

¹Izmir University of Economics

jamila.kalantarova@ieu.ed.tr

Abstract

The initial boundary value problems for the following systems of equations of thermoelasticity type were considered.

and

$$u_{tt} - \nabla \cdot (|\nabla u|^{p-2} \nabla u) + \theta = |u|^{r-2} u, \quad x \in \Omega, \quad t > 0, \quad \theta_t - \Delta \theta = u, \quad x \in \Omega, \quad t > 0, \quad (1)$$

$$v_{tt} - \Delta v + b \nabla \theta = f(v), \quad x \in \Omega, \quad t > 0, \quad \theta_t - \Delta \theta + b \nabla \cdot v_t = 0, \quad x \in \Omega, \quad t > 0, \quad (2)$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain with sufficiently smooth boundary $\partial\Omega$, $r \geq p > 2$, $b > 0$, $v > 0$, $\beta > 0$, $\gamma > 0$, $\delta > 0$ are given parameters, $u, v = (v_1, v_2, v_3)$, q are unknown functions, $f(\cdot) \in C^1(\mathbb{R}^3 \rightarrow \mathbb{R}^3)$ is a given gradient vector with the potential $F(\cdot)$ that satisfies the condition

$$f(z) \cdot z - 2(2\alpha + 1)F(z) \geq -D_0, \quad D_0 \geq 0, \quad \alpha > 0, \quad \forall z \in \mathbb{R}^3. \quad (3)$$

We found sufficient conditions of blow up in finite time of solutions to initial boundary value problems (1) and (2), by using the concavity method and some of its modifications. We showed that solutions even with large enough initial energies of the considered problems may blow up in a finite time.

Key Words: *blow up of solutions, concavity method, coupled system*

References

- [1] H. A. Levine (1974) "Some nonexistence and instability theorems for solutions of formally parabolic equations of the form $Put = -Au + F(u)$," Arch. Rational Mech. Anal. 51, 371–386.
- [2] V. K. Kalantarov and O. A. Ladyzhenskaya (1978) "The occurrence of collapse for quasilinear equations of parabolic and hyperbolic types," J. Math. Sci. 10 (1), 53–70.
- [3] M. O. Korpusov (2012) "Blow-up of the solution of a nonlinear system of equations with positive energy," Theoret. and Math. Phys. 171 (3), 725–738.

Cumulative entropies in terms of moments of order statistics

Maria Longobardi

Abstract

Relations between some kinds of cumulative entropies and moments of order statistics are established. By using some characterizations and the symmetry of a non-negative and absolutely continuous random variable X , lower and upper bounds for entropies are obtained and illustrative examples are given. By the relations with the moments of order statistics, a method is shown to compute an estimate of cumulative entropies and an application to testing whether data are exponentially distributed is outlined.

Reference

Balakrishnan, N., Buono, F., Longobardi, M. (2022). On Cumulative Entropies in Terms of Moments of Order Statistics. *Methodology and Computing in Applied Probability*, 24, 345--359.

SKEWNESS CORRECTION EWMA CONTROL CHART FOR NON-NORMAL PROCESSES

Aminou Tukur Moustapha¹, Derya Karagöz²,

¹Yıldız Technical university; e-mail: moustaphaaminoutukur@gmail.com

²Hacettepe University; e-mail: deryacaliskankaragoz@gmail.com

Abstract

The classic Shewhart control charts are generally used for monitoring the process mean and variability in the characteristics of a random quality variable of interest and are based on the normality assumptions. For skewed distributions, to demonstrate the changes in the population, non-symmetric control limits need to be used. Methods such as the Weighted Variance (WV), Weighted Standard Deviation (WSD) and Skewness Correction (SC) are generally employed. In this paper, the Skewness Correction method is applied to the Exponentially Weighted Moving Average (EWMA) scheme and a control chart called Skewness Correction EWMA (SC_{EWMA}) which is going to be used to monitor non-normal processes is proposed. The performance of the newly proposed SC_{EWMA} is compared with those of the Weighted Variance EWMA (WV_{EWMA}), Weighted Standard Deviation EWMA (WSD_{EWMA}) and the classic EWMA control chart limits based on the degree of skewness and varying smoothing parameters. The comparison is made with respect to their Type-I errors by applying the Monte Carlo simulation technique with data generated from the lognormal, Gamma and Weibull distributions. The newly proposed SC_{EWMA} control chart is expected to produce good results and hence proposed as an alternative to the widely used WV_{EWMA} .

Key Words: *Skewed Distributions, EWMA Control Charts, WV method, WSD method, SC Method.*

Bayesian Inversion into Soil Types with Kernel-Likelihood Models

Selamawit SERKA MOJA¹, Henning OMRE²,

¹Department of Statistics, University of Hawassa, Hawassa, Ethiopia

e-mail: selamawitserka.get@gmail.com

² Department of Mathematical Sciences, Norwegian University of Science and Technology, Norway

e-mail: henning.omre@ntnu.no

Abstract

Knowledge of the sub-surface characteristics is crucial in many engineering activities. Subsurface soil classes must for example be predicted from indirect measurements in narrow drill holes and geological experience. In this study the inversion is made in a Bayesian framework by defining a hidden Markov chain. The likelihood model for the observations is assumed to be in factorial form and they are assessed from a calibration well by kernel estimators. The prior Markov model are defined to be either a traditional stationary Markov chain or a trend Markov chain. The methodology is demonstrated on one case study for offshore fundamentation of wind-mills. We conclude that a suitable choice of kernel likelihood model is of at most importance, and that using a trend Markov prior model improve the predictions even more.

Key Words: *Sub-surface layer prediction, circular uniform kernel, Gaussian kernel.*

References

- [1] Baum LE, Petrie T, Soules G, Weiss N(1970) A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. The Annals of Mathematical Statistics 41(1):164-171.
- [2] Besag J(1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical Society, Series B 36:192-236.

Half-normal curiosities.

Barry C. Arnold

University of California, Riverside, USA

Abstract

If we consider two independent identically distributed half-normal random variables, it may be verified that the sum of the two variables and the maximum of the two variables have distributions of the same type. This surprising occurrence naturally gives rise to a conjectured characterization of the half-normal distribution. This conjecture will be investigated in the current presentation. Attention is also directed to the possibility of similar results for samples of larger sizes.

(based on joint work with Jose Villasenor)

Multi-Sample Rank Tests for Location Against Ordered Alternatives

Eugenia Stoimenova

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences

Abstract

The talk will be concerned with a family of rank tests for the multi-sample hypothesis testing problem. The proposed test statistics are based on the number of observations from one sample that precede or exceed a threshold specified by the other samples. We discuss here distribution properties of such type tests statistics and consistency of the tests against specific ordered alternatives. We will consider an extension of a Lehmann alternative for the two-sample location problem appropriate to the multi-sample problem which leads to a distribution free test under the alternative.

DROUGHT AREA PREDICTION WITH REMOTE SENSING DATA BASED OF MACHINE LEARNING METHODS

Gokce Gok¹, Burcu Hudaverdi²

¹ Dokuz Eylul University, gok.gokce@ogr.deu.edu.tr

² Dokuz Eylul University, burcu.hudaverdi@deu.edu.tr

Abstract

The increasing uncertainty of drought forecasting based on historical meteorological data can be attributed to the influence of climate change. As climate patterns shift and become more erratic, the relationships and trends observed in past data may no longer accurately represent future conditions. The complex and interconnected nature of climate systems further amplifies the difficulty in accurately forecasting droughts. Remote sensing data and machine learning methods have frequently been used in the prediction and forecasting of drought areas, due to the high accuracy and fast results they provide. Detection of drought areas plays an important role in areas such as vegetation health, sustainable management of water resources, forest fire risk detection, long-term and short-term environmental planning.

In this paper, we employed some machine learning methods such as Random Forests, Gradient Boosting Machines, XGBoost and LightGBM to predict drought areas for short-term drought without meteorological data. Landsat 8 satellite images and SRTM 30 m topographic data were used to predict the soil moisture index (SMI) values in the study area covering southwest of Turkey. The land surface factors which are vegetation, topographic, thermal, and water factors were used in drought area prediction. In the learning process, with the help of these methods, the drought function from the training area learned the changes in the soil moisture index (SMI) values. After training and testing with randomly selected 60641 samples, we compared their performances. Moreover we determined the importance of the features in predicting the drought events, so we gained insights into which factors have the greatest influence on drought occurrence.

Key Words: *drought prediction; remote sensing; climate change; random forest; gradient boosting*

References

[1] Haekyung Park, Kyungmin Kim, Dong kun Lee, Prediction of Severe Drought Area Based on Random Forest: Using Satellite Image and Topography Data, *Water* 2019, 11, 705; doi:10.3390/w11040705

[2] Yongnian Zeng, Zhaodong Feng, Nanping Xiang, Assessment of soil moisture using Landsat ETM+ temperature/vegetation index in semiarid environment, 2004, IEEE

[3] Inge Sandholt, Kjeld Rasmussen, Jens Andersen, A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status, *Remote Sensing of Environment* 79, 2002, 213–224

[4] Fei Wang, Zhihao Qin, Caiying Song et al., An Improved Mono-Window Algorithm for Land Surface Temperature Retrieval from Landsat 8 Thermal Infrared Sensor Data, *Remote Sens.*, 2015, 7, 4268-4289; doi:10.3390/rs70404268

[5] Earth Explorer. Available online: <https://earthexplorer.usgs.gov>

Online Statistics Education Using Web-based Software, eStat

Jung Jin LEE¹

¹Professor, ADA University, Baku, Azerbaijan, jlee@ada.edu.az, jjlee@ssu.ac.kr

Abstract

Well known statistical packages such as SAS, SPSS and R are widely used for data processing, but they have not paid attention much to develop modules specialized in Statistics Education. eStat is a free web-based graphical software at www.estat.me which has been developed especially for online Statistics Education. eStat is designed to practice all topics of Introductory Statistics at undergraduate level with many simulation modules, but also to practice data processing online. More specifically, this software includes the modules of all statistical distributions, Law of Large Number, Central Limit Theorem, Confidence Interval, Testing Hypothesis, Analysis of Variance, Nonparametric Tests, Regression Analysis and Forecasting. Moreover, eStat is linked with an eBook which includes many examples, presentation files and lecture movies for students to comprehend the subject of Statistics. This integrated system of eStat was adopted for online teaching by ADA University during pandemic period and it has been very useful for both students and lecturers. It is now used for offline teaching. This paper introduces useful experience of eStat for Statistics Education.

Key Words: Statistics education; statistical package; web-based software; eStat

IMPLEMENTATION OF RECOMMENDER SYSTEMS USING GENETIC ALGORITHM FOR E COMMERCE

Merve POSLU¹, Güvenç ARSLAN², Sevil BACANLI³

¹ Hacettepe University, Department of Statistics, merve.poslu1606@gmail.com

² Kırıkkale University, Department of Statistics, guvenc.arslan@gmail.com

³ Hacettepe University, Department of Statistics, sevil@hacettepe.edu.tr

Abstract

Nowadays, the existence of digitalized services may be very beneficial. For example, the field of e-commerce has created an area of important developments by turning the COVID-19 pandemic into an advantage [1]. One of these developments in e-commerce is to make the best use of recommendation systems. Recommender systems basically work by monitoring the interaction between the users and items. Data gathered during this monitoring process can be used to obtain recommender decisions by applying various data mining techniques [2]. This study aims to investigate the use of genetic algorithms and clustering in recommender systems. Genetic algorithms, which are basically effective optimization algorithms based on natural selection and genetic principles, are one of the heuristic methods used successfully in various fields and applications [3]. Although there are many ways to implement genetic algorithms in recommender systems, we used them to improve the clustering profiles of the users. These clustering profiles are then used to make decisions for recommendation outputs (see, for example [4]). It is expected that using genetic algorithms will provide better profiling of users in the clustering method and improve recommendations.

Key Words: Recommendation systems, Genetic Algorithm, Clustering, E-commerce.

References

- [1] Çiçek, Y., & Muzaffaer, H. (2021). The Impact of Covid-19 Pandemic Crisis on Online Shopping. *AYBU Business Journal*, 1(1), 16-25.
- [2] Gündoğdu, Ö., *Veri Madenciliğinde Genetik Algoritmalar, Yüksek Lisans Tezi, Kocaeli Üniversitesi, Kocaeli, 2007.*
- [3] Reeves, C.R., Rowe, J.E., *Genetic Algorithms: Principles and Perspectives: A Guide to GA Theory*, Kluwer Academic Publishers, Norwell, MA, 1-63, 2002.
- [4] Kim, K. J., & Ahn, H. (2008). A recommender system using GA K-means clustering in an online shopping market. *Expert systems with applications*, 34(2), 1200-1209.

Daily Natural Gas Price Prediction with Long Short-Term Memory

Batuhan ÖZKAN¹, Coşkun PARİM², Fatma NOYAN TEKELİ³

¹ Bitlis Eren University, Faculty of Art & Science, Department of Statistics: bozkan@beu.edu.tr

² Yildiz Technical University, Faculty of Art & Science, Department of Statistics: cparim@yildiz.edu.tr

³ Yildiz Technical University, Faculty of Art & Science, Department of Statistics: fnoyan@yildiz.edu.tr

Abstract

In recent years, natural gas has emerged as a prominent and environmentally friendly energy source. The determination of natural gas spot prices involves the consideration of numerous influential indicators, prompting diverse methodologies in the literature for price prediction.

This study aims to employ the Long Short-Term Memory (LSTM) model in conjunction with twenty-one economic indicators to predict the natural gas price. The dataset utilized comprises daily natural gas spot prices and twenty-one related economic variables. Specifically, historical data spanning 2942 days, encompassing the period from February 27, 2012, to June 6, 2023, is incorporated.

The study proceeds as follows: Firstly, the economic indicators influencing the natural gas spot price for the 7th and 14th days will be determined with the assistance of the genetic algorithm (GA). Subsequently, the gas price will be predicted using the statistically significant variables identified by GA, both 7 and 14 days into the future, employing LSTM. Finally, the prediction performance of the natural gas price will be evaluated by employing metrics such as R-squared, MAE, and RMSE.

Ultimately, the study findings demonstrate that the combined approach of GA and LSTM yields estimations of the 7th and 14th day gas prices with a commendable level of accuracy, thereby underscoring the efficacy of this methodology in natural gas price prediction.

Key Words: *natural gas price, lstm, deep learning, genetic algorithm, prediction*

Analysis of Maximum Precipitation in Thailand Using Model Averaging of Non-Stationary Extreme Value Models

Thanawan PRAHADCHAI¹, Jeong-Soo PARK¹

¹Department of Mathematics and Statistics, Chonnam National University, Gwangju 61186, Korea
tanawanp.st@gmail.com (T.P.); jspark@jnu.ac.kr (J.P.)

Abstracts

Non-stationarity in heavy rainfall time series is often apparent in the form of trends because of long-term climate changes. We have built non-stationary (NS) models for annual maximum daily (AMP) data observed by 79 stations over Thailand. Totally 16 time-dependent functions of the location and scale parameters of the generalized extreme value (GEV) model are considered. At each station, a model is selected by using Akaike information criteria (AIC) among these candidates. The return levels corresponding to some years are calculated and predicted for the future. The model averaging model shows less variance in return level estimation than the best single model.

Key Words: *Generalized extreme value distribution; Model averaging; Non-stationary; Akaike weight; Delta method; Heavy rainfall*

MACHINE LEARNING APPLICATIONS IN DETECTING OBFUSCATED MALWARE

Opetunde IBITOYE¹, Mary AGOYI², Aysegul EREM³,

¹oibitoye@ciu.edu.tr

²magoyi@ciu.edu.tr

³aerem@ciu.edu.tr

Abstract

Malware presents a growing threat to computer systems and networks, with obfuscated malware being particularly insidious due to its ability to mimic the host environment and evade detection [1]. In this study, several machine-learning algorithms for detecting obfuscated malware were examined. The CIC-mem malware'22 dataset [2] was used, and essential data preparation steps were performed, such as handling missing values, data transformation, label encoding, feature selection, and data splitting. Machine learning algorithms, including Decision tree (DT), Random forest (RF), Naïve Bayes (NB), Support vector machine (SVM), and artificial neural network (ANN) algorithms, were used during the training stage. The models' parameters were carefully chosen using grid 10-fold cross-validation. The models were evaluated using accuracy, precision, and kappa-score metrics. The results show that the models produced reliable results using all three metrics. These findings have important implications for detecting obfuscated malware in computer systems.

Key Words: *Machine learning; Malware detection; Obfuscated malware; Naïve Bayes; Support vector machine.*

References

- [1] Q. V. Dang, "Enhancing Obfuscated Malware Detection with Machine Learning Techniques," *Commun. Comput. Inf. Sci.*, vol. 1688 CCIS, pp. 731–738, 2022, doi: 10.1007/978-981-19-8069-5_54.
- [2] T. Carrier, P. Victor, A. Tekeoglu, and A. Lashkari, "Detecting Obfuscated Malware using Memory Feature Engineering," no. Icissp, pp. 177–188, 2022, doi: 10.5220/0010908200003120.

Analyzing Centrality and Connectivity in Airport Networks: A Focus on Turkey and Europe

Cem ERSÖZ¹, Filiz KARAMAN²,

¹ Faculty of Aviation and Aeronautical Science, Ozyegin University, Istanbul, Turkey,
cem.ersoz@ozyegin.edu.tr

² Faculty of Arts and Science, Yildiz Technical University, Istanbul, Turkey, fkaraman@yildiz.edu.tr

Abstract

The main objective of this research is to employ complex network theory to analyze the airport networks of Turkey and Europe. The research question at hand revolves around identifying the topological characteristics of air transport networks and studying their evolution over time. Additionally, the study emphasizes the importance of air transport in the economic and social progress of countries. To examine the structure of Turkey's Airport Network (TAN) and the European Airport Network (EAN), various data sources are utilized. For the analysis of TAN, FlightRadar data is collected during the initial phase following the first Covid-19 outbreak in Turkey. The TAN is evaluated based on its centrality and connectivity to comprehend its structure. Furthermore, the TAN is compared to simulated network models for classification purposes, and it is also compared with other regional and national networks. In the case of Europe, the study relies on 2019 passenger data from Eurostat to explore the structure of the EAN. Network measures are computed to gain insights into the characteristics of the network, and the communities within the EAN are identified using modularity maximization techniques. Additionally, the study investigates binary centrality measures of the network and compares them with their weighted counterparts. The weighted measures incorporate passenger demand and route distances as edge weights in the analysis. This comparison reveals notable distinctions between the two approaches. Moreover, the research considers the findings of previous research papers to analyze the evolution of the air transport networks in Turkey and Europe. By considering the results obtained from these analyses, this study contributes to a better understanding of the air transport networks in Turkey and Europe. It also highlights the similarities and differences between these networks and other air networks worldwide. Ultimately, the insights gained from this research can be valuable in informing policymakers and stakeholders in the aviation industry.

Key Words: *Complex network analysis; Air transport network; weighted betweenness; community detection; connectivity*

A Hybrid Method Based on Wavelet Transform and Neural Network Approach for Meteorological Data Prediction

Gökçe Nur TAŞAĞIL ARSLAN¹, Serpil KILIÇ DEPREN²,

¹Department of Statistics, Faculty of Arts and Sciences, Yildiz Technical University, Turkey,
email:gokcenurtasagil@gmail.com

²Department of Statistics, Faculty of Arts and Sciences, Yildiz Technical University, Turkey,
email:serkilic@yildiz.edu.tr

Abstract

Along with the attention to climate change, the importance of weather forecasting studies has increased. Wavelet transforms (WT) and artificial neural networks (ANN) are among the various methods used to conduct research in this direction. Wavelets are used to denoise the signal. The denoising aims to reduce the noise, flatten, and smooth the original data while maintaining the characteristic structure of the signal. This study examines the use of the ANN and wavelet-ANN hybrid models to predict meteorological data based on the daily averages of four weather factors: temperature, wind speed, humidity, and pressure are gathered by the German meteorological service for the period 2018-2019. It can be seen that the distribution of each variable is composed of noisy data. For this reason, the discrete wavelet transform (DWT) is applied to each variable, and the wavelet coefficients obtained from this transformation using different wavelet types are then utilized for soft universal thresholding. In this way, the ANN and regression analysis are handled over a data set containing critical information instead of the raw data. Thus, the reconstructed noise-free data acquired from the DWT process is used in the ANN in different layers and neurons and linear regression methods. In the comparison of model performance, the coefficient of determination (R^2), mean square error (MSE), and mean absolute percentage error (MAPE) values are evaluated. The result shows that the proposed model has better performance in terms of criteria than the original ANN models, indicating that the WT significantly contributes to modeling based on the data and approaches used. Furthermore, the wavelet Daubechies family outperforms other wavelet families in the decomposition of the meteorological data.

Key Words: *Wavelet Transform; Thresholding; Deep Learning; Artificial Neural Network; Multilayer Perceptron*

References

- [1] Nason, G.P. (2008). *Wavelet Methods in Statistics*, Springer.
- [2] Vidakovic, B. (1999). *Statistical Modeling by Wavelet*, John Wiley & Sons Inc.
- [3] Ciaburro, G., Venkateswaren, B. (2017). *Neural Networks with R: Smart models using CNN, RNN, deep learning, and artificial intelligence principles*. Birmingham: Packt Publishing Ltd.
- [4] Zucatelli, P.J., Nascimento, E.G.S., Santos, A.A.B., Arce, A.M.G., and Moreira, D.M. (2021). An investigation on deep learning and wavelet transform to nowcast wind power and wind power ramp: A case study in Brazil and Uruguay. *Energy*, 230, 120842.

[5] Wang, F., Cen, J., Yu, Z., Deng, S., and Zhang, G. (2022). Research on a hybrid model for cooling load prediction based on wavelet threshold denoising and deep learning. *Energy Reports*, 8, 10950-10962.

[6] DWD, (2019). Deutscher Wetterdienst, Available at: 08.09.2019, https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/.

Stochastic Processes in Insurance Science

Andreas Makrides

Abstract

An innovative approach of loss ratio forecasting is developed using a special type of semi-Markov processes. Three levels of loss ratio are considered as the states of a semi-Markov process, and semi-Markov process methodology is employed for estimating transition probabilities of loss ratio levels transit from a predefined level to another one.

Key Words

Multi-state system, Semi-Markov processes, Reliability analysis, Loss ratio

Odds ratios from logistic, geometric, Poisson, and negative binomial regression models*

H. N. Nagaraja,

The Ohio State University (nagaraja.1@osu.edu)

*Joint work with C. J. Sroka, New Mexico State University (csroka@nmsu.edu).

Abstract

The odds ratio (OR) is used as an important metric of comparison of two or more groups in many applications when the data measure the presence or absence of an event or represent the frequency of its occurrence. In the latter case, researchers often dichotomize the count data into binary form and apply the well-known logistic regression technique to estimate the OR. In the process of dichotomizing the data, however, information is lost. We propose analyzing the count data directly using regression models with the log odds link function. We prove analytically that our approach produces more precise estimates of the OR than logistic regression of the dichotomized data. We also show the gains in precision using simulation studies and real-world datasets, focusing on three related distributions for count data: geometric, Poisson, and negative binomial.

Keywords: Binary data, Confidence intervals, Count data, Fisher information, Maximum likelihood

References

[1] Sroka CJ and Nagaraja HN (2018). Odds ratios from logistic, geometric, Poisson, and negative binomial regression models. *BMC Medical Research Methodology*. 2018 Oct 20;18(1):112. doi: 10.1186/s12874-018-0568-9

A NEW PERSPECTIVE ON THE EVALUATION OF COMORBIDITY INDICES ON SURVIVAL

Alev BAKIR¹, Benan MÜSELLİM², Mustafa Ş. ŞENOCAK²

¹ Istanbul University, Institute of Child Health, Department of Social Pediatrics, Turkey.

alevbakir@yahoo.com

² Istanbul University, Faculty of Cerrahpasa Medicine, Department of Chest Diseases, Turkey (former lecturer).

benanmusellim@gmail.com

² Istanbul University, Faculty of Cerrahpasa Medicine, Department of Biostatistics, Turkey (former lecturer).

mssenocak@gmail.com

Abstract

Mortality and morbidity studies must be evaluated by considering comorbid diseases [1, 2]. Existence, number, and type of comorbid diseases can have an important effect on prognosis, diagnosis, treatment, and outcome [3]. There are some approaches to estimate this affect with realistic and correct predictive power [4]. Comorbid diseases' effect may differ according to comorbidity type and the main disease creating an important issue in survival analysis of patients. Comorbidity indexes are used in the studies or is tried to select the most suitable index for the study by comparing them, but the interaction of comorbid diseases was unobserved. We aim to emphasize evaluation approaches giving different results in judgment results and prediction of prognosis. Retrospective cohort, data were collected from cases with Non-Small Cell Lung Cancer treated in Department of Chest Diseases 1998 to 2012. Comorbid diseases that were projected and selected by senior consultant, are diabetes, chronic obstructive pulmonary disease (COPD), coronary heart disease, renal failure, asthma, interstitial lung disease, previous cancer. Comorbid diseases' effects on the survival time were computed firstly univariate then according to number of comorbidities and finally all together by multivariate analysis and their specific combinations by considering some comorbid indices in the literature. The most valid and reliable and commonly applied indexes (such as Cumulative Illness Rating Scale (CIRS-1968), the Kaplan-Feinstein Classification (KFC-1974), the Charlson Comorbidity Index (CCI-1987) and the Index of Co-Existent Disease (ICED-1987)) were also calculated [3, 4]. Of 247 non-small cell lung cancer (NSCLC) cases analysis, 220 (89%) were men and 27 (11%) were women. Median duration of follow-up was 277 days, at the end of the follow-up 197 cases had died, 50 cases have still lived. The highest rate of these comorbid diseases was coronary heart disease with 13% (n: 32) and the least rate was renal failure with 1% (n: 2). Specific combinations of comorbid diseases' effects on the survival time were found higher than their univariate and number of comorbidities. It is crucial to take into comorbid diseases related to main disease and specially their combination when the risk is estimated in survival research. Reliable judgments of comorbid diseases list related to main disease have a great importance on development of this area.

Key Words: *Survival Analysis; Comorbidity Index; Comorbid Diseases; Interaction Effect; Non-Small Cell Lung Cancer*

References

[1] Birim, Özcan, A. Pieter Kappetein, and Ad JJC Bogers. "Charlson comorbidity index as a predictor of long-term outcome after surgery for nonsmall cell lung cancer." *European journal of cardio-thoracic surgery* 28.5 (2005): 759-762.

- [2] Groll, Dianne L., et al. "The development of a comorbidity index with physical function as the outcome." *Journal of clinical epidemiology* 58.6 (2005): 595-602.
- [3] Hall, Stephen F. "A user's guide to selecting a comorbidity index for clinical research." *Journal of clinical epidemiology* 59.8 (2006): 849-855.
- [4] Extermann, M. "Measuring comorbidity in older cancer patients." *European journal of cancer* 36.4 (2000): 453-471.

Turkish Validation of Meta-analysis Observational Studies in Epidemiology

Arzu Baygöl EDEN¹, Neslihan GÖKMEN İNAN², Alev BAKIR KAYI³

¹Koç University, School of Medicine, Department of Biostatistics, Istanbul, Turkey, abaygul@ku.edu.tr

²Koç University, School Engineering, Istanbul, Turkey, ninan@ku.edu.tr

³Istanbul University, Institute of Child Health, Department of Social Pediatrics, Istanbul, Turkey, alevbakir@yahoo.com

Abstract

Background/aim: Meta-analyses are determined as the most powerful studies in evidence-based medicine. Observational studies give us the opportunity to inspect the patients, the diseases and the treatments in their nature without any intervention. Taking the importance of observational studies into consideration, the quality of meta-analysis of observational studies seems to be very important. A way to raise the quality of a study, is using a validated checklist, which is structured for the specific study type. For this reason, a workshop was held in Atlanta, and as a product of this workshop MOOSE checklist was structured for observational meta-analyses. In this study, we aim to approve the validity of Turkish version of MOOSE checklist.

Materials and methods: MOOSE checklist was translated into Turkish by two authors, after their consensus the translated version was back translated. Then the translated version and the original version of MOOSE were compared. Final Turkish version was applied to 10 observational meta-analyses which are available on databases in Turkish. Two authors scored items independently. The accuracy for the scores of each item, and the agreement for the total scores of two raters were calculated. The agreement was evaluated using the Bland-Altman method.

Results: The accuracy at each item between two authors were excellent (accuracy 0.70-1.00: only one item has 0.70, the others get high values between 0.80 and 1.00). The Cronbach Alpha Coefficient for total scores, which are blindly calculated, was extremely high (0.939).

Conclusion: Turkish version of MOOSE checklist is applicable for Turkish researchers by evaluating observational meta-analyses.

Key Words: *Observational study, Meta-analysis, Check list, MOOSE*

References

- [1] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; Apr 19;283(15):2008-12. doi: 10.1001/jama.283.15.2008. PMID: 10789670
- [2] Simon, A. K., Rao, A., Rajesh, G., Shenoy, R., & Pai, M. B. Reporting guidelines in health research: A review. *Indian journal of public health* 2015, 59(2): 122.
- [3] Moher D, Pham B, Jones A, et al. Does the quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998; 352:609-613.
- [4] Berlin JA, Colditz GA. The role of meta-analysis in the regulatory process for foods, drugs, and devices. *JAMA*. 1999; 281:830-834.
- [5] Yılmaz BM, Kılıçkap M, Abacı A, Barçın C, Bayram F, Karaaslan D, Göksülük H, Kayıkçıoğlu M, Özer N, Süleymanlar G, Şahin M, Tokgözoğlu L, Satman I. Temporal changes in the epidemiology of diabetes mellitus in Turkey: A systematic review and meta-analysis. *Turk Kardiyol Dern Ars*. 2018; 46(7): 546-555
- [6] Ural D, Kılıçkap M, Göksülük H, Karaaslan D, Kayıkçıoğlu M, Özer N, Barçın C, Yılmaz MB, Abacı A, Şengül Ş, Arınsoy T, Erdem Y, Sanisoğlu Y, Şahin M, Tokgözoğlu L. Data on prevalence of obesity and waist circumference in Turkey: Systematic review, meta-analysis and meta-regression of epidemiological studies on cardiovascular risk factors. *Turk Kardiyol Dern Ars*. 2018; 46(7): 577-590

Order Restricted Randomized Block Designs

Omer Ozturk

Department of Statistics

The Ohio State University, Columbus, OH, USA

Field experiments are run under two competing objectives, high precision and minimal cost. The precision can be increased either by using sound experimentation techniques that account for the sources of variation with reasonable statistical models or by increasing the sample size. Large sample sizes usually increase the cost of the experiment and may not be feasible. This paper uses order restricted randomized designs (ORRD) to increase the precision while keeping the sample size and cost of the experiment minimal. The ORRD described here starts with a randomised block design but adds a second layer of blocking by ranking plots within each block. This creates a two-way lay-out, blocks and ranking groups, and uses a restricted randomization to improve the precision of estimation of the treatment parameters. Ranking groups create a correlation structure for within-block units. The restricted randomization uses this correlation structure to reduce the error variance of the experiment. The paper computes the expected mean square for each source of variation in the ORRD design under a suitable linear model. It also provides approximate F-tests for treatment and ranking group effects. The efficiency of the ORRD is investigated through empirical power studies. Finally, an example based on a uniformity field trial illustrates the use of the method in a split-plot experiment.

Key words: Randomized block design, replicated Latin square, row-column design, order restricted randomization

Statistical Analysis of Altitude Effect on Daily Selective Attention Task Performance

Şermin TÜKEL¹, Hasan KAZDAĞLI², G.Yazgi TÛTÛNCÛ³

¹İzmir University of Economics Department of Physiotherapy and Rehabilitation e-mail:sermin.tukel@ieu.edu.tr

²İzmir University of Economics Vocational School of Health Services e-mail: hasan.kazdagli@ieu.edu.tr

³İzmir University of Economics Department of Mathematics e-mail: yazgi.tutuncu@ieu.edu.tr

Abstract

Recreational trekkers are a group of people who might experience the effects of high altitude. The number of people who climb Mount Everest for recreational purposes has been increasing each year [1]. Professional alpinists seem to be a small amount of people who are experienced about the effects of high altitude. However, recreational trekkers are not experienced about the effects of high altitude. Ascend to high altitudes result in exponential decrease in barometric pressure. Therefore, people experience hypobaric hypoxia, e.g. reduced partial pressure of oxygen (PO₂) which lead to reduced oxygen transport cascade from atmosphere to the mitochondria. Hypobaric hypoxia elicits manifold physiological effects in human body resulting primarily from a lower arterial PO₂ (PaO₂) and reduced oxyhemoglobin saturation [2]. Severe and life-threatening conditions might arise as a result of high-altitude exposure such as acute mountain sickness, high-altitude pulmonary, or cerebral oedema especially in those climbers who are altitude naive.

The effects of hypobaric hypoxia on brain functions have been investigated and studies reported a variety of transient neurological consequences, including high altitude headache, sleep disturbances, mental issues and cognitive impairment [3]. Selective attention is one of the core cognitive domains which is a basis for other more complex cognitive functions. Selective attention is tested and found to be limited at different altitudes starting at 3269 m. In contrast, improvements or no changes were found by other studies [3]. Even using the same test, the results were found to be heterogenous, reaching from deterioration, improvement, to no change. Unlike earlier studies which compares two or three measurements at different altitudes, we studied follow-up of daily selective attention task during ascend to Everest Base Camp (EBC) for ten days and compare the results with a sea level task performance in recreational climbers. In addition, we wanted to determine the physiological precursors of high-altitude selective attention task performance. Results revealed that selective attention task performance was limited during EBC trek compared to performance at sea level. Performance during EBC Trek was best explained by systolic and diastolic blood pressure.

Key Words: *high-altitude, selective attention, recreational climbers, Everest Base Camp*

References

- [1] Nyaupane, G., Musa, G., Higham, J., & Thompson-Carr, A. (2015). Mountaineering on Mt Everest: evolution, economy, ecology and ethics. *Mountaineering tourism*, 265-271.
- [2] Cerretelli, P., Marzorati, M., & Marconi, C. (2009). Muscle bioenergetics and metabolic control at altitude. *High Altitude Medicine & Biology*, 10(2), 165-174.
- [3] Bliemsrieder, K., Weiss, E. M., Fischer, R., Brugger, H., Sperner-Unterweger, B., & Hűfner, K. (2022). Cognition and Neuropsychological Changes at Altitude—A Systematic Review of Literature. *Brain Sciences*, 12(12), 1736.