

**BOOK OF ABSTRACTS**



**8th INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS**

**16-18 MAY 2022**

**ICAS CONFERENCE**  
INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

May 16 -18, 2022

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

**ICAS'2022**

8<sup>th</sup> 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](mailto:icasconference.academic@gmail.com)

<http://www.icasconference.com>

ISBN:

**Copyright @ 2022 ICAS 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. 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**

Dr. Ilham AKHUNDOV  
**University of Waterloo – Canada**

## LOCAL SCIENTIFIC COMMITTEE

**Prof. Dr. Aydin ERAR**

Mimar Sinan Fine Arts University – Turkey

**Prof. Dr. Coskun KUS**

Selcuk University – Turkey

**Prof. Dr. I. Esen YILDIRIM**

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**

Istanbul Ayyansaray University – Turkey

**Assoc. Prof. Dr. Baris ASIKGIL**

Mimar Sinan Fine Arts University – Turkey

**Assoc. Prof. Dr. Esra AKDENIZ**

Marmara University – Turkey

## **ORGANIZATION COMMITTEE**

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

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

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

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

**Asst. Prof. Dr. Aysegul EREM**  
Cyprus International University – TRNC

**Instructor PhD Ozlem BERAK KORKMAZOGLU**  
Yıldız Technical University – 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 16-18 May 2022.

All informations are available in conference web site. For more information please do not hesitate to contact us. [info@icasconference.com](mailto:info@icasconference.com)

We cordially invite prospective authors to submit their original papers to ICAS-2022,

- 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](mailto:icasconference.academic@gmail.com)

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

# ICAS 2022 PROGRAM

## 16 May 2022 Monday

Timing: Local Time in Cyprus. GMT+3

**MORNING COFFEE BREAK 11:00 – 11:20**

Welcome Speech :Prof. Dr. Ismihan BAYRAMOGLU, Izmir University of Economics  
11:20 – 11:40 Conference Chair

**REGULAR / FACE TO FACE SESSION A ( 11:40 – 12:40 )**

**SESSION CHAIR: Assoc.Prof. Dr. Aysegul EREM**

*This session will be available in zoom to be followed by participants.*

**11:40 – 12:00**

PAPER TITLE : Bayesian Inference for Skewed Normal Distributions  
AUTHOR(S) : **Narayan SRINIVASAN**<sup>1</sup>, Satyaki MAZUMDER<sup>2</sup>

**12:00 – 12:20**

PAPER TITLE : Generalized Fiducial Inference for the Chen Distribution  
AUTHOR(S) : Cagatay CETINKAYA

**12:20 – 12:40**

PAPER TITLE : Parameter Estimation for a FSDE Model of Glucose Control  
AUTHOR(S) : Alessandro BORRI<sup>1</sup>, **Andrea DE GAETANO**<sup>2</sup>, Simona PANUNZI<sup>3</sup>

**LUNCH BREAK 12:40 – 13:40**

**Information for regular participation; Lunch is included into Salamis Hotel Accommodation**

**E-Plenary Speech A (13:40 – 14:10 )**

**Prof. Dr. Sofia OSTROVSKA**

Speech Title

Independence and Uncorrelatedness:  
From Early Studies to Current Developments

*This speech will be available in zoom to be followed by participants.*

**TECHNIC BREAK 14:10 – 14:20**



<b>ONLINE SESSION A (14:20 – 15:40 )</b>
<b>SESSION CHAIR: Assoc. Prof. Dr. Gulder KEMALBAY</b>

*This session will be available in zoom to be followed by participants.*

**14:20 – 14:40**

PAPER TITLE : On the Generalized Mixed Shock Model  
AUTHOR(S) : Murat OZKUT

**14:40 – 15:00**

PAPER TITLE : Mean Residual Strength of a Multi-state System based on Stress-Strength Modelling  
AUTHOR(S) : **Funda ISCIOGLU**<sup>1</sup>, Firat DEVECI<sup>2</sup>

**15:00 – 15:20**

PAPER TITLE : Comparison of Zero Inflated Regression Models For Application To E-Commerce Data  
AUTHOR(S) : **Duygu KILIC**<sup>1</sup>, Hatice T. Kubra AKDUR<sup>2</sup>, Hulya BAYRAK<sup>3</sup>

**15:20 – 15:40**

PAPER TITLE : Reliability Assessment for Discrete Time Multi-State System Utilizing Level Based Component States  
AUTHOR(S) : **Funda ISCIOGLU**<sup>1</sup>, Ozge ELMASTAS GULTEKIN<sup>2</sup>

<b>COFFEE BREAK 15:40 – 16:00</b>
-----------------------------------

<b>ONLINE SESSION B (16:00 – 17:20 )</b>
<b>SESSION CHAIR: Dr. Krzysztof JASINSKI</b>

*This session will be available in zoom to be followed by participants. Link will be sended on conference day*

**16:00 – 16:20**

PAPER TITLE : Bounds on Expected Order Statistics From the Decreasing Reversed Failure Rate Populations  
AUTHOR(S) : **Agnieszka GORONCY**<sup>1</sup>, Tomasz RYCHLIK<sup>2</sup>

**16:20 – 16:40**

PAPER TITLE : The Number of Failed Components in a Coherent System Consisting of Different Types of Components  
AUTHOR(S) : Krzysztof JASINSKI

**16:40 – 17:00**

PAPER TITLE : Order Statistics Based Training and Scoring Algorithms for Deep Outlier Detection  
AUTHOR(S) : **Ahmet Zahid BALCIOGLU**<sup>1</sup>, Erhan CENE<sup>2</sup>

**17:00 – 17:20**

PAPER TITLE : Assessing Several Strategies for Estimating Average Treatment Effects with Generalized Boosted Models: Matching Weighting via Inverse Probability  
AUTHOR(S) : Hulya KOCYIGIT

TECHNIC BREAK 17:20 – 17:30

E-Plenary Speech B (17:30 – 18:00 )

Prof. Barry C. ARNOLD

Speech Title

Asymmetric Univariate and Bivariate Laplace and  
Generalized Laplace Distributions Revisited

## Gala Dinner at Kybele Restaurant / Bellapais

Salamis Hotel Departure: 18:30

**Gala Dinner invitation is valid only for Full Registration Package holders.  
For Extra Gala Dinner Ticket please contact conference registration desk**

**17 May 2022 Tuesday**

ONLINE SESSION C (10:00 – 11:20 )

SESSION CHAIR: Prof. Dr.Fatma NOYAN TEKELI

*This session will be available in zoom to be followed by participants.*

**10:00 – 10:20**

PAPER TITLE : An it and Statistical Procedure to Merge Administrative Database and Official Statistics

AUTHOR(S) : **Caterina Marini**<sup>1</sup>, Vittorio Nicolardi<sup>2</sup>

**10:20 – 10:40**

PAPER TITLE : A Hybrid Long Short-Term Memory Model For Daily Crude Oil Price Prediction

AUTHOR(S) : Coskun PARIM<sup>1</sup>, **Batuhan Özkan**<sup>2</sup>, Fatma NOYAN TEKELI<sup>3</sup>

**10:40 – 11:00**

PAPER TITLE : Deep Learning In Video Games

AUTHOR(S) : **Alihan ÖZTÜRK**<sup>1</sup>, Fatma NOYAN TEKELI<sup>2</sup>

**11:00 – 11:20**

PAPER TITLE : Portfolio Optimization For Bist 50 Using Local Dependence Approach

AUTHOR(S) : **G.Yazgi TUTUNCU**<sup>1</sup>, Hakan EVGI<sup>2</sup>

**11:20 – 11:40**

PAPER TITLE : Comparison of the claims reserves methods

AUTHOR(S) : **Rümeysa KARATAŞ**<sup>1</sup>, Uğur KARABEY<sup>2</sup>

**COFFEE BREAK 11:40 – 12:00**

**REGULAR / FACE TO FACE SESSION B (12:00 – 13:00 )**

**SESSION CHAIR: Narayan SRINIVASAN**

**12:00 – 12:20**

PAPER TITLE : Bootstrap Aggregation Methods for Feature Selection

AUTHOR(S) : Ayman ALZAATREH

**12:20 – 12:40**

PAPER TITLE : Hybrid Methodology Of Panel Data Analysis Used On Borsa Istanbul Stocks

AUTHOR(S) : **Selim DONMEZ**<sup>1</sup>, Ozer OZAYDIN<sup>2</sup>

**12:40 – 13:00**

PAPER TITLE : A Cross-European Efficiency Assessment of Offshore Wind Farms: A DEA approach

AUTHOR(S) : Negar AKBARI

**13:00 – 13:20**

PAPER TITLE : Hybrid Methodology of Panel Data Analysis Used on Gross Domestic Product of Various Countries

AUTHOR(S) : **Selim DONMEZ**<sup>1</sup>, Ozer OZAYDIN<sup>2</sup>

**LUNCH BREAK 13:20 – 14:30**

**Regular / Face to Face Invited Speech C  
(14:30 – 15:00 )**

**Assoc. Prof. Dr. Aysegul EREM**

Speech Title: Applications of Exceedance Statistics in Control Charts

**CLOSING CEREMONY 15:00 – 15:20**

## Contents

Hybrid Methodology of Panel Data analysis Used on Borsa Istnbul Stocks Selim DÖNMEZ <sup>1</sup> , Özer ÖZAYDIN <sup>2</sup> , .....	14
Order Statistics Based Training and Scoring Algorithms for Deep Outlier Detection .....	
Ahmet Zahid BALCIOGLU, Erhan CENE, .....	15
A Hybrid Long Short-Term Memory Model for Daily Crude Oil Price Predistion .....	
Coşkun PARİM <sup>1</sup> , Batuhan ÖZKAN <sup>2</sup> , Fatma NOYAN TEKELİ <sup>3</sup> .....	16
Parameter estimation for a FSDE model of glucose control.....	
Alessandro BORRI <sup>1</sup> , Andrea DE GAETANO <sup>2</sup> , Simona PANUNZI <sup>3</sup> .....	17
Bootstrap Aggregation Methods for Feature Selection .....	
Ayman ALZAATREH <sup>1</sup> .....	18
An IT and Statistical Procedure to Merge Administrative Database and Official Statistics .....	
Caterina Marini <sup>1</sup> , Vittorio Nicolardi <sup>2</sup> .....	19
GENERALIZED FIDUCIAL INFERENCE FOR THE CHEN DISTRIBUTION.....	
Çağatay ÇETİNKAYA.....	21
COMPARISON OF ZERO INFLATED REGRESSION MODELS FOR APPLICATION TO E-COMMERCE DATA.....	
Duygu KILIÇ <sup>1</sup> , Hatice T. Kübra AKDUR <sup>2</sup> , Hülya BAYRAK <sup>3</sup> .....	22
Reliability Assessment for Discrete Time Multi-State System Utilizing Level Based Component States.....	
Funda ISCIOGLU <sup>1</sup> , Ozge ELMASTAS GULTEKİN <sup>2</sup> , .....	23
Mean Residual Strength of a Multi-state System based on Stress-Strength Modelling.....	
Funda ISCIOGLU <sup>1</sup> , Firat DEVECİ <sup>2</sup> ,.....	25
Bounds on expected order statistics from the decreasing reversed failure rate populations .....	
Agnieszka GORONCY <sup>1</sup> , Tomasz RYCHLIK <sup>2</sup> , .....	27
PORTFOLIO OPTIMIZATION FOR BIST 50 USING LOCAL DEPENDENCE..... APPROACH .....	
G.Yazgı TÜTÜNCÜ <sup>1</sup> , Hakan Evği <sup>2</sup> ,.....	28
DEEP LEARNING in VIDEO GAMES .....	
Alihan ÖZTÜRK <sup>1</sup> , Fatma NOYAN TEKELİ <sup>2</sup> ,.....	29
Assessing Several Strategies for Estimating Average Treatment Effects with Generalized Boosted Models: Matching Weighting via Inverse Probability Weighting.....	
Hulya Kocyigit .....	30
The number of failed components in a coherent system consisting of different types of components .....	
Krzysztof Jasiński <sup>1</sup> .....	31

On the generalized mixed shock model.....	
Murat OZKUT <sup>1</sup> , .....	32
Bayesian Inference for Skewed Normal Distributions .....	
NARAYAN SRINIVASAN1, SATYAKI MAZUMDER2, .....	33
Hybrid methodology of PANEL DATA analysis USED ON GROSS DOMESTIC PRODUCT OF VARIOUS COUNTRIES .....	
Selim DÖNMEZ <sup>1</sup> , Özer ÖZAYDIN <sup>2</sup> , .....	34
Comparison of the claims reserves methods.....	
Rümeysa KARATAŞ <sup>1</sup> , Uğur KARABEY <sup>2</sup> .....	35

# Hybrid methodology of PANEL DATA analysis USED ON BORSA ISTANBUL STOCKS

Selim DÖNMEZ<sup>1</sup>, Özer ÖZAYDIN<sup>2</sup>,

<sup>1</sup>Author's contact address and e-mail: Eskişehir Osmangazi üniversitesi building F1, sdonmez3@gmail.com

<sup>2</sup>Author's contact address and e-mail: Eskişehir Osmangazi üniversitesi building F1,oozaydin@ogu.edu.tr

## Abstract

We used a hybrid methodology of linear regression on panel data and panel data clustering on stocks of Borsa Istanbul. Our aim was to check the efficiency of the panel data clustering in terms of linear regression used on panel data. Providing convincing results from this study would yield the necessary groundwork to forecast the stock movements based on panel data clustering. In our understanding, the base model necessary for forecasting stock movements is the one linked to the one we provide results from. Our model we develop in this study is Platonic in the sense that it describes the general principle of the stock movements. The results we obtained from our model demonstrated the effectiveness of the methodology as it provided high predictivity of stock movements up to %94.

**Key Words:** *Panel Data; Borsa Istanbul; Stock Movements; Clustering; Linear Regression*

# Order Statistics Based Training and Scoring Algorithms for Deep Outlier Detection

Ahmet Zahid BALCIOGLU, Erhan CENE,

<sup>1</sup>Department of Statistics, Faculty of Arts and Sciences, Yildiz Technical University, Turkey;  
zahidbalcioglu@gmail.com

<sup>2</sup>Department of Statistics, Faculty of Arts and Sciences, Yildiz Technical University, Turkey;  
ecene@yildiz.edu.tr

## Abstract

Autoencoder networks are among the most popular deep learning methods for outlier detection. They work as an unsupervised algorithm that aims to reconstruct original instances through learning a low-dimensional feature space. One of the main weaknesses of outlier detection using autoencoders is that the training loss metric is also used for outlier reconstruction errors, which can lead models to learn outliers just as well as normal instances. In this paper, we propose to solve this problem by introducing a robust order statistics based scoring system. We discuss the chosen scoring system and demonstrate its effectiveness through simulation studies. We also show that with our proposed scoring system autoencoder model no longer needs the percentage of anomalies as a hyperparameter to estimate the cut-off point. We also introduce an order statistics augmented loss function, which creates a more robust model by penalizing the ‘leveraging effect’ of outliers in the loss signal. We show that training with our algorithms show a significant improvement when compared to conventional autoencoder training on image and medical datasets.

**Key Words:** *anomaly detection, outlier detection, order statistics, robust statistics, autoencoder*

# A HYBRID LONG SHORT-TERM MEMORY MODEL FOR DAILY CRUDE OIL PRICE PREDICTION

Coşkun PARİM<sup>1</sup>, Batuhan ÖZKAN<sup>2</sup>, Fatma NOYAN TEKELİ<sup>3</sup>

<sup>1</sup> Yıldız Technical University, Faculty of Art & Science, Department of Statistics, [cparim@yildiz.edu.tr](mailto:cparim@yildiz.edu.tr)

<sup>2</sup> Yıldız Technical University, Faculty of Art & Science, Department of Statistics, [baozkan@yildiz.edu.tr](mailto:baozkan@yildiz.edu.tr)

<sup>3</sup> Yıldız Technical University, Faculty of Art & Science, Department of Statistics, [fnoyan@yildiz.edu.tr](mailto:fnoyan@yildiz.edu.tr)

## Abstract

One of the most important indicators affecting the economy of countries is the price of crude oil. For this reason, it is an undeniable fact that estimation will provide significant benefits. In this study, a hybrid model combining wavelet transform (WT) and long short-term memory (LSTM) is proposed to predict daily crude oil price (West Texas Intermediate). Eight different wavelet transform methods were used as Biorthogonal 3.7 (bior3.7), Coiflets 3 (coif3), Daubechies 2 (db2), Discrete Meyer (dmey), Fejer-Korovkin 4 (FK4), Haar, Reverse biorthogonal 5.5 (rbio5.5), and Symlets 7 (sym7). The time series was decomposed by obtaining 4 details and 1 approximation from each transformation. The data set is divided chronologically into 75% training and 25% test datasets, respectively. The one day ahead of the crude oil price (t+1) were estimated using LSTM with the help of the daily values (t) of the wavelet transforms. The LSTM results with all wavelets transform show that 99% of the daily crude oil price is predicted. The proposed hybrid model has could guide experts in daily crude oil price estimation.

***Key Words: crude oil price; lstm; wavelet transform; hybrid models; deep learning***



# Parameter estimation for a FSDE model of glucose control

Alessandro BORRI<sup>1</sup>, Andrea DE GAETANO<sup>2</sup>, Simona PANUNZI<sup>3</sup>

<sup>1</sup>alessandro.borri@biomatematica.it, CNR IASI Roma - ITALY

<sup>2</sup>andrea.degaetano@biomatematica.it, CNR IASI Roma & CNR IRIB Palermo - ITALY

<sup>3</sup>simona.panunzi@biomatematica.it, CNR IASI Roma - ITALY

## Abstract

In the quest for progressively more realistic mathematical representations of biomedical processes, fractional differential equations have the potential of summarizing, with an order that can in principle be estimated from data, different (presumably integer-order) interacting controls or influences upon the observed variable of interest. This is the case, for example, of transcutaneously measured glycemia, where besides glycemia itself (possibly decaying by first-order elimination) also unobserved factors (insulinemia, other hormones) may exert higher order effects. The problem is complicated by the fact that random events (food intake, exercise, emotions) may affect glycemia as well, leading to the formalization of the problem as a Fractional Stochastic Differential Equation. We use a simple FSDE model of glyceemic control to exemplify a possible approach to model parameter estimation in this framework.

**Key Words:** *Continuous Glucose Monitoring; Stochastic Differential Equations; Fractional Differential Equations; Parameter Estimation; Fractional Brownian Motion.*

# Bootstrap Aggregation Methods for Feature Selection

Ayman ALZAATREH<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, American University of Sharjah, aalzaatreh@aus.edu

## **Abstract**

In this talk, a general framework for the ensemble of multiple feature selection methods is proposed. Based on diversified datasets generated from the original set of instances, the importance scores generated by multiple feature selection techniques are aggregated using two methods: Within Aggregation Method (WAM) which refers to aggregating importance scores within a single feature selection and Between Aggregation Method (BAM) which refers to aggregating importance score between multiple feature selection methods. Furthermore, the stability of the aggregation methods will be discussed. The proposed framework on several real datasets with diverse performance and characteristics are applied. The experimental evaluation showed greater effectiveness of WAM than BAM in terms of stability of the selected features. WAM was also shown to provide effective tool to determine the best feature selection method for a given dataset. Furthermore, the combined feature selection framework helped identifying the optimal reduction percentage in the number of important features.

***Key Words: Bootstrap; Feature Selection; Filter Method; Entropy; Stability***

# AN IT AND STATISTICAL PROCEDURE TO MERGE ADMINISTRATIVE DATABASE AND OFFICIAL STATISTICS

Caterina Marini<sup>1</sup>, Vittorio Nicolardi<sup>2</sup>

<sup>1</sup> Department of Economics and Finance, University of Bari Aldo Moro, Bari - Italy, e-mail: caterina.marini@uniba.it

<sup>2</sup> Department of Economics and Finance, University of Bari Aldo Moro, Bari - Italy, e-mail: vittorio.nicolardi@uniba.it

## Abstract

The availability of a complete harmonized database that includes all available information on the phenomenon under investigation is a precious analytical instrument very often unavailable because of issues related to important obstacles that involve both statistical and computer problems, and administrative and legal impediments. At the same time, the enormous flow of information that in an IT world is available needs to be accurately handled because that is a very important opportunity to deeply analyse under a new perspective the social and economic change in the nowadays society. In fact, the official statistics that the National Institutes of Statistics yield are facing their limits in all the research areas and the problem is already recognised and discussed in the international scientific context. Undoubtedly, the Public Administration databases can be a very useful source of additional and detailed data to complete the statistical information about phenomena that are still partially depicted through the official data. And, in some cases, limitations in dealing with the real size of many socio-economic territorial developments are significant. This is the case, for instance, of phenomena such as health, labour forces, well-being, real estate. The debate about the administrative databases and their use to provide a statistical information that is not differently affordable is internationally considered significant and relevant to face the nowadays scientific challenge. It is believed that the administrative data are largely available and readily usable, and this is one the most common errors that is unconsciousness committed. When working with administrative data, the main problem is the integrated use of data that are yielded by independent administrative offices, sometimes belonging to the same organisation, public authority or company, and differently collected for specific administrative purpose. Furthermore, in most cases databases are structurally diverse because of well-established IT and administrative procedures that are not updated to align all information for users. As a consequence of their nature, administrative data can be affected by material and human errors, such as duplication, missing values and erroneous information, and therefore they are not controlled and statistically correct. Another aspect that is important to consider is that administrative data are a little part of the more complex phenomenon known as Big Data. In literature, the Big Data aspect of the problem is widely discussed ([1], [2], [3], [4], [5]). Finally, it is crucial the harmonisation of the administrative data, which are raw information, with the official statistical data as yielded by the National Statistical Institutes and, therefore, consolidated statistically. In some of our recent works, we provided an original fully comprehensive statistical tool to describe the phenomenon of the real estate economy. The database we created is referred to the real estate phenomenon where administrative data and official statistical data are perfectly aligned and the Big Data crucial aspects solved. In this work, we depict the IT and Statistical procedure used to provide the first Full Information Harmonised Real Estate Database in the analysis of the real estate economy.

**Key Words:** *Administrative Data; Big Data; Real Estate Economy; Harmonised Database; IT algorithm.*

## **References**

[1] Hadford, T. (2014) Big data: a big mistake? *Significance*, 11, 14–19.

[2] Kitchin, R. (2014) Data, new epistemologies and paradigm shift. *Big Data & Society.*, 1.

[3] Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data. Data Infrastructures and Their Consequences*: Sage Publications, London.

[4] Marini C., Nicolardi V. (2020) Big Data and Economic Analysis: The Challenge of a Harmonized Database, *Data Science and Social Research II. DSSR 2019. Studies in Classification, Data Analysis, and Knowledge Organization*, Springer Nature.

[5] Marini C., Nicolardi V. (2021) Administrative database and official statistics: The case of the real estate analysis. *Statistica Applicata - Italian Journal of Applied Statistics* - vol. 33 (1), pp.83-95.

# GENERALIZED FIDUCIAL INFERENCE FOR THE CHEN DISTRIBUTION

Cağatay ÇETİNKAYA

Department of Accounting and Taxation, Bingöl University, 12000, Bingöl, Turkey

## Abstract

The fiducial inference idea was firstly proposed by Fisher (1930) as a powerful method in statistical inference. Many authors such as Weeranhandi (1993) and Hannig et. al. (2006) improved this method from different points of view. Since the Bayesian method has some deficiencies such as assuming a prior distribution when there was little or no information about the parameters, the fiducial inference is used to overcome these adversities. This study deals with the generalized fiducial inference for the shape parameters of the Chen' s two-parameter lifetime distribution with bathtub shape or increasing failure rate (Chen, 2000). The method based on the inverse of the structural equation which is proposed by Hannig et. al. (2016) is used. We propose the generalized fiducial inferences of the parameters with their confidence intervals. Then, these estimations are compared with their maximum likelihood and Bayesian estimations. Simulation results show that the generalized fiducial inference is more applicable than the other methods in terms of the performances of estimators for the shape parameters of the Chen distribution. Finally, a real data example is used to illustrate the theoretical outcomes of these estimation procedures.

**Key Words:** Bayesian inference; generalized fiducial inference; interval estimation; Chen distribution; point estimation

## References

- [1] Chen ZA. (2000). New two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics & Probability Letters*, 49(2), 155-161.
- [2] Fisher RA. (1930). Inverse probability. *Proc. Camb. Philos. Soc.*, xxvi: 528–535.
- [3] Hannig J, Iyer HK, Patterson P. (2006). Fiducial generalized confidence intervals. *J. Am. Stat. Assoc.*, 101, 254–269.
- [4] Hannig J, Iyer H, Lai RC, Lee TC. (2016). Generalized fiducial inference: A review and new results. *Journal of the American Statistical Association*, 111 (515), 1346-1361.
- [5] Weeranhandi S. (1993). Generalized confidence intervals. *J. Am. Stat. Assoc.*, 88, 899–905.

# COMPARISON OF ZERO INFLATED REGRESSION MODELS FOR APPLICATION TO E-COMMERCE DATA

Duygu KILIC<sup>1</sup>, Hatice T. Kübra AKDUR<sup>2</sup>, Hülya BAYRAK<sup>3</sup>

<sup>1</sup>Gazi University, Faculty of Sciences, Department of Statistics, Ankara, Turkey e-mail: duygukilic4@gmail.com

<sup>2</sup>Gazi University, Faculty of Sciences, Department of Statistics, Ankara, Turkey e-mail: [hatice\\_senol@wsu.edu.tr](mailto:hatice_senol@wsu.edu.tr)

<sup>3</sup>Gazi University, Faculty of Sciences, Department of Statistics, Ankara, Turkey e-mail: hbayrak@gazi.edu.tr

## Abstract

Count data is often used in regression models. The word "count" is generally used to rank units, items or events [1]. Having more zeros than expected in count datasets is defined as zero inflation. It is more appropriate to analyse such datasets with zero-inflated models that take into account zeros [2]. Zero-inflated models are used in different fields such as econometrics, demography, medicine, public health, epidemiology, biology. Failure to use appropriate methods in the analysis of zero-inflated data can lead to biased parameter estimates, smaller standard errors, and inconsistent results [3]. Zero-inflated models take different names according to the distribution of the response variable. In this study, in addition to the well-known zero-inflated models (zero inflated Poisson and zero inflated negative binomial regression models), the zero inflated Bell model proposed by Lemonte et al., which is quite new in the statistics literature, is utilized for analysing the e-commerce dataset [4].

In this case study, it is aimed to determine the factors affecting the number of problems experienced in e-commerce for Turkey by using zero-inflated count models. For this purpose, zero-inflated Poisson, zero-inflated negative binomial, and zero-inflated Bell regression models are utilized in analysing 2019 TurkStat household information technology survey with the Cran R program. Using Akaike Information Criteria and log-likelihood based goodness-of-fit tests from these models, it is determined which model represented the data set better. According to the results of the analysis, it is observed that the zero-inflated negative binomial model fits the dataset better. In addition, the parameter estimations of the zero-inflated negative binomial model are examined and interpreted as a result of the analysis.

**Key Words:** *Count data, zero inflated models, e-commerce, regression analysis*

## References

- [1] Hilbe, J.M. (2014). Modeling Count Data (First Edition), New York: Cambridge University Press.
- [2] Ridout, M., Hinde, J. and Demetrio, C.G.B. (2001). A Score Test for a Zero-Inflated Poisson Regression Model Against Zero Inflated Negative Binomial Alternatives. *Biometrics* 57, 219-233.
- [3] Miller, J., M., (2007). Comparing Poisson, Hurdle, And Zero Inflated Poisson Model Fit Under Varying Degrees Of Skew And Zero-Inflation. (pHd Thesis). University of Florida.
- [4] Lemonte, A. J., Moreno-Arenas, G., & Castellares, F. (2019). Zero-Inflated Bell Regression Models For Count Data. *Journal of Applied Statistics*.

# Reliability Assessment for Discrete Time Multi-State System Utilizing Level Based Component States

Funda ISCIOGLU<sup>1</sup>, Ozge ELMASTAS GULTEKIN<sup>2</sup>,

<sup>1</sup>Department of Statistics, Ege University, 35040, Bornova, Izmir, Turkey, e-mail:funda.iscioglu@ege.edu.tr

<sup>2</sup>Department of Statistics, Ege University, 35040, Bornova, Izmir, Turkey, e-mail: ozge.elmastas@ege.edu.tr

## Abstract

Multi-state modelling is more flexible to use in many real-life applications. Different from binary-state modelling, the system can have many performance levels from perfect functioning to complete failure. There have been many researches on the reliability analysis of multi-state systems [4]. We consider a system and its components having just three-states in this study. State ‘2, 1 and 0’ indicates, completely working, partially working and failure states, respectively. We especially, consider a generalized three-state k-out-of-n:G system structure[3]. Thus, providing the system to be at least a certain level, we might require a different number of components to be at certain state or above.

Most of the studies deal with the dynamic reliability analysis of systems where continuous lifetime random variables are taken into account [5]. However, a remarkable amount of work in literature deals with some reliability problems where discrete types of distributions are used [1, 2]. Therefore, we consider geometric distributed failure times of the components in the system. Many works using discrete distributions generally examine repairable systems due to the various performance measures such as mean time to system failure (MTSF), availability, etc. However, the components and the system are assumed to be non-repairable in this work.

In most of the engineering applications, the system state can be accurately observed whereas component states are unobservable. Besides, different combinations of component states result with the same state of the system. Thus, the knowledge of the states of the components during the degradation process of the system can enable us to update the reliability function of the system or estimate the system’s further stochastic behaviour and remaining useful life[6]. Also, this knowledge can be beneficial for planning the maintenance strategies. Therefore, in this study we focus on the reliability analysis of a generalized three-state k-out-of-n:G system via considering level based component states. The reliability analysis of this system is conducted through the medium of the transitions between different levels of the component states. Both the generalized and steady state transition probabilities of the states for the related system are obtained. The recursive equations for the reliability functions of the system are provided.

**Key Words:** *multi-state system; k-out-of-n:G system; geometric distribution; reliability analysis, transition probabilities*

## References

[1] Bahatti J, Chitkara AK, Kakkar M (2013) Profit Analysis of Non-Identical Parallel System with Two Types of Failure Using Discrete Distribution, Mathematical Journal of Interdisciplinary Sciences, 1(2), 27-40.

- [2] Eryilmaz S (2013) On mean Residual Life of Discrete Time Multi-State Systems. *Quality Technology&Quantitative Management*, 10(2), 241-250.
- [3] Huang J, Zuo MJ and Wu Y (2000) Generalized multi-state k-out-of-n:G systems. *IEEE Trans Reliab* , 49, 105-111.
- [4] Kuo W and Zuo MJ (2003) *Optimal reliability modeling:principles and applications*. New Jersey:John Wiley&Sons.
- [5] Lisnianski A and Levitin G (2003) *Multi-state system reliability: Assessment, Optimization Applications*. Singapore:World Scientific Pub. Co. Inc.
- [6] Liu Y, Zuo MJ, Li YF, Huang HZ (2015) Dynamic Reliability Assessment for Multi-State Systems Utilizing System-Level Inspection Data, *IEEE Transactions on Reliability*, 64(4), 1287-1299.



# Mean Residual Strength of a Multi-state System based on Stress-Strength Modelling

**Funda ISCIOGLU<sup>1</sup>, Firat DEVECİ<sup>2</sup>,**

<sup>1</sup>Department of Statistics, Ege University, 35040, Bornova, Izmir, Turkey, e-mail:funda.iscioglu@ege.edu.tr

<sup>2</sup>Department of Statistics, Ege University, 35040, Bornova, Izmir, Turkey, e-mail: firat.deveci12@gmail.com

## Abstract

In the reliability analysis, stress-strength modelling has an important area of research especially in real life applications. There has been many researches on stress-strength reliability analysis of systems based on binary structures [3,5]. Multi-state systems and their reliability evaluation problems have also been substantially dealt with in the literature. That is, because multi-state systems are more flexible to use in many engineering systems and they help us to understand the behaviour of many mechanical systems such as production, manufacturing, computer or transportation of oil and gas systems [6]. However, considering multi-state structures, stress-strength reliability has not been investigated much in the literature. There are a few articles considering the problem [2]. Especially for all system types binary or multi-state, dynamic performance measures such as mean residual lifetime, mean past lifetime are examined due to the fact that they are important characteristics for evaluating a system's performance [1].

In stress-strength reliability analysis, mean remaining strength is one of the important performance characteristics of a system and it is defined as the expected lifetime of that system in terms of its strength till its failure. In one of the research articles [4], they evaluated the mean remaining strength for a k-out-of-n, parallel and series systems in the context of binary states of the system and its components. However, within this work, for a multi-state system based on stress-strength reliability, mean remaining strength is newly evaluated. First, we examine the mean remaining strength for a single component. Then, we evaluate the mean remaining strength of a multi-state stress-strength model with a k-out-of-n:G structure. In both analyses, we assume that the stress and strengths are independent random variables. We examine the results based on different parameters of strength and stress. We discuss the results considering the increase in the parameters and the values of  $n$  and  $k$ . The results are also supported with some simulations.

**Key Words:** *multi-state system; stress-strength reliability; mean residual strength; k-out-of-n:G system*

## References

- [1] Asadi, M. and Bairamov, I. (2008) The Mean Residual Life Function of a k-out-of-n Structure at the System Level, IEEE Transactions on Reliability, 55(2), 314-318.
- [2] Eryilmaz, S and Iscioglu, F (2011) Reliability Evaluation for a Multi-State System Under Stress-Strength Setup, Communications in Statistics—Theory and Methods, 40:3, 547-558.

- [3] Eryılmaz, S (2013) On Stress-Strength Reliability with a Time-Dependent Strength, *Journal of Quality and Reliability Engineering*, Article ID 417818, 6 pages.
- [4] Gurler, S (2013) The Mean Remaining Strength of Systems in a Stress-Strength Model, *Hacettepe Journal of Mathematics and Statistics*, 42(2), 181-187.
- [5] Kotz, S and Lumelskii, Y (2003) *The Stress-Strength Model and its Generalizations: Theory and Applications*, World Scientific, Singapore.
- [6] Tian, Z, Zuo, MJ and Yam RCM (2009) Multi-state k-out-of-n systems and their performance evaluation, *IIE Transactions*, 41, 32-44.

# **Bounds on expected order statistics from the decreasing reversed failure rate populations**

**Agnieszka GORONCY<sup>1</sup>, Tomasz RYCHLIK<sup>2</sup>,**

<sup>1</sup>[gemini@mat.umk.pl](mailto:gemini@mat.umk.pl)

<sup>2</sup>[trychlik@impan.pl](mailto:trychlik@impan.pl)

## **Abstract**

We present the sharp positive upper mean-variance bounds on the expectations of order statistics with relatively high ranks based on independent identically distributed random variables with the decreasing reversed failure rate. We also give the attainability conditions. The theoretical results are illustrated with numerical values of the bounds.

***Key Words:*** *order statistic, bound, decreasing reversed failure rate*

# PORTFOLIO OPTIMIZATION FOR BIST 50 USING LOCAL DEPENDENCE APPROACH

**G.Yazgı TÜTÜNCÜ<sup>1</sup>, Hakan Evği<sup>2</sup>,**

<sup>1</sup>Izmir University of Economics, Department of Mathematics, [yazgi.tutuncu@ieu.edu.tr](mailto:yazgi.tutuncu@ieu.edu.tr)

<sup>2</sup> Izmir University of Economics, Department of Mathematics, [hakan.evgi@std.izmirekonomi.edu.tr](mailto:hakan.evgi@std.izmirekonomi.edu.tr)

## **Abstract**

A portfolio is an asset consisting a range of investments that have been put together for a particular goal. For investors to make a great profit, portfolio selection is a very important task. In general, investors want to choose the best portfolio that has the lowest risk and highest return. Given the characteristics of stocks in a stock exchange market, portfolio management is about determining the most desirable collection of stocks to purchase.

In this study, we investigate modern portfolio theory based on Markowitz model [1] and we propose a new portfolio optimization approach based on local dependence function based [2]. Using the well-known Markowitz modern portfolio theory, two optimum portfolios generated with the stocks of companies traded in the Turkish BIST 50 index and the data collected from February 3rd, 2020 to November 1st, 2021. Then we applied our proposed approach to the same data set and proposed two new optimum portfolios. Comparisons and sensitivity analysis presented and some interpretations and future studies suggested.

***Key Words: Portfolio selection, optimization, local dependence function, Markowitz model***

## **References (Times New Roman, 12pt, Bold)**

[1] Steinbach, MC (2001). Markowitz revisited: Mean-variance models in financial portfolio analysis. SIAM REVIEW, Volume:43, Issue:1, Pages:31-85, DOI: 10.1137/S0036144500376650.

[2] Kochan, N, Tutuncu, GY, Giner, G (2021). A new local covariance matrix estimation for the classification of gene expression profiles in high dimensional RNA-Seq data. EXPERT SYSTEMS WITH APPLICATIONS, Volume167, Article Number114200, DOI10.1016/j.eswa.2020.114200.

# DEEP LEARNING in VIDEO GAMES

Alihan ÖZTÜRK<sup>1</sup>, Fatma NOYAN TEKELİ<sup>2</sup>,

<sup>1</sup>Department of Statistics, Faculty of Arts and Sciences, Yildiz Technical University, Turkey  
alihan.masquer@gmail.com

<sup>2</sup>Department of Statistics, Faculty of Arts and Sciences, Yildiz Technical University, Turkey  
fnoyan@yildiz.edu.tr

## Abstract

Reinforcement learning (RL) is currently one of the most dynamic areas in Artificial Intelligence research. It is the technique in which an agent learns how to obtain rewards through interactions with their environment. Some real-world applications, such as robotics and autonomous, are particularly suitable for the RL approach as the environment is unknown and the consequences of actions are uncertain. Because many Reinforced Learning examples in the literature are trained with simple and intuitive reward functions, it struggles with sparse and deceptive feedback. In addition, it may be necessary to search for different ways for trial and error and sometimes return to the previous steps after discovering these ways, so it is important to use the right search algorithms. Discovering all paths in remains one of the challenges of the field. In Reinforced Learning, it may not be the optimum solution to give high values to the activities we want to take place or to apply a penalty system to undesirable activities. However, there is no proven method that selects the appropriate reward function for every situation. The reward functions used are intuitive. In this study, it is planned to train the Reinforced Learning algorithm in a game environment by giving different reward functions. It is aimed to examine the performances and strategies of the learning algorithms and to make comparisons between them. However, for this, it is necessary to have environments where we can directly access the situations and actions in the game, and the diversity of decisions that lead to the result in the game for the reward function. DOOM 2 game is suitable for this study is found by examining the environments in the literature. In this game there are multiple chapters and different obstacles that will ensure that the chapter designs do not reach the goal directly. It is aimed to compare the effects of department diversity and barriers on education, as well as to examine the performance and strategies of representatives trained in the same departments with different reward functions and to make comparisons between them.

*Key Words: deep learning; reinforcement learning; video games; deep q learning*

## References

- [1] Baker, B., Kanitscheider, I., Markov, T., Wu, Y., Powell, G., McGrew, B., & Mordatch, I. (2019). Emergent tool use from multi-agent autocurricula. arXiv preprint arXiv:1909.07528.
- [2] Khorasgani, H., Wang, H., & Gupta, C. (2020). Challenges of applying deep reinforcement learning in dynamic dispatching. arXiv preprint arXiv:2011.05570.
- [3] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.
- [4] Badia, A. P., Piot, B., Kapturowski, S., Sprechmann, P., Vitvitskyi, A., Guo, Z. D., & Blundell, C. (2020, November). Agent57: Outperforming the atari human benchmark. In International Conference on Machine Learning (pp. 507-517). PMLR.

# Assessing Several Strategies for Estimating Average Treatment Effects with Generalized Boosted Models: Matching Weighting via Inverse Probability Weighting

Hulya Kocyigit  
Karamanoglu Mehmetbey University

E-mail: [hk20902@gmail.com](mailto:hk20902@gmail.com)

## Abstract

Observational studies are frequently used to explore the connection between treatments and outcomes in many fields. Nevertheless, unless statistical methods are applied to account for the unbalance of confounders among treatment groups, they do not enable inferences about causal connections to be reached. The propensity score is a measure used to minimize imbalances in the distribution of confounding variables between treatment and control groups in studies evaluating the exposure effects on the outcome. Recent work has suggested that matching weighting (MW) outperforms inverse probability weighting (IPW) in linear settings. Researchers, on the other hand, require guidance in determining the best analytic method for any particular circumstance to estimate treatment effect based on the use generalized boosted model. To address this gap, the objectives of this study are threefold:(1) to present step-by-step instructions for researchers who aim learning IPW and MW for various treatments and outcomes, (2) to examine the two techniques (i.e., IPW and MW) in terms of unnecessary covariates, the complexity of the interaction between treatment or outcome and covariates, and residual variance, bias, and standard mean difference (SMD) in outcome and treatment, and (3) to propose the use of GBM to estimate the appropriate weights. This paper conducted a Monte Carlo simulation to provide clarity to analysts when determining between the two methods in GBM under various scenarios of outcomes or treatment based on the different covariates. Beside, this study describes the conditions under which one approach should be assumed to outperform the other, but we make no blanket assumption about whether one method is usually preferable to the other.

**Keywords:** *matching weighting, balancing weight, observational study, variable selection, GBM*

# The number of failed components in a coherent system consisting of different types of components

Krzysztof Jasiński<sup>1</sup>

<sup>1</sup>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 operating (see [1]) or in a failed coherent system (see [2]) . We assume that its components are of multiple types. That is, the system consists of components having nonidentical failure time distributions. We extend the results which are well-known in the literature for k-out-of-n systems. The obtained results are next used in the classical age replacement policy when the corrective replacement of the system is applied if a system has failed, while the preventive replacement is done before failure. We formulate the optimization problem. We wish to find the optimal replacement time.

*Key Words:* Coherent system; multi types of components; age replacement policy; optimal replacement time; geometric distribution

## References

- [1] Jasiński K (2022) On the number of failed components in a coherent system consisting of multiple types of components. Journal of Computational and Applied Mathematics 410:114189.
- [2] Jasiński K (2022) A study on the number of failed components in a failed coherent system consisting of different types of components, prepared.

# On the generalized mixed shock model

Murat OZKUT<sup>1</sup>,

<sup>1</sup>Izmir University of Economics, Department of Mathematics

murat.ozkut@ieu.edu.tr

## Abstract

A generalized mixed shock model that combines two run shock models was introduced and studied. According to the model, the system subject to shocks fails upon  $k_1$  consecutive shocks whose magnitude is between predefined critical values  $d_1$  and  $d_2$  such that  $d_1 < d_2$ , or  $k_2$  consecutive shocks whose magnitude is above  $d_2$ . In particular, the reliability, mean time to failure of the system, and mean residual lifetime are obtained by assuming that the times between the arrivals of shocks follow the phase-type distribution. The optimal replacement policy is also proposed.

***Key Words: run shock model, reliability, mean residual lifetime, phase-type distribution, optimal replacement time***

## References

- [1] Eryilmaz S and Tekin M (2019) Reliability evaluation of a system under a mixed shock model, *Journal of Computational and Applied Mathematics* 352, 255–261.
- [2] Ozkut M and Eryilmaz S (2019) Reliability analysis under Marshall-Olkin run shock model, *Journal of Computational and Applied Mathematics* 349, 52–59.
- [3] He QM (2006) *Fundamentals of matrix-analytic methods*. Springer-New York.



# Bayesian Inference for Skewed Normal Distributions

NARAYAN SRINIVASAN<sup>1</sup>, SATYAKI MAZUMDER<sup>2</sup>,

<sup>1</sup>Indian Institute of Science Education Research Kolkata, West Bengal, India, e-mail: nss16ms153@iiserkol.ac.in

<sup>2</sup>Indian Institute of Science Education Research Kolkata, West Bengal, India, e-mail: satyaki@iiserkol.ac.in

## Abstract

Standard statistical inference procedures lead to erroneous results with pronounced skewness in data. This has led to the need to model skewness, which has given rise to several skewed probability distributions (see Ghaderinezhad et. al [1], for example). Foremost being the Skew Normal Distribution proposed by Azzalini [2]. However, the distribution has a few problems related to Maximum Likelihood Estimation of the parameters. Kundu [3] has proposed an alternate skew distribution, the Geometric Skew Normal distribution (GSN) for modelling skewed data. It is seen that the Bayesian inferential paradigm has several advantages for the Azzalini Skew Normal Distribution (see Ghaderinezhad et. al [1]). To compare the two skew normal distributions in Bayesian paradigm, in this article, two Bayesian inference techniques have been proposed for the GSN distribution. The first technique uses a MCMC algorithm based on a hybrid Metropolis-Hastings Gibbs sampler. Two approaches have been discussed using a geometric proposal distribution (Redivo et al [4]) and a novel random walk algorithm. The second technique is an approximate inference method using Variational Bayes. A variety of simulation studies have been performed to verify the convergence of the algorithms. An extensive comparison of the Azzalini's skew normal distribution and GSN distribution for modelling skewed data has been undertaken in Bayesian paradigm using simulated as well as real data.

**Key Words:** *Skewness; Bayesian inference; Skew-Normal; MCMC; Variational Bayes*

## References

- [1] Ghaderinezhad, F., Ley, C., and Loperfido, N. (2020). Bayesian inference for skew-symmetric distributions. *Symmetry*, 12:491
- [2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2):171–178.
- [3] Kundu, D. (2014). Geometric skew normal distribution. *Sankhyā: The Indian Journal of Statistics, Series B (2008-)*, 76(2):167–189
- [4] Redivo, E., Nguyen, H. D., and Gupta, M. (2020). Bayesian clustering of skewed and multimodal data using geometric skewed normal distributions. *Computational Statistics Data Analysis*, 152:107040.

# Hybrid methodology of PANEL DATA analysis USED ON GROSS DOMESTIC PRODUCT OF VARIOUS COUNTRIES

Selim DÖNMEZ<sup>1</sup>, Özer ÖZAYDIN<sup>2</sup>,

<sup>1</sup>Author's contact address and e-mail: Eskişehir Osmangazi üniversitesi building F1, sdonmez3@gmail.com

<sup>2</sup>Author's contact address and e-mail: Eskişehir Osmangazi üniversitesi building F1,oozaydin@ogu.edu.tr

## Abstract

We used a hybrid methodology of linear regression on panel data and panel data clustering on gross domestic product per year of various countries between 1968-2015. Our aim was to see the difference between the growth trends of the countries and evaluate their numerically expressed growths according to their significance. The countries used in the data are Afghanistan, Albania, Argentina, Australia, Austria, Burundi, Belgium, Benin, Burkina Faso, Bangladesh, Bulgaria, Bolivia, Brasil, Barbados, Botswana, Central African Republic, Canada, Switzerland, Chile, China, Cote d'Ivor, Cameroon, Congo, Democratic Republic of Congo, Comoros, Cabo Verde, Costa Rica, Czech Republic, Cyprus, Germany, Djibouti, Dominica, Dominican Republic, Equador, Egypt, Ethiopia, Spain, France, Finland, Gabon, United Kingdom, Ghana, Guinea, Guinea-Bissau, Greece, Guatemala, Hong Kong and Honduras. Even though these countries can generally be categorized into developing and developed countries, we wanted to evaluate the clustering algorithm on the basis of coming up with the amount of clusters just enough to make a meaningful distinction between countries other than the aforementioned subjective categorization. Our findings showed us that all countries can be expressed in 6 clusters even though there are minor differences between the clusters.

**Key Words:** *Panel Data; Gross Domestic Product; Growth Trends; Clustering; Linear Regression*

## Comparison of the claims reserves methods

Rümeysa KARATAS<sup>1</sup>, Uğur KARABEY<sup>2</sup>,

<sup>1</sup> Hacettepe University

<sup>2</sup> Karamanoğlu Mehmetbey University, rmyskrts@gmail.com

### Abstract

It is very important for insurance companies to estimate accurately their future requires reserves. Different methods of reserving for claim liabilities of an insurer have been advanced throughout time. There are many methods classified according to the their assumptions. These methods attempt to provide an estimate that can meet future liabilities of insurance companies. In this study we make comparison of the claims reserves methods (chain ladder, bornhuetter ferguson and bornhuetter ferguson methods) with car insurance data.

**Keywords:** Outstanding claims reserves, IBNR, claims reserving methods