

ISDC2018
ABSTRACTS BOOK

ISBN: 978-605-030-311-7

CONTENTS

| | |
|---|----|
| 1. ORGANIZING COMMITTEE..... | 9 |
| 2. SCIENTIFIC COMMITTEE | 10 |
| 3. LIST OF PRESENTER PARTICIPANTS | 13 |
| 4. LIST OF ALL PARTICIPANTS | 18 |
| 5. CONFERENCE PROGRAMME..... | 24 |
| 6. ABSTRACTS OF INVITED SPEAKERS..... | 43 |
| A New Difference-based Weighted Mixed Liu Estimator in Partially Linear Models..... | 43 |
| First Passage Time of Degradation Process..... | 44 |
| DİNAMİK META-SEZGİSEL OPTİMİZASYON VE UYGULAMA ÖRNEKLERİ | 45 |
| Robust Bayesian Relevance Vector Machines in Regression and Supervised Classification Using Information Complexity and the Genetic Algorithm..... | 46 |
| A New Multiple Model Approach based on Artificial Neural Network for Forecasting..... | 47 |
| Compound Sums | 48 |
| Bayesian Joint Estimation of Primary and Substitute Demands wit Lost Sales..... | 49 |
| Alternatives to Least Squares Estimation in the Linear Regression Model..... | 50 |
| Data Analysis in Health and Big Data | 52 |
| Statistics and Industry: Opportunities, Responsibilities and Examples..... | 53 |
| An Overview of Methods for Social Network Monitoring..... | 54 |
| A Stochastic Fuzzy Time Series Forecasting Model..... | 55 |
| Using Concomitants or Order and Record Statistics in Directional Dependence Modeling | 56 |
| 7. ABSTRACTS OF ALL PARTICIPANTS | 57 |
| A Bayesian Generalized Linear Model for Crimean–Congo Hemorrhagic Fever Incidents | 57 |
| A Comparative Application of Multiple Correspondence and Joint Correspondence Analysis: A Case Study of Turkey..... | 58 |
| A Comparative Study of Internal Validity Indices for Correlated Data with K-means Algorithm | 59 |
| A Comparative Study on the Robust Estimators for the Parameters of the General Linear Regression Model via Simulation | 60 |
| A Comparison of Stochastic and Deterministic Algorithms on a Ranking Task..... | 61 |
| A Comparison of Various Parameter Estimation Methods for Mixtures of Von Mises Distributions..... | 62 |
| A Comparison Study of Optimal Bandwidth Selection with Density Functional Estimation for Contaminated Data | 63 |
| A Computer Application for a Decision Making Algorithm Which is Based on the Soft Set Theory..... | 64 |

| | |
|---|----|
| A Dual Response Problem Based on | 65 |
| Median and Median Absolute Deviation | 65 |
| A Load Exchange Mechanism for Less-than-Truckload Operations | 66 |
| A Macro Level Comparison of Hospital Efficiency in Turkey by Using Malmquist Total Factor Productivity Index..... | 67 |
| A Model on the Total Number of Defective Items | 68 |
| A New Approximation for Fuzzy Directional Clustering(FDC) | 69 |
| A New Extended Lindley Distribution..... | 70 |
| A New Method for Measuring Service Quality: An Application in Public Transportation Sector | 71 |
| A New Quantile Penalty Estimation | 72 |
| A Production Inventory Model with Phase-Type Service Time..... | 73 |
| A Queuing-Inventory System with Interruption and Two-Demand Classes | 74 |
| A Shiny App for Subgroup Discovery Using Activity Region Finder (ARF)..... | 75 |
| A Statistical Characterization of a Simulated Earthquake Fault Plane Area by Using Spatial Statistics | 76 |
| A Structural Equation Model for Success Measurement of Development Projects: The Case of ISTKA Projects..... | 77 |
| A Study on the Robust Bootstrap Procedure by Hampel Weighting Function | 78 |
| A Two-stages MCMC Approach to Tree-ring Dating in The Absence of Master Chronologies | 79 |
| Adaptive Principle Components Calibration in Survey Sampling..... | 80 |
| Adjusting of WABL Parameters for Decision Tree Algorithm with Fuzzy Numbers: A Case study.... | 81 |
| Algorithms for Computing the Reliability of r, s -out-of- n Systems..... | 82 |
| Analysis of penalized regression methods in a simple linear model on the high-dimensional data | 83 |
| An Alternative Testing Procedure for The Equality of Two Population Means Using Median Ranked Set Sampling | 84 |
| Application of Extreme Value Theory (EVT) to Inflation Rate Targeting: The Ghanaian case | 85 |
| An Application of Frailty Models in Motor Insurance | 86 |
| An Application of Unrelated Parallel Machine Scheduling with Sequence-Dependent Setups..... | 87 |
| An Efficient Treatment for Best-Worst Method with Multiple Best and/or Worst Criteria | 88 |
| An Investigation on the Efficiency of Research Universities in Turkey | 89 |
| An Optimization Model for Warehouse Load Balancing to Cover High Seasonal Demand in Supply Chain..... | 90 |
| Analysis of PISA-2015 Performance of Turkish Students by Multilevel Structural Equation Modeling | 91 |
| Analyzing Housing Market Dynamics Using Linear and non-Parametric Models | 92 |
| Anomaly Detection in Multivariate Data Using Histogram Based Outlier Score | 93 |
| Artificial Neural Network Approach on Type II Regression Analysis..... | 94 |

| | |
|--|-----|
| Assembling of Structural Equation Model and Fuzzy Analytic Hierarchy Process Methods: An Application for Residence Selection..... | 95 |
| Bayes Estimation of the Reliability Function and Hazard Rate of the Topp–Leone Distribution..... | 96 |
| Bias Types in Genome-Wide Association Studies..... | 97 |
| Bland Altman and Deming Regression Methods from Method Comparison Methods and An Application in Prostate Cancer*..... | 98 |
| Bus Driver Scheduling in Interurban Transportation System..... | 99 |
| Calculating Happiness Index with Fuzzy DSC TOPSIS Method..... | 100 |
| Calculating Value at Risk: Comparing of Different Estimators..... | 101 |
| Calibration of Internal Clustering Validation Indices..... | 102 |
| Cardinality Constrained Mean-Variance-Skewness Model by the Artificial Bee Colony Algorithm: An Application in Istanbul Stock Exchange..... | 103 |
| Cardiovascular Disease Population Risk Tool (CVDPoRT): The Development and Application of a Risk Prediction Model Using Population Health Surveys..... | 104 |
| Characterization Parameters of Domestic Wastewater Samples Collected from Muğla District in Turkey by Decision Tree Method..... | 105 |
| Classification of Author’s Education Degree by Using Three Learning Methods..... | 106 |
| Classification of Data by Using Machine Learning Methods..... | 107 |
| Clustering Active Users and Potential Users of a Mobile Payment Application through Self Organizing Maps..... | 108 |
| Clustering Analysis in Spatio-temporal Data Mining..... | 109 |
| Clustering Bank Customers Using Self-Organizing Maps..... | 110 |
| Combination of Forecasting Methods based on Fuzzy C-Means Clustering..... | 111 |
| Comparing Performances of Artificial Intelligence Techniques based Wavelet Transforms and Feature Selection for Detection of Epileptic Seizures..... | 112 |
| Comparison of Bayes and OLS based Type II Regression..... | 113 |
| Comparison of Classification Performances for Predictiveness Curve and ROC Curve..... | 114 |
| Comparison of Classification Performance of the Galaxies in the Shapley Concentration Region..... | 115 |
| Comparison of Least Squares Estimator with M-Type Regression Estimators..... | 116 |
| Comparison of Optimal Reinsurance Studies on One Side and on Both Sides..... | 117 |
| Comparison of Some Estimators for Parameters of log-Dagum Distribution..... | 118 |
| Comparison of Supervised Disease Diagnosis Methods for Unbalanced Classes: The Case of Diabetes Diagnosis..... | 119 |
| Comparison of the Claims Development Results Using the Chain Ladder and Extended Complementary Loss Ratio Methods for Turkey..... | 120 |
| Comprehensive Meta Analysis..... | 121 |
| Conditional Value-at Risk Modelling on Oil Prices for Emerging Markets..... | 122 |
| Conditional VaR Forecast for Dependent Losses..... | 123 |

| | |
|--|-----|
| Confidence Intervals and Region for Kumaraswamy Distribution Using Wang Pivotal Quantities under Progressive Censoring..... | 124 |
| Conflicts in the Calculation of the Power of a Statistical Test, by Psychology Students. A Semiotic Point of View. | 125 |
| Constructing Discrete Bayesian Networks from Mixed Data in R..... | 126 |
| Continuum Regression and Comparisons of Some Related Estimators..... | 127 |
| Control Charts Based on Com-Poisson Distribution..... | 128 |
| Copula Based Regression Models: An Application..... | 129 |
| Correlation Based Penalty Estimator in Vector Autoregressive Model | 130 |
| Cox Regression Model with Covariates Subject to Measurement Error..... | 131 |
| Creation of a University League Based on the Contribution of their Economics Departments to Economics Education..... | 132 |
| Cryptocurrencies, Currencies and Gold Fractal Volatility Analysis | 133 |
| CT-Visible Doxorubicin-Eluting Clay Particles for Transarterial Catheter-Directed Drug Delivery.. | 134 |
| Customer Satisfaction Survey for Old Age Asylum Using SERVQUEL..... | 135 |
| Detection of Lines on Images with Multimodal Directional Artificial Bee Colony Algorithm | 136 |
| Determining Anomalies of BIST 100 Index by BURR XII Distribution | 137 |
| Diagonal Parameter Symmetry Model Based on Euclidian..... | 138 |
| Distance in Square Contingency Tables..... | 138 |
| Differential Network Analysis with Partial Least Square Regression | 139 |
| Directional k-means Algorithm Using Signed Angular Distance..... | 140 |
| Discriminating between the Lognormal and Weibull Distributions under Progressive Censoring. | 141 |
| Dynamic Measurement of Achieving Europe 2020 Goals in Smart Development – Turkey and European Union Countries | 142 |
| Effect of ionic polymers in the targeted drug delivery applications as the coating materials on superparamagnetic iron oxide nanoparticles | 143 |
| Effect of Sample Size on Learning Performance in Sentiment Analysis..... | 144 |
| Efficiency Analysis of Ankara Rail and Light Rail System Stations..... | 145 |
| Efficiency Evaluation of Ankara Public Transport Bus Lines Using DEA | 146 |
| Efficient Algorithms to Compute Signatures of Binary Systems | 147 |
| Empirical estimator of the Conditional Tail Moment..... | 148 |
| Ensemble Learning Methods as In-silico Model for Prediction of Mutagenicity | 149 |
| Estimation of the Parameters of Two-Fold Weibull Model with Right Censored Reliability Data Using Genetic Algorithm | 150 |
| Estimating First Order Inclusion Probability Based on Percentile Ranked Set Sampling | 151 |
| Estimating the Total Number of Defective Items: A Case Study..... | 152 |
| Estimation of Bone Age from Radiological Images with Soft Computing Methods..... | 153 |

| | |
|---|-----|
| Estimation of Demand for Motor Own Damage Insurance with Neural Networks | 154 |
| Estimation of The Reliability in Multicomponent Stress-Strength Model Based on Chen Lifetime Distribution Using Genetic Algorithm | 155 |
| Estimation of the System Reliability for Generalized Inverse Lindley Distribution under Different Sampling Schemes | 156 |
| Estimation of Turkish Court of Cassation Decisions via Machine Learning and Natural Language Processing Methods | 157 |
| Evaluating The Gage Repeatability and Reproducibility for Industries..... | 158 |
| Evaluation of Epistasis Effect in Genome-Wide Association Studies with Ant Colony Optimization Technique | 159 |
| Examining the Efficiency of Hydroelectric Power Plants in Turkey..... | 160 |
| Examining the Matching Algorithms in Propensity Score Analysis | 161 |
| Experimental Design Applications in Chemistry..... | 162 |
| Extreme Value Theory on Valuation of Actuarial Risks | 163 |
| Factors Affecting the Job Choice of Industrial Engineering Graduates..... | 164 |
| Financial Stress Tests for Measuring Vulnerability of Turkish Economy..... | 165 |
| Forecast Sales With Data Mining In The Public Sector..... | 166 |
| Full Automatic Evaluation of Seismic Refraction Data | 167 |
| Fuzzy ELECTRE I Method for Plant Location Selection: A Case Study | 168 |
| Fuzzy Least Squares Regression Analysis and An Application..... | 169 |
| Fuzzy One-Way ANOVA and An Application | 170 |
| Fuzzy Regression Function Approach Based on Moving Block Bootstrap Method..... | 171 |
| Fuzzy Survival Analysis for Weibull and Gamma Distribution..... | 172 |
| Generalized Linear Model with Hierarchical Likelihood for Sparse Estimation | 173 |
| GIS-Based Approach for Defining Bioenergy Facilities Location: A Case Study in Bismil/Turkey Based on Marginal Delivery Costs and Resources Competition Between Facilities..... | 174 |
| HIV-1 Infection Prognosis: Evolutionary Game Theory Approach | 175 |
| Hierarchical Forecasting | 176 |
| Hybrid Differential Evolutionary Algorithms for Global Optimization Problems..... | 177 |
| Hybrid Model Approach for Feature Selection in Content Analysis..... | 178 |
| Hybrid Type-1 Fuzzy Time Series Functions Approaches: The Application in Stock Exchanges..... | 179 |
| Improved RALS(2)-LM Cointegration Test in Presence of Structural Breaks and GARCH Innovations | 180 |
| Improvement of Vehicles' Production by Creating the Intelligent Information System to Verify Manufacturability of Design Documentation..... | 181 |
| Imputation Strategies for Dealing with Missing Values in Multi-item Scales | 182 |
| Inventory Control Policy Decisions in the Presence of Servers Subject to Random Failure | 183 |

| | |
|---|-----|
| Inversion of Electrical Resistivity Data: An Application from Gümüşhane University Football Field | 184 |
| Investigation of The Attitudies of Academicians About The Individual Pension System: The Case of Eskisehir Osmangazi University | 185 |
| Investigating the Impact of Polysomy 17 in Breast Cancer Patients Whit Amplification through Meta-Analysis | 186 |
| Investigating the Use of Mobile Banking Application by the Technology Acceptance Model (TAM) | 187 |
| Investigation of European Football Clubs Efficiencies with Bi-Objective Multi Criteria Data Envelopment Analysis (BiO-MCDEA) | 188 |
| Investigation of The Effect of The Phase Space Parameters on Prediction Performance in Chaotic Time Series | 189 |
| Likelihood Function for Zero-inflated Models with Fuzzy Counts..... | 190 |
| Liu Pretest and Stein-Type Shrinkage Estimation in Linear Regression Models | 191 |
| Market Segmentation With Clustering Analysis –SSO Sample..... | 192 |
| Markovian Queueing System with Ordered Entry | 193 |
| M-Competitions: ATA Method and Others | 194 |
| Measurement of the Region's Functions on the Example of the Analysis of the Level of Innovation | 195 |
| Measuring Product Sustainability: A Risk-Focused Approach..... | 196 |
| Measuring the Financial Performances of Companies Listed on BIST-Services Index with Data Envelopment Analysis | 197 |
| Mix EWMA-CUSUM Control Chart In the Presence Of Measurement Error..... | 198 |
| Modeling the Financial Attitudes and Behaviors of Y Generation via Structural Equation Modeling | 199 |
| Modelling of Wind Speed Data for Konya | 200 |
| Modeling Nonlinear Relation by Using Running Interval Smoother, LOWESS and Different Quantile Estimators..... | 201 |
| Modifying Clements' Method with Alternative Skewness and Kurtosis Estimators | 202 |
| Monitoring of Blood Parameters of Dogs and Cats Using Control Chart Techniques..... | 203 |
| More Efficient Estimator for Unit Root and Cointegration: RALS Approach..... | 204 |
| Multicomponent Stress-strength Reliability Estimation for A Model Based on Bivariate Generalized Exponential Distribution..... | 205 |
| Multiple Regression Analysis Based on ANN Using with The Different Type of Error Function | 206 |
| Network Meta-Analysis for Decision Making in Medicine | 207 |
| New (r, Q) Inventory Control Models in the Presence of Misplaced items | 208 |
| Nonlinear Neural Network for Portfolio Optimization based on ARIMA..... | 209 |
| On the Efficiency of Proportion Estimators Under Modified Ranked Set Sampling Methods..... | 210 |
| On the General Formula of Renewal Function of Erlang Distribution | 211 |

| | |
|---|-----|
| On the Renormalized Entropy-based Complexity Analysis by BIST100 Data..... | 212 |
| Optimal Bandwidth Selection for a Kernel Density with a Location-Scale Property..... | 213 |
| Optimal Determination of the Parameters of Some Biased Estimators Using Genetic Algorithm . | 214 |
| Optimising the Biasing Parameters of Liu Type Estimator Based on Ridge Regression Using with Genetic and Differential Evolution Algorithm | 215 |
| Outlier Detection and Variable Selection for Count Data | 216 |
| Parameter Estimation for Uniform-Geometric Distribution Based on Censored Sample | 217 |
| Parametric Joint Modelling for Longitudinal and Survival Data..... | 218 |
| Pattern Recognition in Control Charts with Neural Networks | 219 |
| Penalized Logistic Regression for Classification: A Case Study for non-OECD Countries..... | 220 |
| Penalized L_2E Estimator for Gaussian Mixture Model | 221 |
| Performance Comparisons of Some Classification Algorithms Used in Machine Learning | 222 |
| Practicality of Some Variations of Ranked Set Sampling..... | 223 |
| Precursory Seismicity Rate Changes associated with the earthquakes in the East Anatolian Region of Turkey: An Application of Region-Time-Length (RTL) Algorithm | 224 |
| Preparation MRI imageable novel HEC-PVP-Magnetic Embolic Particles with anticancer drug release for Transcatheter Arterial Embolization | 225 |
| Pricing Of Long-Term Care Insurance For Turkey..... | 226 |
| Probability Density Function Estimation Using Scaled Inverse Chi-Squared Kernel | 227 |
| Regression Mixture Model with Two Components for Count Data: An Application to Traffic Accidents | 228 |
| Reliability Analysis of a Multi-state System with Identical Units Having Two Dependent Components | 229 |
| Restricted Bridge Estimator..... | 230 |
| Results of a Project of Marriages Between Disabled People | 231 |
| Risk Evaluation for Support: Predictions for Elder-Life in Community Tool | 232 |
| Robust Estimators in Linear Regression: A Simulation Study and an Application on Healty Care.. | 233 |
| Robust Estimation of Response Surface..... | 234 |
| Robust Estimators in Generalized Linear Models with Binary Responses in the case of Influential Points..... | 235 |
| Robust Estimators in Portfolio Optimization..... | 236 |
| Searching the Differences Through the Tails of Distributions..... | 237 |
| Shiny App to Real Time Tweet Sentiment Classification Using Deep Learning..... | 238 |
| Simulated Annealing Based Simulation Optimization Method for Resource Constrained Project Scheduling Problem with Multiple Modes | 239 |
| Site Effect Classification Estimated from H/V Spectral Ratio of Single Station Microtremor Data in Kürtün-Özkürtün Districts of Gümüşhane, Turkey..... | 240 |
| Spatial Econometric Modeling of Social Capital Effect on Internal Migration | 241 |

| | |
|---|-----|
| Spatio Temporal Bayesian Beta Model to Estimate the Forced Displacement Ratio by Municipalities in Colombia 2002-2016 | 242 |
| Spatio-Temporal Multivariate Imputation of Missing Air Quality Data | 243 |
| Stacking Ensemble Approach for the Classification of Survival Status: HCC patients..... | 244 |
| Statistical Investigation of the Effect of Burnout and Job Satisfaction on Academicians | 245 |
| Statistical Methods in Optimization of Food Materials | 246 |
| Statistical Power in Meta-Analysis Using Correlation Coefficient as Effect Size..... | 247 |
| Statistical Shape Analysis of Hand and Wrist in Children..... | 248 |
| Stock Certificate Selection for Optimal Investment by Using Path Analysis and Multi Criteria Decision Making Methods..... | 249 |
| Study and modeling of the extreme rainfalls (case: Khemis Miliana, Algeria)..... | 250 |
| Temperature interpolation using spatio-temporal Kriging for incomplete data | 251 |
| The Effect of Biodiesel Market and Policies in European Union and Turkey on Economic Growth | 252 |
| The effect of Methods Used for Missing Data Imputation on Classification Success..... | 253 |
| The Effects of Ranking Error Models on Mean Estimators Based on Ranked Set Sampling | 254 |
| The Impact of Logistics on the Economic Series in Turkey..... | 255 |
| The Iterative Method for Multidimensional Objects Ranking..... | 256 |
| The Lindley- Frechet Distribution and Properties with Lifetime Data Application | 257 |
| The Portfolio Optimization based on Genetic Algorithm Approach | 258 |
| The Proposed F Approach for Estimation of Missing Landmark | 259 |
| The Use of Joinpoint Regression Analysis in Core Healthcare Indicators in Turkey, 2003-2016. ... | 260 |
| Time Series Based Decision Support System to Reduce Unexpected Airline Diverts | 261 |
| Time Series Forecasting of Intermittent Demand by Using ATA Method and Computational Intelligence | 262 |
| Turkish Energy Sector Analysis: A Bootstrap Data Envelopment Analysis Approach | 263 |
| Understanding Drought with Copula Functions: Case Study for Konya Province..... | 264 |
| Use of Heijunka Technique to Reduce Production Variability by Leveling Excessive Demand Increase/Decrease..... | 265 |
| Using Multi-Response Taguchi Method (MRSN) for the Achievement of Product Optimization with Six Sigma Methodology in the Automotive Industry..... | 266 |
| What sample sizes is needed in Validation Studies?..... | 267 |
| Workforce Forecasting | 268 |

1. ORGANIZING COMMITTEE

Head of the Organizing Committee

Atilla GÖKTAŞ

Members of the Organizing Committee

Dursun AYDIN

Özge AKKUŞ

Serdar DEMİR

Eralp DOĞU

Özlem GÜRÜNLÜ ALMA

Öznur İŞÇİ GÜNERİ

Nida GÖKÇE NARİN

Nevin GÜLER DİNCER

Mehmet KARAHASAN

Aytaç PEKMEZCİ

Volkan SEVİNÇ

Çağatay BAL

Selman MERMİ

Muhammet Oğuzhan YALÇIN

Ersin YILMAZ

Ö. Ozan EVKAYA

Pınar GÖKTAŞ

2. SCIENTIFIC COMMITTEE

| | |
|------------------------|---|
| Abdullah Fırat Özdemir | Dokuz Eylül University, TURKEY |
| Adil Baykasoğlu | Dokuz Eylül University, TURKEY |
| Ahmet Özmen | Anadolu University, TURKEY |
| Ali Rıza Firuzan | Dokuz Eylül University, TURKEY |
| Algın Okursoy | Adnan Menderes University, TURKEY |
| Ali İhsan Genç | Çukurova University, TURKEY |
| Arzu Altın Yavuz | Osmangazi University, TURKEY |
| Asma Gul | Shaheed Benazir Bhutto Women University, PAKISTAN |
| Atalay Çağlar | Pamukkale University, Turkey |
| Atıf Evren | Yıldız Technical University, TURKEY |
| Aydın Erar | Mimar Sinan University of Fine Arts, TURKEY |
| Aylin Alın | Dokuz Eylül University, TURKEY |
| Ayşen Orhan Apaydın | Ankara University, TURKEY |
| Ayşen Dener Akkaya | Middle East Technical University, TURKEY |
| Ayten Yiğiter | Hacettepe University, TURKEY |
| Barış Sürücü | Middle East Technical University, TURKEY |
| Berna Yazıcı | Anadolu University, TURKEY |
| Bill Woodall | Virginia Tech University, USA |
| Birdal Şenoğlu | Ankara University, TURKEY |
| Birsen Eygi Erdoğan | Marmara University, TURKEY |
| Bülent Altunkaynak | Gazi University, TURKEY |
| Bülent Güloğlu | Istanbul Technical University, TURKEY |
| Canan Hamurkaroğlu | Karabuk University, TURKEY |
| Celal Aydın | Gazi University, TURKEY |
| Cem Kadılar | Hacettepe University, TURKEY |
| Cemil Çolak | İnönü University, TURKEY |
| Ceyhan İnal | Hacettepe University, TURKEY |
| Coşkun Kuş | Selçuk University, TURKEY |
| Çağdaş Hakan Aladağ | Hacettepe University, TURKEY |
| Dariush Akbarian | İslamic Azad University, IRAN |
| Deniz İnan | Marmara University, TURKEY |
| Dilek Altaş | Marmara University, TURKEY |
| Durdu Karasoy | Hacettepe University, TURKEY |
| Ebru Çağlayan Akay | Marmara University, TURKEY |
| Ekaterina Chimitova | Novosibirsk State Technical University, RUSSIA |
| Eleanor Stillman | Sheffield University, UK |
| Emel Çankaya | Sinop University, TURKEY |
| Engin Yıldıztepe | Dokuz Eylül University, TURKEY |
| Erhard Cramer | Aachen University, GERMANY |
| Erol Eğrioğlu | Giresun University, TURKEY |
| Fatih Çemrek | Osman Gazi University, TURKEY |
| Feridun Taşdan | Western Illinois University, USA |
| Fikri Akdeniz | Çağ University, TURKEY |
| Fikri Gökpınar | Gazi University, TURKEY |
| Fotios Milienos | Ioannina University, GREECE |

| | |
|----------------------------|---|
| Gamze Özel Kadılar | Hacettepe University, TURKEY |
| Gökhan Silahtaroğlu | Medipol University, TURKEY |
| Güçkan Yapar | Dokuz Eylül University, TURKEY |
| Gül Ergün | Hacettepe University, TURKEY |
| Gülay Başarır | Mimar Sinan University of Fine Arts, TURKEY |
| Gülhayat Gölbaşı Şimşek | Yıldız Technical University, TURKEY |
| Güneş Çetin Gerger | Celal Bayar University, TURKEY |
| Güzin Yüksel | Çukurova University, TURKEY |
| Habib Shah | King Khalid University, SAUDI ARABIA |
| Hadi Alizadeh Noughabi | Birjand University, IRAN |
| Hamparsum Bozdoğan | Tennessee University, USA |
| Hamza Gamgam | Gazi University, TURKEY |
| Hasan Bal | Gazi University, TURKEY |
| Hasan Örkcü | Gazi University, TURKEY |
| Hatice Hicret Özkoç | Muğla Sıtkı Koçman University, TURKEY |
| Heather Turner | Warwick University, UK |
| Hülya Bayrak | Gazi University, TURKEY |
| Hülya Çingı | Hacettepe University, TURKEY |
| Hüseyin Tatlıdil | Hacettepe University, TURKEY |
| Hon Keung Tony Ng | Southern Methodist University, USA |
| İbrahim Demir | Yıldız Technical University, TURKEY |
| İhsan Alp | Gazi University, TURKEY |
| İlhan Usta | Anadolu University, TURKEY |
| İlker Ercan | Uludag University, TURKEY |
| İnci Batmaz | Middle East Technical University, TURKEY |
| İsmihan Bayramoğlu | İzmir Economy University, TURKEY |
| Kurtuluş Bozkurt | Adnan Menderes University, TURKEY |
| Lara Lusa | University of Ljubljana, SLOVENIA |
| M. Akif Bakır | Gazi University, TURKEY |
| M. Ali Cengiz | Samsun 19 Mayıs University, TURKEY |
| Meral Çetin | Hacettepe University, TURKEY |
| Meral Ebegil | Gazi University, TURKEY |
| Minoo Aminnejad | Razi University, IRAN |
| Miruna Mazurencu Marinescu | Bucharest University of Economic Studies, ROMANIA |
| Mohd Saifullah Rusiman | Tun Hussein University on Malaysia, MALAYSIA |
| Müjgan Tez | Marmara University, TURKEY |
| Mükerrem Bahar Başkır | Bartın University, TURKEY |
| Narayanaswamy Balakrishnan | McMaster University, CANADA |
| Nazan Danacıoğlu | Sinop University, TURKEY |
| Necla Gündüz Tekin | Gazi University, TURKEY |
| Nihal Ata Tutkun | Hacettepe University, TURKEY |
| Oğuz Akbilgiç | Tennessee University, USA |
| Olcay Aslan | Ankara University, TURKEY |
| Onur Köksoy | Ege University, TURKEY |
| Ozan Kocadağlı | Mimar Sinan University of Fine Arts, TURKEY |
| Özgür Yeniay | Hacettepe University, TURKEY |
| Özlem Türkşen | Ankara University, TURKEY |
| Philippe Castagliola | Université de Nantes, FRANCE |

Pınar Göktaş
Rama Shanker
Reza Arabi Beloghi
Rezzan Uslu
S. Ezaj Ahmed
Salvatore Ingrassia
Selahattin Kaçıranlar
Semra Erbaş
Semra Türkan
Serdar Kurt
Serkan Eryılmaz
Serpil Aktaş Altunay
Serpil Cula
Sevil Şentürk
Sukhdev Singh
Suvra Pal
Süleyman Günay
Sylwia Gwozdziwicz
Şahamet Bülbül
Şenol Erdoğan
Taner Tunç
Türkan Erbay Dalkılıç
Ufuk Yolcu
Ummahan Gökovalı
Ülkü Gürler
Vedat Sağlam
Veli Yılancı
Veysel Yılmaz
Wali Khan Mashwani
Yaprak Arzu Özdemir
Yasemin Gençtürk
Yasemin Kayhan Atılğan
Yeliz Mert Kantar
Yüksel Akay Ünvan
Yılmaz Akdi
Yüksel Terzi
Zehra Muluk
Zeki Yıldız
Zeynep Filiz

Muğla Sıtkı Koçman University, TURKEY
Eritrea Institute of Technology University, AFRICA
Tabriz University, IRAN
Samsun 19 Mayıs University, TURKEY
Brock University, CANADA
University of Catania, ITALY
Çukurova University, TURKEY
Gazi University, TURKEY
Hacettepe University, TURKEY
Dokuz Eylül University, TURKEY
Atılım University, TURKEY
Hacettepe University, TURKEY
Başkent University, TURKEY
Anadolu University, TURKEY
Indian Institute of Technology Patna, INDIA
Texas University, USA
Hacettepe University, TURKEY
Jacob of Paradies University, POLAND
Marmara University, TURKEY
Osmangazi University, TURKEY
Ondokuz Mayıs University, TURKEY
Karadeniz Technical University, TURKEY
Giresun University, TURKEY
Muğla Sıtkı Koçman University, TURKEY
Bilkent University, TURKEY
Ondokuz Mayıs University, TURKEY
Sakarya University, TURKEY
Osmangazi University, TURKEY
KOHAT University, PAKISTAN
Gazi University, TURKEY
Hacettepe University, TURKEY
Hacettepe University, TURKEY
Anadolu University, TURKEY
Yıldırım Beyazıt University, TURKEY
Ankara University, TURKEY
Ondokuz Mayıs University, TURKEY
Başkent University, TURKEY
Osmangazi University, TURKEY
Osmangazi University, TURKEY

3. LIST OF PRESENTER PARTICIPANTS

Adem Akgül, TURKEY
Adem Dođaner, TURKEY
Adil Baykasoglu, TURKEY
Adnan Karaibrahimođlu, TURKEY
Agata Mesjasz-Lech, POLAND
Ahmet Kara, TURKEY
Ahmet Toy, TURKEY
Alev Yüksel Aydar, TURKEY
Ali Ekici, TURKEY
Ali Rıza Firuzan, TURKEY
Ali Zafer Dalar, TURKEY
Andrés Felipe Rodriguez Perez, COLOMBIA
Andrzej Sokołowski, POLAND
Arzu BAYGÜL, TURKEY
Asiye Zühal Küçükmustafa, TURKEY
Aslı Günay, TURKEY
Asli Akilli, TURKEY
Atakan GERGER, GERMANY
Atalay Çađlar, TURKEY
Atif Evren, TURKEY
Ayça Ölmez, TURKEY
Ayda Amniattalab, TURKEY
Aylin Göçođlu, TURKEY
Ayşe İři, TURKEY
Ayşegül Yabancı TURKEY
Ayten Yiđiter, TURKEY
Azize Zehra Çelenli Başaran, TURKEY
Bahadır Yüzbaşı, TURKEY
Baki Unal, TURKEY
Berhan Çoban, TURKEY
Berihan R. Elemary, EGYPT
Berkalp Tunca, TURKEY
Berrin Gültay, TURKEY
Betül Zehra Karagül, TURKEY
Beyza Kızılkaya, TURKEY
Bilgi Yılmaz, TURKEY
Bulent Oktay Akkoyunlu, TURKEY
Burak Dilber, TURKEY
Burcu Aytaçođlu, TURKEY
Burcu Durmuş, TURKEY
Burçin Öner, TURKEY
Buse Dedeakayođulları, TURKEY
Bükre Yıldırım Külekci, TURKEY
Bünyamin Sarıbacak, TURKEY
Caner Tanıř, TURKEY

Ceren Dirik, TURKEY
Ceyhun Mazlum, TURKEY
Coşkun Kuş, TURKEY
Çağdaş Hakan Aladağ, TURKEY
Çiğdem Lazoğlu, TURKEY
Damla Beril Özyazgan, TURKEY
Daniela López Sarmiento, COLOMBIA
Demet Sezer, TURKEY
Derviş Topuz, TURKEY
Derya Alkın, TURKEY
Devrim Bilgili, USA
Didem Derici Yıldırım, TURKEY
Dilek Veysikarani, TURKEY
Duru Sertkaya Karasoy, TURKEY
Duygu Demiray, TURKEY
Ebru Kaya Başar, TURKEY
Ebru Özgür Güler, TURKEY
Ece Özgören, TURKEY
Ece Uzunsakal, TURKEY
Eleanor Stillman, UNITED KINGDOM
Elif Çoker, TURKEY
Elif Dil, TURKEY
Elif Erçelik, TURKEY
Engin Sungur, USA
Engin Taş, TURKEY
Enis Gümüştas, TURKEY
Eren Bas, TURKEY
Erhan Ustaoglu, TURKEY
Erkut Tekeli, TURKEY
Erol Eğrioğlu, TURKEY
Esin Avcı, TURKEY
Esin Firuzan, TURKEY
Esra Betül Kınacı, TURKEY
Esra Satıcı, TURKEY
Fatih Çemrek, TURKEY
Fatma Ezgi Can, TURKEY
Fatma Sevinç Kurnaz, TURKEY
Femin Yalcin, TURKEY
Fikri Akdeniz, TURKEY
Funda Iscioglu, TURKEY
G.Yazgı Tütüncü, TURKEY
Gizem Erol, TURKEY
Godrick Oketch, TURKEY
Gökhan Aydın, TURKEY
Gökhan Gökdere, TURKEY
Gökhan Silahtaroglu, TURKEY
Gözde Navruz, TURKEY

Gözde Özçırpan, TURKEY
Gözde Ulu Metin, TURKEY
Gözde Yaylalı Umul, TURKEY
Guermah Toufik Rassoul Abdelaziz, ALGERIA
Güçkan Yapar, TURKEY
Güenal Bilek, TURKEY
Hakan Savaş Sazak, TURKEY
Hamparsum Bozdoğan, USA
Hande Konşuk Ünlü, TURKEY
Harun Kınacı, TURKEY
Havva Didem Celikcan, TURKEY
Hayriye Esra Akyüz, TURKEY
Hazal Berve Doğan, TURKEY
Hüseyin Güler, TURKEY
Ilgım Yaman, TURKEY
İlkay Altındağ, TURKEY
İlker Ercan, TURKEY
İlker Türker, TURKEY
İrem Yanık, TURKEY
İsmail Kınacı, TURKEY
Kadir Karakaya, TURKEY
Ksenia Shubenkova, RUSSIA
Kumru Urgancı Tekin, TURKEY
Liaquat Ahmad, PAKISTAN
Mahmut Sarı, TURKEY
Maide Gökçe Bekaroğlu, TURKEY
Małgorzata Markowska, POLAND
Marek Szajt, POLAND
Mária Bohdalová, SLOVAKIA
Marina Fađoš, SLOVAKIA
Masoud Muhammed Hassan, IRAQ
Mehmet Arif Demirci, TURKEY
Mehmet Gülşen, TURKEY
Mehmet Nasih Tağ, TURKEY
Melih Ağraz, TURKEY
Melis Zeybek, TURKEY
Meltem Gülsüm Tuna, CANADA
Meltem Yontar, TURKEY
Merve Akdede, TURKEY
Merve Türkegün, TURKEY
Metehan Sönmez, TURKEY
Metin Yangın, TURKEY
Monjed H. Samuh, PALESTINE
Muhammad Noor-ul-amin, PAKISTAN
Muhammet Burak Kılıç, TURKEY
Murat Gül, TURKEY
Mustafa Çağatay Korkmaz, TURKEY

Mustafa Nadar, TURKEY
Müge Yeldan, TURKEY
Mürüvvet Deniz Sezer, TURKEY
Narayanaswamy Balakrishnan, CANADA
Natalia Niño Martínez, COLOMBIA
Nazlı Deniz Ergüç, TURKEY
Necla Gündüz, TURKEY
Neslihan Avcu, TURKEY
Neslihan Gökmen, TURKEY
Nesrin Güler, TURKEY
Nida Gökçe Narin, TURKEY
Nihat Tak, TURKEY
Nimet Yapıcı Pehlivan, TURKEY
Nur Uylaş Satı, TURKEY
Nurhayat Şahinkaya, TURKEY
Nuri Celik, TURKEY
O. Ozan Evkaya, TURKEY
Oğuz Akpolat, TURKEY
Okan Örsan Özener, TURKEY
Orhan Kesemen, TURKEY
Oya Kalaycıoğlu, TURKEY
Ozan Kocadağlı, TURKEY
Ozge Cagcag Yolcu, TURKEY
Ozlem Sevinc (Ardic), TURKEY
Ömer Akgüller, TURKEY
Özlem Akay, TURKEY
Özlem GÖCEN, TURKEY
Özlem Gürünlü Alma, TURKEY
Özlem Kaymaz, TURKEY
Özlem Türkşen, TURKEY
Pakize Yiğit, TURKEY
Rassoul Abdelaziz, ALGERIA
Richard Minkah, SUDAN
Risti Nur'aini, INDONESIA
Rümeysa Karataş, TURKEY
Sami AKDENİZ, TURKEY
Saurabh Kumar, INDIA
Sebnem Er, SOUTH AFRICA
Selahattin Kaçıranlar, TURKEY
Selen Çakmakyapan, TURKEY
Selman Mermi, TURKEY
Semra Erbaş, TURKEY
Semra Erpolat Taşabat, TURKEY
Senem Şahan Vahaplar, TURKEY
Serhat Emre Akhanlı, UNITED KINGDOM
Serkan Eryılmaz, TURKEY
Serkan Öztürk, TURKEY

Serpil Aktaş Altunay, TURKEY
Serpil Kılıç Depren, TURKEY
Shehzad Muhammad Ahmed, PAKISTAN
Sinan Saraçlı, TURKEY
Songül Bulut, TURKEY
Sultan Turhan, TURKEY
Suzan Kantarci Savaş, TURKEY
Şerife Özkar, TURKEY
Şeyma Yaşar, TURKEY
Talha Arslan, TURKEY
Talip Arsu, TURKEY
Tolga Yamut, TURKEY
Tuba Adar, TURKEY
Tufan Baydemir, TURKEY
Tuğçe Ekiz Yılmaz, TURKEY
Tuğçe Leventoğlu, TURKEY
Turgut Özaltındış, TURKEY
Ufuk Yolcu, TURKEY
Uğur Binzat, TURKEY
Uğur Karabey, TURKEY
Umay Uzunoğlu Koçer, TURKEY
Ülkü Gürler, TURKEY
Ümit Kuvvetli, TURKEY
Vildan Yıldırım, TURKEY
Volkan Alptekin, TURKEY
Volkan Yeşildere, TURKEY
Wali Khan Mashwani, PAKISTAN
William H. Woodal, USA
Yaprak Arzu Özdemir, TURKEY
Yonca Yazırlı, TURKEY
Yunus Akdoğan, TURKEY
Yusuf Can Sevil, TURKEY
Yüksel Akay Ünvan, TURKEY
Zeynep Tunç, TURKEY

4. LIST OF ALL PARTICIPANTS

| | | | |
|---------------|------------------|--------------------------------------|----------------|
| ADEM | AKGÜL | GAZIANTEP UNIVERSITY | TURKEY |
| ADEM | DOĞANER | KAHRAMANMARAŞ SUTÇU IMAM UNIVERSITY | TURKEY |
| ADIL | BAYKASOĞLU | DOKUZ EYLUL UNIVERSITY | TURKEY |
| ADNAN | KARAIBRAHIMOĞLU | SULEYMAN DEMIREL UNIVERSITY | TURKEY |
| A.FIRAT | ÖZDEMİR | DOKUZ EYLUL UNIVERSITY | TURKEY |
| AGATA | MESJASZ-LECH | CZESTOCHOWA UNIVERSITY OF TECHNOLOGY | CZECH REPUBLIC |
| AHMET | KARA | ISTANBUL COMMERCE UNIVERSITY | TURKEY |
| AHMET | TOY | ONDOKUZ MAYIS UNIVERSITY | TURKEY |
| ALEV YÜKSEL | AYDAR | MANISA CELAL BAYAR UNIVERSITY | TURKEY |
| ALI | EKICI | OZYEGİN UNIVERSITY | TURKEY |
| ALI RIZA | FIRUZAN | DOKUZ EYLUL UNIVERSITY | TURKEY |
| ALI ZAFER | DALAR | GIRESUN UNIVERSITY | TURKEY |
| ANDRÉS FELIPE | RODRIGUEZ PEREZ | UNIVERSIDAD SANTO TOMAS | COLOMBIA |
| ANDRZEJ | SOKOLOWSKI | CRACOW UNIVERSITY OF ECONOMICS | POLAND |
| ARZU | BAYGÜL | BEYKENT UNIVERSITY | TURKEY |
| ASIYE ZUHAL | KÜÇÜKMUSTAFA | THE UNIVERSITY OF AVRASYA | TURKEY |
| ASLI | GÜNAY | SOCIAL SCIENCES UNIVERSITY OF ANKARA | TURKEY |
| ASLI | AKILLI | AHI EVRAN UNIVERSITY | TURKEY |
| ATAKAN | GERGER | PFW AEROSPACE GMBH | TURKEY |
| ATALAY | ÇAĞLAR | PAMUKKALE UNIVERSITY | TURKEY |
| ATIF | EVREN | YILDIZ TEKNİK UNIVERSITY | TURKEY |
| ATILA | GÖKTAŞ | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| AYÇA | ÖLMEZ | DOKUZ EYLUL UNIVERSITY | TURKEY |
| AYDA | AMNIATTALAB | SABANCI UNIVERSITY | TURKEY |
| AYDIN | KARAKOCA | NECMETTİN ERBAKAN UNIVERSITY | TURKEY |
| AYLIN | GÖÇOĞLU | DOKUZ EYLUL UNIVERSITY | TURKEY |
| AYŞE | IŞI | GAZI UNIVERSITY | TURKEY |
| AYŞEGÜL | YABACI | BEZMIALEM VAKIF UNIVERSITY | TURKEY |
| AYTAÇ | PEKMEZCI | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| AYTEN | YIĞITER | HACETTEPE UNIVERSITY | TURKEY |
| AZIZE ZEHRA | ÇELENLİ | ONDOKUZ MAYIS UNIVERSITY | TURKEY |
| BAHADIR | YÜZBAŞI | INONU UNIVERSITY | TURKEY |
| BAKI | ÜNAL | HACETTEPE UNIVERSITY | TURKEY |
| BERHAN | ÇOBAN | DOKUZ EYLUL UNIVERSITY | TURKEY |
| BERİHAN R. | ELEMARY | DAMIETTA UNIVERSITY | EGYPT |
| BERKALP | TUNCA | AFYON KOCATEPE UNIVERSITY | TURKEY |
| BERRİN | GÜLTAY | ÇANAKKALE ONSEKİZ MART UNIVERSITY | TURKEY |
| BETÜL ZEHRA | KARAGÜL | HACETTEPE UNIVERSITY | TURKEY |
| BEYZA | KIZILKAYA | DOKUZ EYLUL UNIVERSITY | TURKEY |
| BILGI | YILMAZ | MIDDLE EAST TECHINACAL UNIVERSITY | TURKEY |
| BUĞRA KAAAN | TIRYAKI | KARADENİZ TECHNICAL UNIVERSITY | TURKEY |
| BURAK | DILBER | DOKUZ EYLUL UNIVERSITY | TURKEY |
| BURCU | AYTAÇOĞLU | EGE UNIVERSITY | TURKEY |
| BURÇİN | ÖNER | GAZI UNIVERSITY | TURKEY |
| BUSE | DEDEAKAYOĞULLARI | OBASE BILGISAYAR DANIŞMANLIK | TURKEY |
| BÜKRE | YILDIRIM KÜLEKÇİ | MIDDLE EAST TECHINACAL UNIVERSITY | TURKEY |

| | | | |
|---------------|------------------|---|---------------|
| BÜLENT OKTAY | AKKOYUNLU | MARMARA UNIVERSITY | TURKEY |
| CANER | TANIŞ | SELÇUK UNIVERSITY | TURKEY |
| CEM | KOÇAK | HITIT UNIVERSITY | TURKEY |
| CEREN | DIRIK | KIRIKKALE UNIVERSITY | TURKEY |
| CEYHAN | INAL | HACETTEPE UNIVERSITY | TURKEY |
| CEYHUN | MAZLUM | TRABZON PUBLIC HEALTH STATE | TURKEY |
| CHAIRUN NISAH | ARDIANTARI | UNIVERSITY OF INDONESIA | INDONESIA |
| COŞKUN | KUŞ | SELÇUK UNIVERSITY | TURKEY |
| ÇAĞATAY | BAL | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| ÇAĞDAŞ HAKAN | ALADAĞ | HACETTEPE UNIVERSITY | TURKEY |
| ÇİĞDEM | LAZOĞLU | HACETTEPE UNIVERSITY | TURKEY |
| DAMLA BERİL | ÖZYAGAN | APLUS ENERGY INVESTMENT CONSULTING | TURKEY |
| DANIELA | LOPEZ SARMIENTO | UNIVERSIDAD SANTO TOMAS | COLOMBIA |
| DEMET | SEZER | SELÇUK UNIVERSITY | TURKEY |
| DERVIŞ | TOPUZ | ÖMER HALISDEMİR UNIVERSITY | TURKEY |
| DERYA | ALKIN | NECMETTİN ERBAKAN UNIVERSITY | TURKEY |
| DERYA | GÖZEN | STM | TURKEY |
| DESSIE | WANDA | UNIVERSITY OF INDONESIA | INDONESIA |
| DEVİRİM | BİLGİLİ | UNIVERSITY OF NORTH FLORIDA | UNITED STATES |
| DIDEM | DERİCİ YILDIRIM | MERSİN UNIVERSITY | TURKEY |
| DİLEK | VEYSİKARANI | MUNZUR UNIVERSITY | TURKEY |
| DOĞAN | YILDIZ | YILDIZ TEKNİK UNIVERSITY | TURKEY |
| DURŞUN | AYDIN | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| DURU | SERTKAYA KARASOY | HACETTEPE UNIVERSITY | TURKEY |
| DUYGU | DEMİRAY | YEDİTEPE UNIVERSITY | TURKEY |
| EBRU | ÖZGÜR GÜLER | ÇUKUROVA UNIVERSITY | TURKEY |
| EBRU | YÜKSEL HALILOĞLU | TOBB UNIVERSITY OF ECONOMICS AND TECHNOLOGY | TURKEY |
| EBRU KAYA | BAŞAR | AKDENİZ UNIVERSITY | TURKEY |
| ECE | ÖZGÖREN | AFYON KOCATEPE UNIVERSITY | TURKEY |
| EDA | ÖZKUL | KARADENİZ TECHNICAL UNIVERSITY | TURKEY |
| ELANOR | STILLMAN | UNIVERSITY OF SHEFFIELD | TURKEY |
| ELİF | ÇOKER | MİMAR SİNAN FINE ARTS UNIVERSITY | TURKEY |
| ELİF | ERÇELİK | İSTANBUL TECHNICAL UNIVERSITY | TURKEY |
| ELİF | DİL | HACETTEPE UNIVERSITY | TURKEY |
| ENGİN | TAŞ | AFYON KOCATEPE UNIVERSITY | TURKEY |
| ENGİN | YILDIZTEPE | DOKUZ EYLUL UNIVERSITY | TURKEY |
| ENGİN | SUNGUR | UNIVERSITY OF MINNESOTA MORRIS | UNITED STATES |
| ENİS | GÜMÜŞTAŞ | MİMAR SİNAN GÜZEL SANATLAR UNIVERSITY | TURKEY |
| ERALP | DOĞU | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| EREN | BAŞ | GİRESUN UNIVERSITY | TURKEY |
| ERHAN | USTAOĞLU | MARMARA UNIVERSITY | TURKEY |
| ERKİN | SARIBACAK | ONDOKUZ MAYIS UNIVERSITY | TURKEY |
| ERKUT | TEKELİ | ÇUKUROVA UNIVERSITY | TURKEY |
| EROL | EĞRİOĞLU | GİRESUN UNIVERSITY | TURKEY |
| ESİN | AVCI | GİRESUN UNIVERSITY | TURKEY |
| ESİN | FIRUZAN | DOKUZ EYLUL UNIVERSITY | TURKEY |
| ESRA BETÜL | KINACI | GAZİ UNIVERSITY | TURKEY |
| ESRA | SATICI | KARAYOLLARI GENEL MÜDÜRLÜĞÜ | TURKEY |

| | | | |
|--------------------|-----------------|--|-----------|
| EZGI | ÇABUK | GAZI UNIVERSITY | TURKEY |
| FATİH | ÇEMREK | OSMAN GAZI UNIVERSITY | TURKEY |
| FATMA EZGI | CAN | ULUDAĞ UNIVERSITY | TURKEY |
| FATMA GÜL | AKGÜL | ARTVIN ÇORUH UNIVERSITY | TURKEY |
| FATMA SEVINÇ | KURNAZ | YILDIZ TECHNICAL UNIVERSITY | TURKEY |
| FEMİN | YALÇIN | İZMİR KATİP CELEBİ UNIVERSITY | TURKEY |
| FIKRI | AKDENİZ | ÇAĞ UNIVERSITY | TURKEY |
| FUNDA | İŞÇİOĞLU | EGE UNIVERSITY | TURKEY |
| GİZEM | EROL | STM | TURKEY |
| GODRICK | MARADONA OKETCH | YILDIZ TECHNICAL UNIVERSITY | TURKEY |
| GÖKHAN | AYDIN | İSTANBUL MEDİPOL ÜNİVERSİTESİ | TURKEY |
| GÖKHAN | GÖKDERE | FIRAT UNIVERSITY | TURKEY |
| GÖKHAN | SILAHTAROĞLU | İSTANBUL MEDİPOL ÜNİVERSİTESİ | TURKEY |
| GÖZDE | NAVRUZ | DOKUZ EYLUL UNIVERSITY | TURKEY |
| GÖZDE | ÖZÇIRPAN | ANKARA UNIVERSITY | TURKEY |
| GÖZDE | ULU METİN | MINISTRY OF INDUSTRY AND TECHNOLOGY | TURKEY |
| GÖZDE YAZGI | TÜTÜNCÜ | İZMİR UNIVERSITY OF ECONOMICS | TURKEY |
| GUERMAH | TOUFIK | UNIVERSITE HASSIBA BENBOUALI DE CHLEF | ALGERIA |
| GÜÇKAN | YAPAR | DOKUZ EYLUL UNIVERSITY | TURKEY |
| GÜLAY | BAŞARIR | MIMAR SINAN GÜZEL SANATLAR UNIVERSITY | TURKEY |
| GÜNAL | BİLEK | BITLİS EREN UNIVERSITY | TURKEY |
| HAKAN SAVAŞ | SAZAK | EGE UNIVERSITY | TURKEY |
| HAMPARSUM | BOZDOĞAN | UNIVERSITY OF TENNESSEE | TURKEY |
| HANDE | KONŞUK ÜNLÜ | HACETTEPE UNIVERSITY | TURKEY |
| HANIFI OKAN | İŞGÜDER | DOKUZ EYLUL UNIVERSITY | TURKEY |
| HANS | FRENK | SABANCI UNIVERSITY | TURKEY |
| HARUN | KINACI | ERCIYES UNIVERSITY | TURKEY |
| HATİCE | YENİAY | HACETTEPE UNIVERSITY | TURKEY |
| HAVVA DİDEM | OVLA ÇELİKCAN | MERSİN UNIVERSITY | TURKEY |
| HAYRİYE ESRA | AKYÜZ | BITLİS EREN UNIVERSITY | TURKEY |
| HAZEL BERVE | DOĞAN | STM | TURKEY |
| HÜSEYİN | GÜLER | ÇUKUROVA UNIVERSITY | TURKEY |
| İLGİM | YAMAN | GİRESUN UNIVERSITY | TURKEY |
| İLKAY | ALTINDAĞ | SELÇUK UNIVERSITY | TURKEY |
| İLKER | ERCAN | ULUDAĞ UNIVERSITY | TURKEY |
| İLKER | TÜRKER | KARABUK UNIVERSITY | TURKEY |
| İREM | KARABEBE | HACETTEPE UNIVERSITY | TURKEY |
| İREM | YANIK | HACETTEPE UNIVERSITY | TURKEY |
| İRİNA | MAKAROVA | KAZAN FEDERAL UNIVERSITY | RUSSIA |
| İSMAIL | KINACI | SELÇUK UNIVERSITY | TURKEY |
| KADIR | KARAKAYA | SELÇUK UNIVERSITY | TURKEY |
| KHRESHNA | SYUHADA | INSTITUTE TEKNOLOGI BANDUNG | INDONESIA |
| KSENİA | SHUBENKOVA | KAZAN FEDERAL UNIVERSITY | RUSSIA |
| KUMRU | URGANCI TEKİN | ÇANAKKALE ONSEKİZ MART UNIVERSITY | TURKEY |
| LİAQUAT | AHMAD | UNIVERSITY OF VETERINARY AND ANIMAL SCIENCES | PAKISTAN |
| LİSTİARİNİ JANVARI | MUHTARI | INSTITUTE TEKNOLOGI BANDUNG | INDONESIA |
| MAHMUT | SARI | GUMUŞHANE UNIVERSITY | TURKEY |
| MAİDE GÖKÇE | BEKAROĞLU | İSTANBUL TECHNICAL UNIVERSITY | TURKEY |

| | | | |
|---------------------|-----------------|---------------------------------------|----------------|
| MALGORZATA | MARKOWSKA | WROCLAW UNIVERSITY OF ECONOMICS | POLAND |
| MAREK | SZAJT | CZESTOCHOWA UNIVERSITY OF TECHNOLOGY | CZECH REPUBLIC |
| MARIA | BOHDALOVA | COMENIUS UNIVERSITY | SLOVAKIA |
| MARINA | FADOS | COMENIUS UNIVERSITY | SLOVAKIA |
| MARTIN | TEJKAL | CZECH UNIVERSITY | CZECH REPUBLIC |
| MASOUD MUHAMMED | HASSAN | ZAKHO UNIVERSITY | IRAG |
| MEHMET | GÜLŞEN | BAŞKENT UNIVERSITY | TURKEY |
| MEHMET | KARAHASAN | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| MEHMET NASİH | TAĞ | MERSIN UNIVERSITY | TURKEY |
| MEHMET ARIF | DEMIRCI | UNIVERSITY OF MUŞ ALPASLAN | TURKEY |
| MEHMET AYDIN | ERAR | MIMAR SINAN GUZEL SANATLAR UNIVERSITY | TURKEY |
| MELİH | AĞRAZ | MIDDLE EAST TECHINACAL UNIVERSITY | TURKEY |
| MELİS | ZEYBEK | EGE UNIVERSITY | TURKEY |
| MELTEM | YONTAR | ISTANBUL TECHNICAL UNIVERSITY | TURKEY |
| MELTEM GÜLSÜM | TUNA | OTTAWA HOSPITAL RESEARCH INSTITUTE | CANADA |
| MERAL | EBEGİL | GAZI UNIVERSITY | TURKEY |
| MERVE | AKDEDE | UŞAK UNIVERSITY | TURKEY |
| MERVE | TÜRKEGÜN | MERSIN UNIVERSITY | TURKEY |
| MESERET MOLLA | KASSAHUN | SUSSEX UNIVERSITY | UNITED KINGDOM |
| MESUDE | ARABACIOĞLU | BOTAŞ | TURKEY |
| METEHAN | SÖNMEZ | YILDIZ TEKNİK UNIVERSITY | TURKEY |
| METİN | YANGIN | MIMAR SINAN GUZEL SANATLAR UNIVERSITY | TURKEY |
| MİCHAL | GREGUS | COMENIUS UNIVERSITY | SLOVAKIA |
| MONJED | SAMUH | PALESTINE POLYTECHNIC UNIVERSITY | SAUDI ARABIA |
| MUHAMMAD | NOOR-UL-AMIN | COMSTATS UNIVERSITY | PAKISTAN |
| MUHAMMET BURAK | KILIÇ | MEHMET AKİF ERSOY UNIVERSITY | TURKEY |
| MUHAMMET OĞUZHAN | YALÇIN | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| MURAT | AKŞİT | AFYON KOCATEPE UNIVERSITY | TURKEY |
| MURAT | GÜL | GİRESUN UNIVERSITY | TURKEY |
| MUSTAFA | NADAR | ISTANBUL TECHNICAL UNIVERSITY | TURKEY |
| MUSTAFA ÇAĞATAY | KORKMAZ | ARTVIN ÇORUH UNIVERSITY | TURKEY |
| MÜGE | YELDAN | HACETTEPE UNIVERSITY | TURKEY |
| MÜRÜVVET DENİZ | SEZER | DOKUZ EYLUL UNIVERSITY | TURKEY |
| NARAYANASWAMY | BALAKRISHNAN | MCMASTER UNIVERSITY | CANADA |
| NATALIA NIÑO | MARTÍNEZ | UNIVERSIDAD SANTO TOMAS | COLOMBIA |
| NECLA | GÜNDÜZ | GAZI UNIVERSITY | TURKEY |
| NESLIHAN | AVCU | DOKUZ EYLUL UNIVERSITY | TURKEY |
| NESLIHAN | GÖKMEN | ISTANBUL TECHNICAL UNIVERSITY | TURKEY |
| NESRİN | GÜLER | SAKARYA UNIVERSITY | TURKEY |
| NEVIN | GULER DINCER | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| NIDA | GÖKÇE NARIN | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| NIHAL | ATA TUTKUN | HACETTEPE UNIVERSITY | TURKEY |
| NIHAT | TAK | KIRKLARELI UNIVERSITY | TURKEY |
| NİMET | YAPICI PEHLIVAN | SELCUK UNIVERSITY | TURKEY |
| NIYAZI ONUR | BAKIR | ALTINBAŞ UNIVERSITY | TURKEY |
| NUR | UYLAŞ | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| NURAN | ÇAKIR YILDIZ | YILDIZ TEKNİK UNIVERSITY | TURKEY |

| | | | |
|-------------------------|-----------------|---------------------------------------|----------------|
| NURHAYAT | ŞAHINKAYA | DOKUZ EYLUL UNIVERSITY | TURKEY |
| NURI | ÇELİK | GEBZE TEKNİK UNIVERSITY | TURKEY |
| OĞUZ | AKPOLAT | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| OKAN ORSAN | ÖZENER | OZYEGIN UNIVERSITY | TURKEY |
| ORHAN | KESEMEN | KARADENİZ TECHNICAL UNIVERSITY | TURKEY |
| ORHAN MAZLUM | YAZAR | RECEP TAYYIP ERDOĞAN UNIVERSITY | TURKEY |
| OSMAR DARIO | VERA | NATIONAL UNIVERSITY OF QUILMES | ARGENTINA |
| OYA | KALAYCIOĞLU | ABANT İZZET BAYSAL UNIVERSITY | TURKEY |
| OZAN | EVKAYA | MIDDLE EAST TECHNICAL UNIVERSITY | TURKEY |
| OZAN | KOCADAĞLI | MİMAR SİNAN UNIVERSITY | TURKEY |
| ÖMER | AKGÜLLER | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| ÖZGE | AKKUŞ | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| ÖZGE | CAĞCAĞ YOLCU | GİRESUN ÜNİVERSİTESİ | TURKEY |
| ÖZGE | TEZEL | KARADENİZ TECHNICAL UNIVERSITY | TURKEY |
| ÖZLEM | AKAY | ÇUKUROVA UNIVERSITY | TURKEY |
| ÖZLEM | ARIK | KÜTAHYA SAĞLIK BİLİMLERİ ÜNİVERSİTESİ | TURKEY |
| ÖZLEM | GOCEN | HACETTEPE UNIVERSITY | TURKEY |
| ÖZLEM | KAYMAZ GÜLLÜ | ANKARA UNIVERSITY | TURKEY |
| ÖZLEM | GÜRÜNLÜ ALMA | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| ÖZLEM | SEVİNÇ | MERKEZ BANKASI | TURKEY |
| ÖZLEM | TÜRKŞEN | ANKARA UNIVERSITY | TURKEY |
| ÖZNUR | İŞÇİ GÜNERİ | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| PAKİZE | YIGİT | İSTANBUL MEDİPOL UNIVERSITY | TURKEY |
| PINAR | GÖKTAŞ | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| RASSOUL | ABDELAZİZ | NATIONAL HIGH SCHOOL OF HYDRAULICS | ALGERIA |
| RICHARD | MINKAH | UNIVERSITY OF GHANA | GHANA |
| RİSTİ | NUR'AINI | İNSTITUT TEKNOLOGI BANDUNG | INDONESIA |
| RÜMEYSA | KARATAŞ | HACETTEPE UNIVERSITY | TURKEY |
| S. EJAZ | AHMED | BROCK UNIVERSITY | CANADA |
| SAMI | AKDENİZ | DOKUZ EYLUL UNIVERSITY | TURKEY |
| SELAHATTİN | KAÇIRANLAR | ÇUKUROVA UNIVERSITY | TURKEY |
| SELİM BAHA | YILDIZ | CELAL BAYAR UNIVERSITY | TURKEY |
| SELMAN | MERMI | MUGLA SITKI KOÇMAN UNIVERSITY | TURKEY |
| SEMRA | ERBAŞ | GAZİ UNIVERSITY | TURKEY |
| SEMRA | ERPOLAT TAŞABAT | MİMAR SİNAN GÜZEL SANATLAR UNIVERSITY | TURKEY |
| SENEM | VAHAPLAR | DOKUZ EYLUL UNIVERSITY | TURKEY |
| SERAP | ULUSAM SEÇKİNER | GAZİANTEP UNIVERSITY | TURKEY |
| SERHAT EMRE | AKHANLI | UNIVERSITY OF COLLEGE LONDON | UNITED KINGDOM |
| SERKAN | ERYILMAZ | ATILIM UNIVERSITY | TURKEY |
| SERKAN | ÖZTÜRK | GUMUŞHANE UNIVERSITY | TURKEY |
| SERPİL | KILIÇ DEPREN | YILDIZ TEKNİK UNIVERSITY | TURKEY |
| SERPİL AKTAŞ ALTUNAY | AKTAŞ ALTUNAY | HACETTEPE UNIVERSITY | TURKEY |
| SHEHZAD MUHAMMAD | AHMED | BAHAUDDIN ZAKARIYA UNIVERSITY | PAKISTAN |
| SİBEL | ERTÜRK | KARADENİZ TECHNICAL UNIVERSITY | TURKEY |
| SİNAN | SARAÇLI | AFYON KOCATEPE UNIVERSITY | TURKEY |
| SONGÜL | BULUT | THE STATE SUPPLY OFFICE | TURKEY |

| | | | |
|---------------|----------------|--|--------------|
| SULTAN | TURHAN | MUGLA SITKI KOCMAN UNIVERSITY | TURKEY |
| SUZAN | KANTARCI | KIRKLARELI UNIVERSITY | TURKEY |
| ŞEBNEM | BURNAZ | ISTANBUL TECHNICAL UNIVERSITY | TURKEY |
| ŞEBNEM | ER | UNIVERSITY OF CAPE TOWN | SOUTH AFRICA |
| ŞERİFE | ÖZKAR | BALIKESİR UNIVERSITY | TURKEY |
| ŞEYMA | YAŞAR | INONU UNIVERSITY | TURKEY |
| TALHA | ARSLAN | VAN YUZUNCU YIL UNIVERSITY | TURKEY |
| TALIP | ARSU | AKSARAY UNIVERSITY | TURKEY |
| TOLGA | YAMUT | DOKUZ EYLUL UNIVERSITY | TURKEY |
| TUBA | ADAR | ATATURK UNIVERSITY | TURKEY |
| TUFAN | BAYDEMİR | OBASE RESEARCH AND DEVELOPMENT TEAM | TURKEY |
| TUĞÇE | EKİZ YILMAZ | DOKUZ EYLUL UNIVERSITY | TURKEY |
| TUĞÇE | LEVENTOĞLU | DOKUZ EYLUL UNIVERSITY | TURKEY |
| TURGUT | ÖZALTINDIŞ | MIMAR SINAN GÜZEL SANATLAR UNIVERSITY | TURKEY |
| UFUK | YOLCU | GİRESUN ÜNİVERSİTESİ | TURKEY |
| UGUR | KARABEY | HACETTEPE UNIVERSITY | TURKEY |
| UĞUR | BINZAT | DOKUZ EYLUL UNIVERSITY | TURKEY |
| UMAY | UZUNOĞLU KOÇER | DOKUZ EYLÜL UNIVERSITY | TURKEY |
| ÜLKÜ | GÜRLER | BILKENT UNIVERSITY | TURKEY |
| ÜMIT | KUVVETLİ | DOKUZ EYLUL UNIVERSITY | TURKEY |
| VEDİDE REZZAN | USLU | ONDOKUZ MAYIS UNIVERSITY | TURKEY |
| VILDAN | YILDIRIM | THE STATE SUPPLY OFFICE | TURKEY |
| VOLKAN | ALPTEKİN | İZMİR KATİP ÇELEBİ UNIVERSITY | TURKEY |
| VOLKAN | YEŞİLDERE | ANADOLU UNIVERSITY | TURKEY |
| WALI KHAN | MASHWANI | KOHAT UNIVERSITY OF SCIENCE AND TECHNOLOGY | PAKISTAN |
| WILLIAM H. | WOODALL | VIRGINIA TECH | TURKEY |
| YAPRAK ARZU | ÖZDEMİR | GAZİ UNIVERSITY | TURKEY |
| YASEMİN | GENÇTÜRK | HACETTEPE UNIVERSITY | TURKEY |
| YONCA | YAZIRLI | ESKİSEHIR TEKNİK UNIVERSİTESİ | TURKEY |
| YUNUS | AKDOĞAN | SELÇUK UNIVERSITY | TURKEY |
| YUSUF CAN | SEVİL | DOKUZ EYLUL UNIVERSITY | TURKEY |
| YÜKSEL AKAY | ÜNVAN | YILDIRIM BEYAZIT UNIVERSITY | TURKEY |
| ZEYNEP | TUNÇ | INONU UNIVERSITY | TURKEY |

5. CONFERENCE PROGRAMME

Programme of 11th ISDC2018 For October 3 (Wednesday)

| MUGLA HALL | |
|--|---|
| 08:30-10:00 | REGISTRATION |
| 10:00 - 10:30 | OPENING CEREMONY |
| 10:30 - 10:40 | COFFEE BREAK |
| 10:40-12:20 INVITED SESSION 1 | A New Difference-based Weighted Mixed Liu Estimator in Partially Linear Models Fikri AKDENİZ |
| 12:20 - 14:00 | LUNCH TIME |
| 14:00-14:40 INVITED SESSION 2 | Statistics and Industry: Opportunities, Responsibilities and Examples Eleanor STILLMAN |
| 14:40-15:20 INVITED SESSION 3 | Compound Sums Serkan ERYILMAZ |
| 15:20 - 15:40 | COFFEE BREAK |
| 15:40-17:40 INVITED SESSION 4 | Statistical Learning In Big Data Analytics Ejaz AHMED Advanced Inferences for High Dimensional Data Analytic in Bioinformatics Halima BENSMAIL |
| 17:50-19:00 | POSTER PRESENTATIONS |
| 20:30-23:00 | COCKTAIL PARTY |

Programme of 11th ISDC2018 For October 4 (Thursday)

| | MUGLA HALL | BODRUM HALL |
|---|---|--|
| 09:00-09:40 INVITED SESSION 1 | First Passage Time of Degradation Process Narayanaswamy BALAKRISHNAN | POSTER PRESENTATIONS |
| 09:40-10:20 INVITED SESSION 2 | Robust Bayesian Relevance Vector Machines in Regression and Supervised Classification Using Information Complexity and the Genetic Algorithm Hamparsum BOZDOĞAN | |
| 10:20-10:40 | COFFEE BREAK | |
| Session Chair | Erol EĞRİOĞLU | Khreshna SYUHADA |
| SESSION 1 10:40- 12:20 | A New Intuitionistic High Order Fuzzy Time Series Method Based on Robust Regression | Risk Sharing Prediction of Indonesian Co-Operative |
| | Cem KOÇAK, Erol EĞRİOĞLU, Eren BAŞ, Ufuk YOLCU | Khreshna SYUHADA, Listiarini MUHTARI |
| | An Intuitionistic Fuzzy Time Series Prediction Method based on Sigma-Pi Neural Network | Modifying Clements' Method with Alternative Skewness and Kurtosis Estimators |
| | Ufuk YOLCU, Erol EĞRİOĞLU, Eren BAŞ, Özge CAGCAG YOLCU | Burcu AYTAÇOĞLU, Cemil Giray GENÇ |
| | Combination of Forecasting Methods based on Fuzzy C-Means Clustering | Adaptive Modified Maximum Likelihood Estimation for AR(1) Model under Generalized Logistic Innovations |
| | Nihat TAK, Erol EĞRİOĞLU, Eren BAŞ and Ufuk YOLCU | Ayşen DENER AKKAYA, Özlem TÜRKER BAYRAK |
| | A Novel Fuzzy TOPSIS Method: Fuzzy DSC TOPSIS | Use of Relative Entropy Statistics in Contingency Tables |
| | Semra ERPOLAT TAŞABAT | Atif EVREN |
| | Calculating Happinies Index With Fuzzy DSC TOPSIS Method | Understanding Hospital-Acquired Malnutrition in Pediatric Population through Statistical Modelling |
| Semra ERPOLAT TAŞABAT, Nalan CİNEMRE | Dessie WANDA, Khreshna IA SYUHADA, Nani NURHAENI, Chairun NARDIANTARI | |
| 12-20-14:00 | LUNCH TIME | |
| 14:00-14:40 INVITED SESSION 3 | An Overview of Methods for Social Network Monitoring William H. WOODALL | |
| 14:40-15:20 INVITED SESSION 4 | Data Analysis in Health and Big Data Gökhan SİLAHTAROĞLU | |
| 15:20-15:40 | COFFEE BREAK | |

Programme of 11th ISDC2018 For October 4 (Thursday)

| | MUGLA HALL | BODRUM HALL |
|--|--|--|
| Session Chair | Gözde YAZGI TÜTÜNCÜ | Hakan Savaş SAZAK |
| SESSION 2 15:40- 17:00 | Assembling of Structural Equation Model and Fuzzy Analytic Hierarchy Process Methods: An Application for Residence Selection | Creation of a University League Based on the Contribution of their Economics Departments to Economics Education |
| | Özlem AKAY, İlkay ALTINDAĞ | S. Baha YILDIZ, Volkan ALPTEKİN, Sibel SELİM |
| | Investigating the Use of Mobile Banking Application by the Technology Acceptance Model (TAM) | Bootstrap and Jackknife Approximate Confidence Intervals and Performance Comparisons for Population Coefficient of Variation |
| | İlkay ALTINDAĞ Ebru ÖZER TOPALOĞLU | Hayriye ESRA AKYÜZ, Hamza GAMGAM |
| | The Use of Some Loss Functions in Fuzzy Time Series Forecasting Problem | A Comparative Study on the Robust Estimators for the Parameters of the General Linear Regression Model via Simulation |
| | Özge CAGCAG YOLCU, Eren BAŞ, Erol EĞRİOĞLU, Ufuk YOLCU | Hakan Savaş SAZAK, Nalan MUTLU |
| | New (r, Q) Inventory Control Models in the Presence of Misplaced items | Inventory Control Policy Decisions in the Presence of Servers Subject to Random Failure |
| | Gözde YAZGI TÜTÜNCÜ, Ceki FRANKO, Linda ZHANG | Şerife ÖZKAR, N. Onur BAKIR |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Ceyhan İNAL | Güçkan YAPAR |
| SESSION 3 17:20- 18:40 | Hybrid Type-1 Fuzzy Time Series Functions Approaches: The Application in Stock Exchanges | An Application of Unrelated Parallel Machine Scheduling with Sequence-Dependent Setups |
| | Ali Zafer DALAR, Erol EĞRİOĞLU | Merve Burcu SARIKAYA, Okan Örsan ÖZENER, Ali EKİCİ |
| | Fuzzy Regression Function Approach based on Moving Block Bootstrap Method | A Load Exchange Mechanism for Less-than-Truckload Operations |
| | Eren BAŞ, Erol EĞRİOĞLU, Ufuk YOLCU | Başak ALTAN, Okan Örsan ÖZENER |
| | Using Multi-Response Taguchi Method (MRSN) for the Achievement of Product Optimization with Six Sigma Methodology in the Automotive Industry | Time Series Forecasting of Intermittent Demand by Using Ata Method and Computational Intelligence |
| | Atakan GERGER, Ali Rıza FİRUZAN | Tuğçe EKİZ YILMAZ, Güçkan YAPAR |
| | A New Method for Measuring Service Quality: An Application in Public Transportation Sector | M-Competitions: ATA Method and Others |
| Ümit KUVVETLİ, Ali Rıza FİRUZAN, Atakan GERGER | Güçkan YAPAR | |

Programme of 11th ISDC2018 For October 4 (Thursday)

| | MARMARIS HALL | KÖYCEĞİZ HALL |
|---|---|---|
| Session Chair | Esin FİRUZAN | Tahir KHANIYEV |
| SESSION 1 10:40- 12:20 | The Impact of Logistics on the Economic Series in Turkey | Cardiovascular Disease Population Risk Tool (CVDPoRT): The development and application of a risk prediction model using population health surveys |
| | Berhan ÇOBAN, Şerife ÖZKAR | Meltem TUNA |
| | More Efficient Estimator for Unit Root and Cointegration: RALS Approach | A Macro Level Comparison of Hospital Efficiency in Turkey by Using Malmquist Total Factor Productivity Index |
| | Berhan ÇOBAN, Esin FİRUZAN | Ceren DİRİK, Serap ŞAHİN |
| | Financial Stress Tests for Measuring Vulnerability of Turkish Economy | Investigation of the Effect of the Phase Space Parameters on Prediction Performance in Chaotic Time Series |
| | Tuğçe LEVENTOĞLU, Esin FİRUZAN | Ayşe İŞİ |
| | Comparison of Bayes and OLS based Type II Regression | Time Series Based Decision Support System to Reduce Unexpected Airline Diverts |
| | Ece ÖZGÖREN, Sinan SARAÇLI | Hazal BERVE DOĞAN, Derya GÖZEN, Başak GEVER, Umut DEMIREZEN, Tahir KHANIYEV |
| Precursory Seismicity Rate Changes associated with the earthquakes in the East Anatolian Region of Turkey: An Application of Region-Time-Length (RTL) Algorithm | A Generalization of Normal Distribution with Linear Model Application | |
| Serkan ÖZTÜRK, Santi PAILOPLEE | Demet SEZER, Nuri ÇELİK | |
| 12:20-14:00 | LUNCH TIME | |
| Session Chair | Serkan ÖZTÜRK | Orhan KESEMEN |
| SESSION 2 15:40- 17:00 | Inversion of Electrical Resistivity Data: An Application from Gümüşhane University Football Field | Fuzzy One-Way ANOVA and An Application |
| | Mahmut SARI Serkan ÖZTÜRK | Derviş TOPUZ, Suat ŞAHİNLER |
| | Improved RALS(2)-LM Cointegration Test in presence of structural breaks and GARCH innovations | Fuzzy Least Squares Regression Analysis and An Application |
| | Esin FİRUZAN, Berhan ÇOBAN | Derviş TOPUZ, İsmail KESKİN, Suat ŞAHİNLER |
| | The Use of Joinpoint Regression Analysis in Core healthcare Indicators in Turkey, 2003-2016 | Estimation and forecasting with stationary vector autoregressive (VAR) model of financial data in Turkey |
| | Pakize YİĞİT, Seda KUMRU | Dursun AYDIN, Öznuri ŞÇİ GÜNERİ, Burcu DURMAZ |
| Estimation of Turkish Court of Cassation Decisions via Machine Learning and Natural Language Processing Methods | Detection of Lines on Images with Multimodal Directional Artificial Bee Colony Algorithm | |
| Enes DİLBER, Melih AĞRAZ | Orhan KESEMEN, Eda ÖZKUL, Buğra Kaan TIRYAKI, Özge TEZEL | |
| 17:00-17:20 | COFFEE BREAK | |

Programme of 11th ISDC2018 For October 4 (Thursday)

| | MARMARIS HALL | KÖYCEĞİZ HALL |
|--|--|---|
| Session Chair | Oğuz AKPOLAT | Serpil AKTAŞ ALTUNAY |
| SESSION 3 17:20- 18:40 | Modelling of Wind Speed Data for Konya | A New Approximation for Fuzzy Directional Clustering(FDC) |
| | İsmail KINACI, Emine KARAKAYA , Coşkun KUŞ | Orhan KESEMEN, Özge TEZEL, Eda ÖZKUL, Buğra Kaan TIRYAKI |
| | Experimental Design Applications in Chemistry | Evaluation of Fuzzy Linear Regression Models-Based on Mathematical Programming with Application to Animal Science |
| | Oğuz AKPOLAT | Aslı AKILLI, Özkan GÖRGÜLÜ, Sabri GÜL, Mahmut KESKIN |
| | A Comparison Study of Optimal Bandwidth Selection with Density Functional Estimation for Contaminated Data | A Shiny App for Subgroup Discovery Using Activity Region Finder (ARF) |
| | Celal AYDIN, Necla GÜNDÜZ | Serpil KILIÇ DEPREN, Praneeth VEPAKOMMA, Birol EMIR, Javier CABRERA |
| | Optimal Bandwidth Selection for a Kernel Density with a Location-Scale Property | Estimation of Bone Age from Radiological Images with Soft Computing Methods |
| Celal AYDIN, Necla GÜNDÜZ, Jale BALİBEYOĞLU | Nida GÖKÇE NARİN, İbrahim Önder YENİÇERİ, Gamze YÜKSEL | |
| | FETHİYE HALL | DALYAN HALL |
| Session Chair | Ali Rıza FIRUZAN | Fatih ÇEMREK |
| SESSION 1 10:40- 12:20 | Liu Pretest and Stein-Type Shrinkage Estimation in Linear Regression Models | Directional k-means Algorithm Using Signed Angular Distance |
| | Bahadır YÜZBAŞI S. Ejaz Ahmed Yasin ASAR | Orhan KESEMEN, Buğra Kaan TIRYAKI, Özge TEZEL, Eda ÖZKUL |
| | Parametric Joint Modelling for Longitudinal and Survival Data | Economic Institutions and FDI Flows: A Cross-country Investigation |
| | Elif DİL, Duru KARASOY | Mehmet NasihTAĞ, Süleyman DEĞİRMEN |
| | Use of Heijunka Technique to Reduce Production Variability by Leveling Excessive Demand Increase/Decrease | Spatial Econometric Modeling of Social Capital Effect on Internal Migration |
| | Atakan GERGER, Ali Rıza FIRUZAN, Ümit KUVVETLİ | Esra SATICI, Halis KIRAL, Barış ALPASLAN |
| | Regression Mixture Model with Two Components for Count Data: An Application to Traffic Accidents | Semimarkovian Model for Prediction of the Severity of a Terrorist Threat |
| | Hande KONŞUK ÜNLÜ, Ayten YIĞİTER , Hilal ÖZCEBE, Burcu KÜÇÜKBIÇER | Martin TEJKAL, Jaroslav MICHÁLEK |
| Pattern Recognition in Control Charts with Neural Networks | The Effect of Biodiesel Market and Policies in European Union and Turkey on Economic Growth | |
| Alper VAHAPLAR Senem ŞAHAN VAHAPLAR | Fatih ÇEMREK, Hüseyin Naci BAYRAÇ | |
| 12-20-14:00 | LUNCH TIME | |

Programme of 11th ISDC2018 For October 4 (Thursday)

| | FETHİYE HALL | DALYAN HALL |
|---|--|---|
| Session Chair | Selahattin KAÇIRANLAR | Sinan SARAÇLI |
| SESSION 2 15:40- 17:00 | Workforce Forecasting | Comparison of Classification Methods for Galaxies in the Shapley Concentration Region |
| | Buse DEDEKAYOĞULLARI | Nida GÖKÇE NARIN, Nazlı Deniz ERGÜÇ |
| | Adjusting of WABL Parameters for Decision Tree Algorithm with Fuzzy Numbers: A case study | Comparison of L1, L2 and Least Squares Ratio Estimators |
| | Suzan KANTARCI SAVAŞ, Efendi NASIBOV | Ahmet TOY, Erol TERZI |
| | Continuum Regression and Comparisons of Some Related Estimators | Interval Estimation for Order Statistics by Chebyshev Type of Inequalities |
| | Berrin GÜLTAY, Hüseyin GÜLER, Selahattin KAÇIRANLAR | Erhan USTAOĞLU |
| | A Structural Equation Model For Success Measurement Of Development Projects: The Case of Istka Projects | Modeling the Financial Attitudes and Behaviors of Y Generation via Structural Equation Modeling |
| Seda YANIK OZBAY, Meltem YONTAR | Murat AKŞIT, Sinan SARAÇLI | |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Serhat Emre AKHANLI | Ayten YİĞİTER |
| SESSION 3 17:20- 18:40 | Cox Regression Model with Covariates Subject to Measurement Error | A Queuing Inventory System with Interruption and Two-Demand Classes |
| | Hatice IŞIK, Duru KARASOY | Umay UZUNOĞLU KOCER, Serife OZKAR |
| | Calibration of Internal Validation Indexes for Estimating the Number of Clusters | Modified Tukey's Control Chart Based on Pairwise Distance Scale Estimator |
| | Serhat Emre AKHANLI | Hayriye Esra AKYÜZ, Moustafa Omar Ahmed ABU-SHAWIESH |
| | Comparison of Least Squares Estimator with M-Type Regression Estimators | Control Charts Based on Com-Poisson Distribution |
| | Ahmet TOY, Erol TERZI, Cemil ÇOLAK | Canan HAMURKAROĞLU, Ayten YİĞİTER, Yasemin GENÇTÜRK |
| | Results of A Project of Marriages Between Disabled People | Site Effect Classification Estimated from H/V Spectral Ratio of Single Station Microtremor Data in Kürtün-Özkürtün Districts of Gümüşhane, Turkey |
| Ceyhun MAZLUM, Selma YILDIZ | Serkan ÖZTÜRK, Yasemin BEKER, Levent PEHLIVAN, Mahmut SARI | |
| | MİLAS HALL | DATÇA HALL |
| Session Chair | Yaprak Arzu ÖZDEMİR | Angelita RUVALCABA |
| SESSION 1 10:40- 12:20 | An Alternative Testing Procedure for The Equality of Two Population Means Using Median Ranked Set Sampling | Ranking multi-dimensional poverty in Mexico for better, prioritized policies: through a partial order algorithm |
| | Meral EBEGİL, Yaprak Arzu ÖZDEMİR, Fikri GÖKPINAR | Angelita RUVALCABA, Maurizio VICHI, Enrico di BELLA |
| | Some Shrinkage Estimators Based on Median Ranked Set Sampling | Characterization of hazard function of Some Beta-G distributions |
| Meral EBEGİL, Yaprak Arzu ÖZDEMİR, Fikri GÖKPINAR | Sharqa HASHMI, Ahmed Zogo MEMON | |

Programme of 11th ISDC2018 For October 4 (Thursday)

| | MİLAS HALL | DATÇA HALL |
|------------------------------|---|---|
| Session Chair | Yaprak Arzu ÖZDEMİR | Angelita RUVALCABA |
| SESSION 1 10:40- 12:20 | Evaluation of Epistasis Effect in Genome-Wide Association Studies with Ant Colony Optimization Technique | An application of disjoint principal component analysis based on sustainable livelihood framework: evidence from a cash transfer program data in Sub-Saharan Africa |
| | Havva DİDEM ÇELİKCAN, Bahar TAŞDELEN | Angelita RUVALCABA, Enrico di BELLA |
| | Outlier Detection and Variable Selection for Count Data | Mix EWMA-CUSUM Control Chart In the Presence Of Measurement Error |
| | Fatma Sevinç KURNAZ | Muhammad NOOR-UL-AMIN, Afshan RIAZ |
| | Turkish Energy Sector Analysis: A Bootstrap Data Envelopment Analysis Approach | A Bayesian Analysis of Order Change in Autoregressive Time Series Model |
| | Esra Betül KINACI, H. Hasan ÖRKÇÜ, Hasan BAL | Saurabh KUMAR, Jitendra KUMAR, Vikas Kumar SHARMA |
| 12-20-14:00 | LUNCH TIME | |
| Session Chair | Ozan KOCADAĞLI | Zahid KHAN |
| SESSION 2 15:40- 17:00 | Efficiency Evaluation of Ankara Public Transport Bus Lines Using DEA | Truncated Based Median Ranked Set Sampling for Estimation of Population Mean |
| | Harun KINACI, Hayri ULVI | Zahid KHAN, Muhammad ISMAIL |
| | Efficiency Analysis of Ankara Rail and Light Rail System Stations | Fourier Series Estimator in Nonparametric Regression Approach for Longitudinal Data based on Generalized of Penalized Weighted Least Square |
| | Harun KINACI, Abdullah ORMAN | M. Fariz Fadillah MARDIANTO, Sri Haryatmi KARTIKO, Herni UTAMI |
| | Comparing performances of artificial intelligence techniques based wavelet transforms and feature selection for detection of epileptic seizures | Estimation of the Tail Index of a Generalised Pareto Distribution using a Pivot from a Transformed Pareto Distribution |
| | Ozan KOCADAĞLI, Ezgi ÖZER | Richard MINKAH, Tertius de WET |
| | Fuzzy Survival Analysis for Weibull and Gamma Distribution | Conflicts in the Calculation of the Power of a Statistical Test, by Psychology Students. A Semiotic Point of View. |
| | Gözde Ulu METİN, Nevin GÜLER DİNCER | Osmar Darío VERA |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Gülay BASARIR | Marek SZAJT |
| SESSION 3 17:20- 18:40 | Comparison of Supervised Disease Diagnosis Methods for Unbalanced Classes: The Case of Diabetes Diagnosis | Application of Extreme Value Theory (EVT) to Inflation Rate Targeting: The Ghanaian case |
| | Sultan TURHAN, Yüksel ÖZKAN, B. Sarer YUREKLI, Aslı SUNER, Eralp DOĞU | Ezekiel Nii Noye NORTEY, Eric OCRAN, Abeku Atta Asare-KUMI, Richard MINKAH |
| | Ranking Communication Apps based on Android OS by Applying an Integrated Decision-Making Approach | Determinants and Regional Variations of Cereal Crops Productivity in Ethiopia |
| | Sevgi ABDALLA | Hayimro MERIE, Girma TAYE |
| | BLS Knowledge: A comparison between clinical A comparison between clinical and basic medical professionals | Measurement of the region's functions on the example of the analysis of the level of innovation |
| | Anam REHMAN, Ujala SHUJAT, Nuha MAHMOOD | Marek SZAJT, Agata MESJASZ-LECH |
| | Determining Anomalies of BIST 100 Index by BURR XII Distribution | Improvement of Vehicles' Production by Creating the Intelligent Information System to Verify Manufacturability of Design Documentation |
| | Esra Betül KINACI, H. Hasan ÖRKÇÜ, Hasan BAL, Harun KINACI | Irina MAKAROVA, Ksenia SHUBENKOVA, Vadim MAVRIN |

Programme of 11th ISDC2018 For October 5 (Friday)

| | MUGLA HALL | BODRUM HALL | |
|-------------------------------------|--|--|--|
| 09:00-09:40 INVITED SESSION 1 | Bayesian Joint Estimation of Primary and Substitute Demands with Lost Sales Ükü GÜRLER | | |
| 09:40-10:20 INVITED SESSION 2 | A Stochastic Fuzzy Time Series Forecasting Model Ç. Hakan ALADAĞ | | |
| 10:20-10:40 | COFFEE BREAK | | |
| Session Chair | Yunus AKDOĞAN | Mehmet Nasih TAĞ | |
| SESSION 1 10:40-12:20 | Conditional VaR Forecast for Dependent Losses Risti NUR'AINI, Khreshna Syuhada MAHFUDHOTIN | Examining the Efficiency of Hydroelectric Power Plants in Turkey Metehan SÖNMEZ, Doğan YILDIZ, Mehmet Samil GÜNEŞ | |
| | The Effects of Ranking Error Models on Mean Estimators Based on Ranked Set Sampling Sami AKDENİZ, Tuğba Özkal YILDIZ, Yusuf Can SEVİL | Evaluation of Criteria for Selection of Personnel Capable Work Rotation Using Autocratic MAGDM Based on Interval Valued Intuitionistic Fuzzy Set (IVIF) Tuba ADAR, Elif KILIÇ DELİCE | |
| | Parameter Estimation for Uniform-Geometric Distribution Based on Censored Sample Mehtap Koca YILMAZ, Yunus AKDOĞAN, Kadir KARAKAYA | What Sample Sizes is Needed in Validation Studies? Arzu BAYGÜL, Neslihan GÖKMEN | |
| | Exploring Relationship among Quantitative Traits of Sugarcane Varieties Using Principal Component Analysis Irum RAZA | Financial Time Series Analysis with parametric and non parametric methods: a comparative study Selman MERMİ, Dursun AYDIN | |
| | Correlation Based Penalty Estimator in Vector Autoregressive Model Bahadır YÜZBAŞI, Ahmet DEMİRALP, Mehmet GÜNGÖR | Antecedents of High-Impact Entrepreneurship: Evidence from Multi-level Analysis Mehmet Nasih TAĞ | |
| | LUNCH TIME | | |
| | 14:00-14:40 INVITED SESSION 3 | A New Multiple Model Approach based on Artificial Neural Network for Forecasting Erol EĞRİOĞLU | |
| | 14:40-15:20 INVITED SESSION 4 | | |
| 15:20-15:40 | COFFEE BREAK | | |
| Session Chair | Özlem TÜRKŞEN | Johannes BG FRENK | |
| SESSION 2 15:40-17:00 | GIS-based Approach for Defining Bioenergy Facilities Location: A Case Study in Bismil/Turkey Based on Marginal Delivery Costs and Resources Competition Between Facilities Serap ULUSAM SEÇKİNER, Adem AKGÜL | Cryptocurrencies, Currencies and Gold Fractal Volatility Analysis Mária BOHDALOVA, Michal GREGUS | |

Programme of 11th ISDC2018 For October 5 (Friday)

3

| | MUGLA HALL | BODRUM HALL | |
|--|---|--|---|
| Session Chair | Özlem TÜRKŞEN | Johannes BG FRENK | |
| SESSION 2 15:40-17:00 | Estimating First Order Inclusion Probability Based on Percentile Ranked Set Sampling Yusuf Can SEVİL, Tuğba Özkal YILDIZ | Investigation of the Employment by Educational attainment: Evidence from the 15 European Union Member Countries Marina FAD'OS, Mária BOHDALOVA | |
| | A Statistical Characterization of a Simulated Earthquake Fault Plane Area by Using Spatial Statistics Özlem TÜRKŞEN | Estimating the Total Number of Defective Items: A Case Study Ayda AMNİATTALALAB, . J.B.G Frenk | |
| | Conditional Value-at Risk Modelling on Oil Prices for Emerging Markets Tolga YAMUT, Burcu HÜDAVERDİ UÇER | Temperature Interpolation Using Spatio-temporal Kriging for Incomplete Data Daniela López SARMIENTO, Wilmer PINEDA-RIOS | |
| | COFFEE BREAK | | |
| | Session Chair | İlker ERCAN | Berihan ELEMARY |
| SESSION 3 17:20-18:40 | A Dual Response Problem Based on Median and Median Absolute Deviation Gözde ÖZÇIRPAN, Ahmet KOCATÜRK, Elif ŞEN, Özlem TÜRKŞEN | Comparison of two imputation methods for estimation for the right-censored nonparametric regression model Ersin YILMAZ, Dursun AYDIN | |
| | On the Efficiency of Proportion Estimators Under Modified Ranked Set Sampling Methods Aylin GÖÇOĞLU, Neslihan DEMİREL | Algorithms for Computing the Reliability of (r,s) out of n Systems Gökhan GÖKDERE, Selahattin GÜNEŞ, Duygu TANAY | |
| | Statistical Shape Analysis of Hand and Wrist in Children Ural KOÇ, İlker ERCAN, Semih BOLU, Ayşegül YABACI, Onur TAYDAŞ | Evaluation and Improvement of Promising Rubber Recycling OT Machine Using Fractional Factorial and Response Surface Statistical Design Berihan ELEMARY | |
| | The Proposed F Approach for Estimation of Missing Landmark Fatma Ezgi CAN, İlker ERCAN | Robust Estimators in Linear Regression : A Simulation Study and An Application on Health Care Neslihan GÖKMEN, Aydın ERAR | |
| | MARMARİS HALL | | KÖYCEĞİZ HALL |
| | Session Chair | Selahattin KAÇIRANLAR | Ebru ÖZGÜR GÜLER |
| | SESSION 1 10:40-12:20 | Efficient Algorithms to Compute Signatures of Binary Systems Gökhan GÖKDERE, Fatih AYDIN, Ayşenur SÖNMEZ | Statistical Investigation of the Effect of Burnout and Job Satisfaction on Academicians Ebru ÖZGÜR GÜLER, Dilek VEYSİKARANI |
| Hierarchical Forecasting Mehmet GÜLŞEN, Pelin TOKTAŞ | | Pricing Of Long-Term Care Insurance For Turkey Çiğdem LAZOĞLU, Murat BÜYÜKYAZICI | |
| HIV-1 Infection Prognosis: Evolutionary Game Theory Approach Bilge BAŞER, Metin YANGIN | | Investigation of the Attitudes of Academicians About the Individual Pension System: The Case of Eskişehir Osmangazi University Fatih ÇEMREK | |

Programme of 11th ISDC2018 For October 5 (Friday)

| | MARMARİS HALL | KÖYCEĞİZ HALL |
|---|---|--|
| Session Chair | Selahattin KAÇIRANLAR | Ebru ÖZGÜR GÜLER |
| SESSION 1 10:40- 12:20 | Optimal Determination of the Parameters of Some Biased Estimators Using Genetic Algorithm | Performance Evaluation Of ISE30 (Istanbul Stock Exchange) Stock Certificates and Formation of Portfolio By Using Multi - Criteria Decision Making Techniques |
| | Erkut TEKELİ, Selahattin KAÇIRANLAR | Yüksel Akay ÜNVAN, Hüseyin TATLIDİL |
| | Evaluating the Gage Repeatability and Reproducibility for Industries | |
| | Ali Rıza FİRUZAN, Ümit KUVVETLİ, Atakan GERGER | |
| 12-20-14:00 | LUNCH TIME | |
| Session Chair | Mustafa NADAR | Yasemin GENÇTÜRK |
| SESSION 2 15:40- 17:00 | Probability Density Function Estimation Using Scaled Inverse Chi-Squared Kernel | An Interactive Web Application for Kruskal Wallis H Test with R Shiny Package |
| | Elif ERÇELİK, Mustafa NADAR | A.Kadir ARSLAN, Şeyma YAŞAR, Cemil ÇOLAK, Saim YOĞLU |
| | Differential Network Analysis with Partial Least Square Regression | Comparison of the Claims Development Results Using the Chain Ladder and Extended Complementary Loss Ratio Methods for Turkey |
| | Aylin ALIN, Ayça ÖLMEZ, Gökhan KARAKÜLAH, Aslı SUNER | Müge YELDAN, Yasemin GENÇTÜRK |
| | Risk Evaluation for Support: Predictions for Elder-Life in Community Tool | A Computer Application for a Decision Making Algorithm Which is Based on the Soft Set Theory |
| | Meltem TUNA | Gözde YAYLALI UMUL, Nazan ÇAKMAK POLAT, Bekir TANAY |
| | Fuzzy ELECTRE I Method for Plant Location Selection: a Case Study | Analysis of PISA-2015 Performance of Turkish Students by Multilevel Structural Equation Modeling |
| Kezban BULUT, Suna ÇETİN, Metehan ZOĞAL, Kemalcan BEDİZ | Ertan AKGENÇ, Nimet YAPICI PEHLİVAN | |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Gökhan SİLAHTAROĞLU | Talha ARSLAN |
| SESSION 3 17:20- 18:40 | Stacking Ensemble Approach for the Classification of Survival Status: HCC patients | Estimation of Diabetic Patients Data by Semi-Parametric Regression Methods |
| | Metin YANGIN, Bilge BAŞER, Ayça ÇAKMAK PEHLİVANLI | Dursun AYDIN, Özlem ARIK |
| | Comparison of Some Estimators for Parameters of log-Dagum Distribution | Robust Estimation of Response Surface |
| | Caner TANIŞ , Merve ÇOKBARLI, Buğra SARAÇOĞLU | Talha ARSLAN, Şükrü ACITAŞ, Birdal ŞENOĞLU |

Programme of 11th ISDC2018 For October 5 (Friday)

| | MARMARİS HALL | KÖYCEĞİZ HALL |
|--|---|---|
| Session Chair | Gökhan SİLAHTAROĞLU | Talha ARSLAN |
| SESSION 3 17:20- 18:40 | Characterization Parameters of Domestic Wastewater Samples Collected from Muğla District in Turkey by Decision Tree Method | Performance Comparison of Some Imputation Methods Used in Missing Value(s) Analysis: A Simulation Study |
| | Sait GÜLLERİ, Gökhan SİLAHTAROĞLU, Oğuz AKPOLAT | A. Kadir ARSLAN, Zeynep TUNÇ, Emek GÜLDOĞAN, Cemil ÇOLAK |
| | Restricted Bridge Estimator | Classification of Hydrocephalus and Determination of Related Factors by Machine Learning Methods |
| | Hüseyin GÜLER, Ebru ÖZGÜR GÜLER | Zeynep TUNÇ, Cemil ÇOLAK, Ramazan ÖZDEMİR |
| | FETHİYE HALL | DALYAN HALL |
| Session Chair | Hülya ÇINGİ | Masoud Muhammed HASSAN |
| SESSION 1 10:40- 12:20 | Bayes Estimation of the Reliability Function and Hazard Rate of the Topp–Leone Distribution | A Two-stages MCMC Approach to Tree-ring Dating in The Absence of Master Chronologies |
| | İlhan USTA, Merve AKDEDE | Masoud Muhammed HASSAN |
| | Estimation of the Inverse Rayleigh Distribution: A Comparison Study | Empirical Estimator of the Conditional Tail Moment |
| | İlhan USTA, Merve AKDEDE | Laidi MOHAMED, Rassoul ABDELAZİZ, Ould Rouis HAMID |
| | An Interactive Web Application for Sample Size and Power Analysis with R using "Shiny" for Quantitative Variables | Iranian Women's Childbearing Desire; Cart Approach |
| | A.Kadir ARSLAN, Şeyma YAŞAR, Cemil ÇOLAK, Saim YOLOĞLU | Arezoo BAGHERI, Mahsa SAADATI |
| | Analysis Of Worldwide Exchange Rates Via Financial Network Techniques | Study and Modeling of the Extreme Rainfalls (case: Khemis Miliana, Algeria) |
| | Yusuf Yargı BAYDILLİ, İlker TÜRKER | Guermah TOUFIK, Rassoul ABDELAZIZ |
| Factors Affecting the Job Choice of Industrial Engineering Graduates | A Model on the Total Number of Defective Items | |
| | Damla Beril ÖZYAZGAN, Erkan İŞIKLI | J.B.G Frenk, Ayda AMNIATTALAB |
| 12-20-14:00 | LUNCH TIME | |
| Session Chair | Elif ÇOKER | Atalay ÇAĞLAR |
| SESSION 2 15:40- 17:00 | Estimation of Reliability for Consecutive k-out-of-n System in Stress-Strength Setup Based on Proportional Hazard Rate Family | An Efficient Treatment for Best-Worst Method with Multiple Best and/or Worst Criteria |
| | Duygu DEMİRAY, Fatih KIZILASLAN | Hüseyin KOÇAK, Atalay ÇAĞLAR, Gülin Zeynep ÖZTAŞ |
| | A Comparative Application of Multiple Correspondence and Joint Correspondence Analysis: A Case Study of Turkey | The Iterative Method for Multidimensional Objects Ranking |
| | Elif ÇOKER | Małgorzata MARKOWSKA, Danuta STRAHL |
| | Examining the Matching Algorithms in Propensity Score Analysis | Discriminating between the Lognormal and Weibull Distributions under Progressive Censoring |
| | Meral YAY, Elif ÇOKER, Elif Özge ÖZDAMAR | Coşkun KUŞ, Ahmet PEKGÖR, İsmail KINACI |
| 17:00-17:20 | COFFEE BREAK | |

Programme of 11th ISDC2018 For October 5 (Friday)

| | FETHİYE HALL | DALYAN HALL |
|-------------------------------------|---|--|
| Session Chair | Mustafa Çağatay KORKMAZ | Bünyamin SARIBACAK |
| SESSION 3 17:20- 18:40 | Investigation of European Football Clubs Efficiencies with Bi-Objective Multi Criteria Data Envelopment Analysis (BiO-MCDEA) | Clustering Analysis in Spatio-Temporal Data Mining |
| | Talip ARSU | Turgut ÖZALTINDIŞ, Elif Özge ÖZDAMAR |
| | A New Extended Lindley Distribution | Performance Comparisons of Some Classification Algorithms Used in Machine Learning |
| | Mustafa Çağatay KORKMAZ | Bünyamin SARIBACAK, Erol TERZİ |
| | The Effect of Methods Used for Missing Data Imputation on Classification Success | Constructing Discrete Bayesian Networks from Mixed Data in R |
| | Hüseyin BUDAK, Enis GÜMÜŞTAŞ | Günel BİLEK, Efehan ULAŞ |
| | Penalized Logistic Regression for Classification : A Case Study for Non-OECD Countries | Bus Driver Scheduling in Interurban Transportation System |
| | Ezgi NAZMAN, Semra ERBAŞ, Hülya OLMUŞ | Ali Rıza GÜLLER, Mehmet Ulaş ÇAKIR, Gizem EROL |
| | MİLAS HALL | DATÇA HALL |
| Session Chair | Vedide Rezzan USLU | Wali Khan MASHWANI |
| SESSION 1 10:40- 12:20 | Understanding Drought with copula functions: Case study for Konya Province | Hybrid Differential Evolutionary Algorithms for Global Optimization Problems |
| | O. Ozan EVKAYA, Ceylan YOZGATLIGİL, A. Sevtap SELÇUK KESTEL | Wali Khan MASHWANI |
| | Optimising the biasing parameters of Liu Type estimator based on Ridge Regression using with Genetic and Differential Evolution Algorithm | Adaptive Principle Components Calibration in Survey Sampling |
| | Mehmet Arif DEMİRCİ, Vedide Rezan USLU | Shehzad MUHAMMAD AHMED, Khurram HARIS, Sajjad Haider BHATTI |
| | Forecasting of per capital consumption of electricity with Autoregressive models based on wavelet decomposition | Lymphoedema Management to Prevent Acute Dermatolymphangioadenitis in Podoconiosis in Northern Ethiopia (GoLBeT): A Pragmatic Randomised Controlled Trial |
| | Aytaç PEKMEZCİ, M. Oğuzhan YALÇIN | Henok NEGUSSIE, Meseret Molla KASSAHUN, Moses NGARI |
| | | Stacking Method for Determining Weights in Partial Least Squares Model Averaging |
| | | Anang KURNIA, Bagus SARTONO, Muhammad A. RAMADHAN |
| 12-20-14:00 | LUNCH TIME | |
| Session Chair | Mehmet GÜLŞEN | Mohamed BOUABAZ |
| SESSION 2 15:40- 17:00 | Clustering Bank Customers Using Self-Organizing Maps | Using of Artificial Intelligence and Statistical Methods in Projects Management |
| | Seda YANIK, Abdelrahman ELMORSY, Meltem YONTAR | Mohamed BOUABAZ, Otmani AMİRA |
| | An Intuitionistic Failure Mode and Effect Analysis Proposal to Determine the Most Important Corrective and Preventive Strategies | Estimation of Diseases in Students Environment |
| | Gülin Feryal CAN, Mehmet GÜLŞEN | Malika BOUGHANDJIOUA, Mohamed BOUABAZ |

Programme of 11th ISDC2018 For October 5 (Friday)

| | MILAS HALL | DATÇA HALL |
|---|--|---|
| Session Chair | Mehmet GÜLŞEN | Mohamed BOUABAZ |
| SESSION 2 15:40- 17:00 | A Comparative Study on Robust Design from the Perspective of Estimators | Generating Discrete Analogues of Continuous Bivariate Probability Distributions |
| | Onur KÖKSOY, Melis ZEYBEK | Alessandro BARBIERO |
| | Clustering Active Users and Potential Users of a Mobile Payment Application through Self Organizing Maps | Time Series Modeling of Merger Series: An Autoregressive Process |
| | Gökhan AYDIN, Pakize YİĞİT, Şebnem BURNAZ | Jitendra KUMAR, Varun AGIWAL |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Wilmer PINEDA RIOS | Montasir OSMAN |
| SESSION 3 17:20- 18:40 | Exponentiality Test Based on Tsallis Divergence Between Equilibrium Distributions on Progressively Censored Data | Cointegrated Panel Data Model for Per Capita Gross Domestic Product and Electric Power Consumption of 11 South American Countries 1990-2014 |
| | Arezou HABIBIRAD, Vahideh AHRARI | Mayron Esteban GUTIERREZ, Heivar Yesid RODRIGUEZ |
| | Limiting Spectral Distribution of High Dimensional Hayashi's Estimator in Presence of Asynchronicity | Youth Status in Egyptian Labor Market |
| | Arnab CHAKRABARTI, Rituparna SEN | Said Ahmed SAID |
| | Limit Theorem for a Semi-Markovian Random Walk with Two Barriers | Spatio Temporal Bayesian Beta Model to Estimate the Forced Displacement Ratio by Municipalities in Colombia 2002-2016 |
| | Tahir KHANIYEV, Rovshan ALIYEV, Zulfiye HANALIOGLU | Natalia Niño MARTINEZ, Wilmer PINEDA-RIOS, Dagoberto Bermúdez RUBIO, Andrés CRUZ PEREZ |
| | Shiny App to Real Time Tweet Sentiment Classification Using Deep Learning | The Efficiency of Artificial Neural Networks in Continuous Data Modeling |
| Andrés Felipe RODRIGUEZ PEREZ, Wilmer PINEDA-RIOS | Montasir OSMAN | |

| | MUGLA HALL | BODRUM HALL | |
|-------------------------------------|--|--|--|
| 09:00-09:40 INVITED SESSION 1 | Alternatives to Least Squares Estimation in the Linear Regression Model Selahattin KAÇIRANLAR | | |
| 09:40-10:20 INVITED SESSION 2 | Dinamik Meta-Sezgisel Optimizasyon ve Uygulama Örnekleri Adil BAYKASOĞLU | | |
| 10:20-10:40 | COFFEE BREAK | | |
| Session Chair | Aynur İNCEKIRIK | Cağdaş Hakan ALADAĞ | |
| SESSION 1 10:40- 12:20 | A Research on Efficiency of Puplic Library Utilization in Turkey Rıdvan KESKİN, Aynur İNCEKIRIK | The Portfolio Optimization based on Genetic Algorithm Approach Vedide Rezan USLU, Azize Zehra ÇELENLİ BAŞARAN | |
| | A Comparison of Reliability Tests in Field Researches and An Application on Agricultural Data Ece UZUNSAKAL, Doğan YILDIZ | Robust Estimators in Portfolio Optimization Baki ÜNAL, Cağdas Hakan ALADAĞ | |
| | Copula Based Regression Models: An Application Rümeysa KARATAŞ, Uğur KARABEY | Multiple Regression Analysis based on ANN using with the different type of error function Asiye Zühal KÜÇÜKMUSTAFA, Mehmet Arif DEMİRCİ, Vedide Rezan USLU | |
| | On Tauberian Theorems for and Statistically Summable Sequences of Fuzzy Number Zerrin ÖNDER, İbrahim ÇANAK | Nonlinear Neural Network for Portfolio Optimization based on ARIMA İlgım YAMAN | |
| | Classification of The Companies Traded in The IMKB by Financial Ratios Burçin ÖNER, Yüksel ÖNER, Azize Zehra ÇELENLİ BAŞARAN, Çağlar SÖZEN | An Optimization Model for Warehouse Load Balancing to Cover High Seasonal Demand in Supply Chain Tufan BAYDEMİR | |
| | 12:20-14:00 | LUNCH TIME | |
| | 14:00-14:40 INVITED SESSION 3 | Using Concomitants or Order and Record Statistics in Directional Dependence Modeling Engin Sungur | |
| | 14:40-15:20 INVITED SESSION 4 | The Role of Academic Statisticians in USA Engin Sungur | |
| | 15:20-15:40 | COFFEE BREAK | |

| | MUGLA HALL | BODRUM HALL |
|-------------------------------------|--|---|
| Session Chair | Uğur KARABEY | A. Fırat ÖZDEMİR |
| SESSION 2 15:40- 17:00 | Calculating Value at Risk: Comparing of Different Estimators | Modeling Nonlinear Relation by Using Running Interval Smoother, LOWESS and Different Quantile Estimators |
| | Uğur KARABEY, Derya KARAGÖZ | Burak DİLBER, A. Fırat ÖZDEMİR, Gözde NAVRUZ |
| | Statistical Methods in Kansei Engineering: A Review | Estimation of Demand for Motor Own Damage Insurance with Neural Networks |
| | Burcu Devrim İÇTENBAŞ | Betül Zehra KARAGÜL, Samet GENÇGÖNÜL |
| | Statistical Methods in Optimization of Food Materials | Penalized L2E Estimator for Gaussian Mixture Model |
| | Alev Yüksel AYDAR | Fikriye KABAĞÇI, Yüksel ÖNER, Umashanger THAYSİVAM |
| | Extreme Value Theory on Valuation of Actuarial Risks | |
| | Bükre Yıldırım KÜLEKÇİ, Uğur KARABEY, A. Sevtap SELÇUK-KESTEL | |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Oğuz AKPOLAT | Femin YALÇIN |
| SESSION 3 17:20- 18:40 | Confidence Intervals and Region for Kumaraswamy Distribution Using Wang Pivotal Quantities under Progressive Censoring | Hybrid Model Approach for Feature Selection in Content Analysis |
| | Kadir KARAKAYA, İsmail KINACI, Coşkun KUŞ, Yunus AKDOĞAN | Adem DOĞANER, Cemil ÇOLAK |
| | Classification of Data by Using Machine Learning Methods | On Generalized Run Shock Models |
| | Bünyamin SARIBACAK, Erol TERZİ | Femin YALÇIN, Serkan ERYILMAZ, Ali Rıza BOZBULUT |
| | A Production Inventory Model with Phase-Type Service Time | Investigation of the Clustering Structure of Cities in Turkey with the Functional Clustering Analysis Using Some Climatic Variables |
| | Serife OZKAR, Umay UZUNOĞLU KOCER | Çağlar SÖZEN, Yüksel ÖNER, Hasan BULUT, Burçin ÖNER |
| | | |

| | MARMARİS HALL | KÖYCEĞİZ HALL |
|--|---|--|
| Session Chair | Aydın KARAKOCA | Nihal ATA TUTKUN |
| SESSION 1 10:40- 12:20 | Ensemble Learning Methods as In-silico Model for Prediction of Mutagenicity | Forecast Sales With Data Mining In The Public Sector |
| | Enis GÜMÜŞTAŞ, Ayça ÇAKMAK PEHLİVAN | Vildan YILDIRIM Ümit Cengiz UYSAL |
| | Multicomponent Stress-Strength Reliability Estimation for a Model Based on Bivariate Generalized Exponential Distribution | A Study On Predicting The Success of Bank Telemarketing Via WEKA Tools |
| | Mustafa NADAR | Nur Uylaş SATI |
| | Estimation of the Parameters of Two-Fold Weibull Model with Right Censored Reliability Data using Genetic Algorithm | Empirical Bayes Estimation of Proportion of Cause of Death Among Age Groups: Comparison of 2012 and 2017 |
| | Erkut TEKELİ , Güzin YÜKSEL | Esin AVCI |
| | A Comparative Study of Internal Validity Indices for Correlated Data with k-means Algorithm | Artificial Neural Network approach on Type II Regression Analysis |
| | Derya ALKIN, Aydın KARAKOCA | Berkalp TUNCA, Sinan SARAÇLI |
| | Comparison of optimal reinsurance studies on one side and on both sides | A Comparison of Stochastic and Deterministic Algorithms on a Ranking Task |
| Betül Zehra KARAGÜL, Murat BÜYÜKYAZICI | Engin TAS | |
| 12-20-14:00 | LUNCH TIME | |
| Session Chair | Ömer AKGÜLLER | Nuri Çelik |
| SESSION 2 15:40- 17:00 | Classification of Author's Education Degree by Using Three Learning Methods | Effect of Sample Size on Learning Performance in Sentiment Analysis |
| | Betül Kan KILINÇ , Yonca YAZIRLI | Adem DOĞANER |
| | The Analysis of Interlocking Directors via Hypergraphs: The Case of Turkey's Listed Companies | Simple Linear Regression Example of Uniform-Normal Distribution |
| | Ömer AKGÜLLER, Mehmet Ali BALCI | Nuri ÇELİK, Çiğdem TOPÇU GULOKSUZ |
| | Robust Estimators in Generalized Linear Models with Binary Responses in the case of Influential Points | Comparison of Classification Performances for Predictiveness Curve and ROC Curve |
| | Burcu MESTAV, Kumru URGANCI TEKİN, Neslihan İYİT | Merve TURKEGUN, Bahar TASDELEN, Selma UNAL, Damla Hazal SARAL |
| Cardinality Constrained Mean-Variance-Skewness Model by the Artificial Bee Colony Algorithm: An Application in Istanbul Stock Exchange | Spatio-Temporal Multivariate Imputation of Missing Air Quality Data | |
| İlhan USTA, Volkan YEŞİLDERE , Mustafa SARIKAVAK | Sebnem ER | |
| 17:00-17:20 | COFFEE BREAK | |

| | MARMARİS HALL | KÖYCEĞİZ HALL |
|--|--|--|
| Session Chair | Engin Yıldıztepe | Didem DERİCİ YILDIRIM |
| SESSION 3 17:20- 18:40 | Anomaly Detection in Multivariate Data using Histogram Based Outlier Score | Network Meta-Analysis for Decision Making in Medicine |
| | Beyza KIZILKAYA, Engin YILDIZTEPE | Didem DERİCİ YILDIRIM, Damla Hazal SARAL, Bahar TASDELEN |
| | A Study on the Robust Bootstrap Procedure by Hampel Weighting Function | A New Quantile Penalty Estimation |
| | Uğur BİNZAT, Engin YILDIZTEPE | Bahadır YÜZBAŞI, Yasin ASAR, Zühal KÜÇÜKARASLAN |
| | Converting Median and Interquartile to Mean and Standard Deviation in Meta-Analysis | A Simulation Study on Comparison of Alternative Cointegration Tests |
| | Esin AVCI | Aytaç PEKMEZCİ |
| | | |
| | FETHİYE HALL | DALYAN HALL |
| Session Chair | Özlem TÜRKER | Funda İŞÇİOĞLU |
| SESSION 1 10:40- 12:20 | Estimation of the System Reliability for Generalized Inverse Lindley Distribution under Different Sampling Schemes | Reliability Analysis of a Multi-state System with Identical Units Having Two Dependent Components |
| | Fatma Gul AKGUL Keming YU Birdal SENOGLU | Funda İŞÇİOĞLU, Aysegül EREM |
| | On the General Formula of Renewal Function of Erlang Distribution | Simulated Annealing Based Simulation Optimization Method for Resource Constrained Project Scheduling Problem with Multiple Modes |
| | Ozlem ARDIÇ Tahir KHANIYEV | Nurhayat ŞAHİNKAYA, Gokalp YILDIZ |
| | Examination of Accuracy in Fuzzy Rule-Based Classification and Decision Tree: An Application | Spatiotemporal Variability and Trends of Snowfall in Turkey |
| | Betül Kan KILINÇ, Yonca YAZIRLI | Hakki BALTACI, Bulent Oktay AKKOYUNLU, Mete TAYANÇ |
| | Searching the Differences Through the Tails of Distributions | An Investigation on the Efficiency of Research Universities in Turkey |
| | Gözde NAVRUZ, A. Fırat ÖZDEMİR | Aslı GÜNAY, Ebru YÜKSEL HALİLOĞLU |
| | Full Automatic Evaluation of Seismic Refraction Data | Managerial Efficiency and Stochastic Decision Support Systems in Universities |
| Orhan KESEMEN, Sibel ERTÜRK, Ebru NAZ | Ahmet KARA | |
| 12-20-14:00 | LUNCH TIME | |

| | FETHİYE HALL | DALYAN HALL |
|---|--|--|
| Session Chair | Selen ÇAKMAKYAPAN | Yunus AKDOĞAN |
| SESSION 2 15:40- 17:00 | Market Segmentation With Clustering Analysis –SSO Sample | Customer Satisfaction Survey for Old Age Asylum Using SERVQUEL |
| | Songül BULUT | Suna ÇETİN, Kezban BULUT, Hamza KUZUCU |
| | The Lindley- Frechet Distribution and Properties with Lifetime Data Application | Inference statistics of the Gini index for loss distributions |
| | Selen ÇAKMAKYAPAN, Gamze ÖZEL | Amina BARI, Abdelaziz RASSOUL |
| | Measuring Product Sustainability: A Risk-Focused Approach | Dynamic Measurement of Achieving Europe 2020 Goals in Smart Development – Turkey and |
| | Mürüvvet Deniz SEZER , Hasan SELİM | Andrzej SOKOLOWSKI, Małgorzata MARKOWSKA |
| | Effect of Prior Specification and Sample Size on Posterior Inferences in a Mixed Linear Model in Animal Breeding: A Simulation Study | Likelihood Function for Zero-Inflated Models with FuzzyCounts |
| Ebru KAYA BAŞAR, Mehmet Ziya FIRAT | Godrick OKETCH Filiz KARAMAN | |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | Adnan KARAİBRAHİMOĞLU | Özlem KAYMAZ |
| SESSION 3 17:20- 18:40 | A Threshold Method for Financial Networks and Geometric Scattering of Agents | A comparison of various parameter estimation methods for mixtures of von Mises distributions |
| | Ömer AKGÜLLER | Muhammet Burak KILIÇ |
| | On the Renormalized Entropy-based Complexity Analysis by BIST100 Data | An Application of Frailty Models in Motor Insurance |
| | Adnan KARAİBRAHİMOĞLU, Özgür AFŞAR, Şenay ÇELİKÖRS | Uğur KARABEY, Nihal ATA TUTKUN |
| | Statistical Power in Meta-Analysis Using Correlation Coefficient as Effect Size | Generalized Linear Model with Hierarchical Likelihood for Sparse Estimation |
| | Burçin ÖNER, Bülent ÇELİK | Özlem KAYMAZ, Henry WODD, Arief GUSNATO |
| | A Bayesian Generalized Linear Model For Crimean- Congo Hemorrhagic Fever Incidences | |
| Devrim BİLGİLİ | | |
| | MİLAS HALL | DATÇA HALL |
| Session Chair | Murat GÜL | Liaquat AHMAD |
| SESSION 1 10:40- 12:20 | Measuring the Financial Performances of Companies Listed on BIST-Services Index with Data Envelopment Analysis | Joint Modeling of Longitudinal Response and Time-to-event Data using Bartlett Decomposition : A Bayesian Perspective |
| | Murat GÜL, Olcay KARAKUŞ | Arindom CHAKRABORTY Srimanti DATTA |
| | Evaluation of Development Agencies with Data Envelopment Analysis | Monitoring of Blood Parameters of Dogs and Cats Using Control Chart Techniques |
| Murat GÜL, Nagihan AKSU | Liaquat AHMAD, Sabqa KHAN, Muhammad ASLAM | |

| | MİLAS HALL | DATÇA HALL |
|--|---|--|
| Session Chair | Murat GÜL | Liaquat AHMAD |
| SESSION 1 10:40- 12:20 | An Intuitionistic Failure Mode and Effect Analysis Proposal to Determine the Most Important Corrective and Preventive Strategies | Generalized Ratio-Type Estimator for Population Variance Using Auxiliary Information in Simple Random Sampling |
| | Gülin Feryal CAN, Mehmet GÜLŞEN | Muhammad ISMAIL, Sumbal ZURWA, Nazia KANWAL |
| | Estimation of the Reliability in Multicomponent Stress-Strength Model based on Chen Lifetime Distribution Using Genetic Algorithm | Copula Density Estimation of Iranian Household's Income and Expenditure by using Model Selection Method |
| | Muhammet Burak KILIÇ, Gökhan GÖKDERE | Sedigheh SHAMS |
| | Analyzing Housing Market Dynamics using Linear and non-Parametric Models | Generalized Improved Exponential Estimator of Population Variance |
| Bilgi YILMAZ, Fatma YERLİKAYA-Özkurt, A. Sevtap SELÇUK KESTEL | Tanveer AKHLAQ, Muhammad ISMAIL, Aamir SANAULLAH | |
| 12-20-14:00 | LUNCH TIME | |
| Session Chair | Nesrin GÜLER | Monjed H. SAMUH |
| SESSION 2 15:40- 17:00 | Some Rank Formulas for Covariance Matrices of Mixed Effects under General Linear Mixed Models | Statistical Evaluation of DWT-PCA/SVD and FFT-PCA/SVD Face Recognition Algorithms under varying Head-poses |
| | Nesrin GÜLER | Louis Asiedu, E. N.N. Nortey, F. N. YIRENKYI |
| | Imputation strategies for dealing with missing values in multi-item scales | A Long Short-term Memory based MSM Model for Short-term Volatility Forecasting |
| | Oya KALAYCIOĞLU | Khalidoun KHASHANAH, Yue LI ; Chenjie SHAO |
| | Markovian Queueing System with Ordered Entry | DEA Methodology in Logistics Decision Support |
| | H. Okan IŞGÜDER, Neslihan AVCU | Agata MESJASZ-LECH, Marek SZAJT |
| | Stock Certificate Selection for Optimal Investment by Using Path Analysis and Multi Criteria Decision Making Methods | Practicality of Some Variations of Ranked Set Sampling |
| Özlem GÖCEN İrem KARABEBE | Monjed H. SAMUH, M. HAFİDZ Omar M. PEAR HOSSAIN | |
| 17:00-17:20 | COFFEE BREAK | |
| Session Chair | | |
| SESSION 3 17:20- 18:40 | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

6. ABSTRACTS OF INVITED SPEAKERS

A New Difference-based Weighted Mixed Liu Estimator in Partially Linear Models

Fikri Akdeniz

*Department of Mathematics and Computer Science, Çağ University,
33800 Tarsus-Mersin, Turkey*

ABSTRACT

We consider a generalized difference-based Liu estimator (GDLE) of the parameters in the partially linear model, $y = X\beta + f + \varepsilon$, when the errors are correlated. Under the linear stochastic constraint $r = R\beta + e$, a new weighted mixed Liu estimator (GDWMLE) is extended. Also, the performance of this estimator over the generalized difference-based weighted mixed estimator (GDWME) and the GDLE in terms of the mean square error matrix (MSEM) criterion is investigated. A method to estimate the biasing parameter d and non-stochastic weight ω is also considered. The efficiency properties of the new estimator is illustrated by a simulation study. Finally, the performance of the new estimator is evaluated for a real dataset.

Keywords: Difference-based estimator; Generalized Liu estimator; Generalized difference-based weighted mixed Liu estimator; Partially linear model; Weighted mixed estimator.

AMS Subject Classification: Primary: 62G08; Secondary: 62J07

First Passage Time of Degradation Process

Narayanaswamy Balakrishnan
McMaster University
Hamilton, Ontario
Canada

ABSTRACT

In this talk, I will begin by introducing some popular degradation processes, such as Wiener process and gamma process. After presenting their basic properties, I will discuss the derivation of first passage time distribution and its approximation. I will finally present some numerical results, both for parametric and nonparametric evaluations, to illustrate the methods developed.

DİNAMİK META-SEZGİSEL OPTİMİZASYON VE UYGULAMA ÖRNEKLERİ

Adil Baykasoğlu

Endüstri Mühendisliği, Dokuz Eylül Üniversitesi

ÖZET

Optimizasyon problemlerinde çoğunlukla problemin tanım kümesinin, parametrelerin, kısıtların ya da değişkenlerin önceden kesin olarak bilindiği ve sabit kaldığı varsayılmaktadır. Ancak gerçek hayatta karşılaşılan problemlerin önemli kısmında önceden kestirilemeyen olaylar problem yapısında değişikliklere sebep olabilmekte ve buna bağlı olarak önceden hazırlanmış bir planın baştan oluşturulması ya da hızlı bir şekilde revize edilmesi gerekmektedir. Sonuç olarak nerede, ne zaman ve hangi şiddette oluşacağı önceden kesin olarak bilinmeyen ya da sağlıklı bir şekilde tahmin edilmesi güç olan dinamik olaylar, genellikle birbirlerine bağımlı bir dizi yeni problemin oluşmasına neden olmaktadır. Bunun gibi zamana bağlı tanım kümesi ya da değişken parametreleri olan problemler, ilgili bilimsel yazında dinamik optimizasyon problemi olarak isimlendirilmektedir. Burada amaç, optimizasyon probleminin tek bir çözümünü bulmanın yanı sıra, olaylara ya da zamana bağlı olarak değişen eniyi çözümlerin hızlı şekilde takibini sağlayabilmektir. Dinamik optimizasyon problemlerinin klasik optimizasyon yaklaşımları ile ele alınması oldukça zor olduğundan bu alanda meta-sezgisel yöntemlerin geliştirilmesi ön plana çıkmıştır. Son yıllarda çok sayıda başarılı meta-sezgisel algoritma dinamik optimizasyon problemlerine başarıyla uyarlanmıştır. Özellikle sürekli dinamik optimizasyon problemlerinin çözümünde anlamlı bir literatür oluşmuştur. Ancak kesikli dinamik optimizasyon alanındaki uygulamalar oldukça nadirdir. Hala pek çok açıdan gelişime ve yeni fikirlere ihtiyaç duyulan bu konunun Endüstri 4.0 kavramının ön plana çıktığı modern üretim sistemleri alanında yeterince uygulanmadığı görülmektedir. Mevcut sunumda dinamik meta-sezgisel optimizasyon konusu çeşitli yönleri ile irdelenerek, konu ile ilgili yapmış olduğumuz çok sayıda çalışma ve uygulamalar özetlenecektir.

Robust Bayesian Relevance Vector Machines in Regression and Supervised Classification Using Information Complexity and the Genetic Algorithm

Hamparsum Bozdogan

The University of Tennessee, U.S.A.

E-mail: bozdogan@utk.edu

ABSTRACT

Support Vector Machines (SVMs) have been popular kernel methods in regression and classification applications. However, *SVMs* suffer from a number of limitations. In this talk, we propose a new and novel model selection in *Bayesian Relevance Vector Machines (BRVMs)* in regression and classification problems (Tipping, 2001). *BRVM* is a sparse kernel technique, which is an improvement over the *SVMs* from the Bayesian learning point of view, while avoiding the limitations that exist in *SVMs*.

Unresolved model selection issues in *BRVM* regression and classification include: *choosing the optimal form of the kernel function* among a portfolio of kernel choices for a substantive data set; *the parameters of the kernel function*; and *the subset selection of the best predictor variables* in regression and classification.

We introduce novel statistical modeling techniques based on the *information-theoretic measure of complexity* called *ICOMP criterion* developed by Bozdogan (1990, 1994, 2000-2018) as the fitness function hybridized with the *genetic algorithm (GA)* as our *optimizer* to perform the model selection. *ICOMP* allows the identification of the best fitting kernel function or functions among a large portfolio of kernel functions. It measures both the *lack-of-fit (LOF)* and the *complexity* of the *BRVM* models. The *genetic algorithm (GA)* enables the rapid computation of models that would otherwise be impossible in a reasonable amount of time for subset selection of best predictor variables for high-dimensional data.

We illustrate the advantages of this new approach on simulated and on real benchmark data sets in regression and classification problems including the classification of cardiac imaging of diseased aortic tissues for early detection of the cause of heart attack.

As a conclusion, we discuss how to robustify *BRVMs* using general distributional models along with smooth and flexible priors to enforce a stronger sparsity in the model to achieve further *Occam's Razor* in regression and classification problems.

Key Words: *Bayesian Relevance Vector Machines, Model Selection, ICOMP, and Genetic Algorithm.*

A New Multiple Model Approach based on Artificial Neural Network for Forecasting

Erol Eğrioğlu
Giresun University, Forecast
Research Laboratory, Giresun,
Turkey
E-mail:erol.egrioglu@giresun.edu.tr

ABSTRACT

In the literature, there are many of types forecasting methods. Each method has its own advantages and disadvantages. Although some methods can work well for a specific time series, other methods may not produce good forecasts for the same time series. Multiple model approaches use different models to obtain forecasts from a specific time series but this approach is not a model combination approach. It is possible to use different methods for each time point. Selecting the correct model among used methods for the specific time point is the main problem. In this study, pi-sigma artificial neural network based on artificial bee colony is used to select correct method in multiple model approach. In the multiple model approach, classical and soft computing forecasting methods are used. The proposed approach is applied for meteorological time series data sets. The strong and weak sides of the proposed method are explored.

Key Words: *Pi-Sigma Neural Network, Multiple Model Approach, Forecasting, Artificial Bee Colony Algorithm.*

Compound Sums

Serkan Eryilmaz

Atilim University, Turkey

E-mail: serkan.eryilmaz@atilim.edu.tr

ABSTRACT

The distribution of the compound sum $S = \sum_{t=1}^N Y_t$, where N is an integer-valued random variable and Y_1, Y_2, \dots is a sequence of random variables, has attracted great deal of attention in probability and statistics due to its wide applications including reliability, actuarial science, and risk management.

Various techniques have been proposed to study the distribution of S especially when N and Y_1, Y_2, \dots are independent.

In this talk, distributional properties of the random variable S will be presented when N is phase-type random variable that possibly depends on Y_1, Y_2, \dots .

The new results cover many cases in the literature since the random variable N and the sequence of random variables Y_1, Y_2, \dots are assumed to be dependent.

The extension of the results to the bivariate compound sums is also mentioned.

Key Words: *Compound sum, Dependence, Phase-type distributions, Probability generating function*

Bayesian Joint Estimation of Primary and Substitute Demands wit Lost Sales

Ülkü Gürler

Bilkent University

E-mail: ulku@bilkent.edu.tr

Burcu Tekin

Aselsan

E-mail: burcutekin@bilkent.edu.tr

Emre Berk

Bilkent University

E-mail: eberk@bilkent.edu.tr

ABSTRACT

In this study, we consider the joint estimation of the demand arrival rate, primary and substitute demand rates in a lost sales environment with two products, using a Bayesian methodology. It is assumed that demand arrivals for the products follow a Poisson process where the unknown arrival rate is a random variable. An arriving customer requests one of the two products with a certain probability and may substitute the other product if the first choice (primary demand) is not available. The strength of our work lies both in (i) its generality regarding the simultaneous estimation of all relevant demand parameters in a dynamic fashion and (ii) its generality regarding the prior distribution choice of the related random variables.

Key Words: *Bayes Estimation, Inventory, Substitutable Products*

Alternatives to Least Squares Estimation in the Linear Regression Model

Selahattin Kaçiranlar

Department of Statistics, Faculty of Science and Letters, Çukurova University,
Adana 01330, Turkey

skacir@cu.edu.tr

ABSTRACT

The multiple linear regression model, least squares method and the optimality of the ordinary least squares estimator (OLSE) within the class of unbiased linear estimators are widely used in many fields of natural and social sciences. Inference under a linear regression model is carried out with respect to certain model assumptions. In practice, an analysis via linear models should make these assumptions as safe as possible, but clearly a complete certainty about each and every assumptions cannot be achieved.

The ordinary multiple linear regression model is described by the equation

$$y = X\beta + \varepsilon,$$

where fundamental assumptions are as follows:

- (i) X is an $n \times p$ model matrix of full column rank of observations on p non-stochastic explanatory variables ($\text{rank}(X) = p < n$),
- (ii) y is an $n \times 1$ vector of observations on the dependent variable,
- (iii) β is a $p \times 1$ vector of unknown parameters,
- (iv) ε is an $n \times 1$ vector of disturbances with expectation $E(\varepsilon) = 0$ and variance-covariance matrix $\text{Cov}(\varepsilon) = \sigma^2 I_n$.

The OLSE obtained by minimizing the error sum of squares is given by

$$\hat{\beta} = (X'X)^{-1} X'y.$$

The variance-covariance matrix for $\hat{\beta}$ is given by

$$\text{Cov}(\hat{\beta}) = \sigma^2 (X'X)^{-1}.$$

$\hat{\beta}$ is the best linear unbiased estimator (BLUE) for β under the above assumptions. When there are linear dependencies between the regressor variables, the problem of multicollinearity is said to exist. In the existence of multicollinearity, the method of least squares will generally produce poor estimates of parameters. The variance of the least squares estimates of the regression coefficients may be inflated, the absolute value of the least squares estimates will be too large and they will be unstable. Although the OLSE $\hat{\beta}$ is still the BLUE for β , with near multicollinearity this property is of little comfort.

In the sequel, a number of various estimators are introduced, such as restricted least squares, principal components, ridge, Liu, shrinkage, minimax, Bayes and robust estimators. The relationships between the estimators, risk comparisons, and the possible applications are widely discussed. There have been a number of interesting developments in the study of biased regression estimators. In this article, some of these methods have been considered.

Key Words: *Least Squares Estimation, Multicollinearity, Ridge Estimator, Liu Estimator, Principal Components Estimator.*

1. Akdeniz F. , Kaçiranlar S. (1995). On the almost unbiased generalized Liu estimator unbiased estimation of the bias and mse. *Communications in Statistics: Theory and Methods.* 24 (7): 1789-1797.
2. Gross J. (2003). *Linear regression.* Springer. New York.
3. Hoerl AE. , Kennard RW. (1970 a). Ridge regression : biased estimation for nonorthogonal problems. *Technometrics.* 12 : 55-67.
4. Hoerl AE. , Kennard RW. (1970 b). Ridge regression : application for nonorthogonal problems. *Technometrics.* 12 : 69-82.
5. Kaçiranlar S., Sakallıoğlu S. , Akdeniz, F. , Styan, GPH. ve Werner, HJ. (1999). A New biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland Cement . *Sankhya : The Indian Journal of Statistics, Vol.61, Series B, Pt.3,* 443-459.
6. Liu K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics: Theory and Methods.* 22 (2) , 393-402.
7. Liu K. (2003). Using Liu-type estimator to combat collinearity. *Communications in Statistics: Theory and Methods.* 32 (5), 1009-1020.
8. Liu K. (2004). More on the Liu-type estimator in linear regression. *Communications in Statistics: Theory and Methods.* 33 (11-125), 2723-2733.
9. Özkale MR. , Kaçiranlar S. (2007). The restricted and unrestricted two parameter estimators. *Commun. in Statistics: Theory and Methods.* 36 (13-16), 2707-2725.
10. Sakallıoğlu S. , Kaçiranlar S. , Akdeniz, F. (2001). Mean squared error comparisons of some biased regression estimators . *Communications in Statistics, Theory and Meth.* 30 (2), 347-361.
11. Theil H. , Goldberger AS.(1961). On pure and mixed estimation in economics. *International Economic Review.* 2, 65-77.
12. Trenkler,G.(1981). *Biased estimators in the linear regression model.* Cambridge, Massachusetts.
13. Vinod HD. and Ullah, A. (1981). *Recent advances in regression methods.* Marcel Dekker. New York.

Data Analysis in Health and Big Data

Gökhan Silahtaroglu

Istanbul Medipol University

gsilahtaroglu@medipol.edu.tr

ABSTRACT

For years, statistical analysis techniques have been successfully applied on data which are manually collected. Although it seems that statistical techniques used are more important at first, it is well known by scholars that amount, reliability and obtaining quality data have got much more importance. In the current digital era, we are exposed to a lot data every day. These are mostly unstructured data such as chats in online forums by patients or their family members, questions to medical professionals in social media and their answers, search engine data related to medicine, diagnosis and treatment. When it comes to big data in health, it embraces all laboratory data, public health records kept by government, health insurance records, scientific research results and reports. Today, when we simply cluster the searches done via search engines like Google, Yandex or Yahoo, we can detect an acute infectious disease in a geographic area and also see how fast and which direction it is spreading to. What is more, records in health institutions, diagnosis and laboratory reports may be analysed through data mining methods like artificial intelligence and decision trees besides conventional statistical methods. In addition, merging these data with others collected on the Internet will probably attach another dimension to the analysis. Analysing these big data in afore mentioned ways may well contribute to preventive health care. It is quite possible to provide health professionals with a decision support system merely with analysing any kind of patients' records such as video, image or voice via machine learning algorithms. These data may easily be obtained as soon as they step in the health institution. Comparing the patient's data with thousands of others in seconds, the machine may present or offer alternatives of diagnosis and treatment with a known probability value. Also, all these may be used for scientific research and health education. On the other hand, patients' privacy should not be violated for the sake of these analyses and research. It should also be remembered that all these are not to keep the health professionals especially doctors out of the system totally or partially, but to serve them so that they could do their jobs better and with less flaws. Just imagine a doctor who drags and drops a patient's results on a program and sees all possible diagnoses and risks. Isn't it amazing? In this study, two mostly confused diseases have been analysed by artificial neural networks and decision tree algorithms in order to extract the hidden patterns behind these diseases and find out information for a right diagnosis. In the study, analyses which are done on Tensorflow platform using machine learning algorithms and repeated by data mining tools as well will be presented and results will be discussed.

Statistics and Industry: Opportunities, Responsibilities and Examples

Eleanor Stillman

University of Sheffield, UK
e.c.stillman@sheffield.ac.uk

ABSTRACT

Statistics has a long and fruitful tradition of interaction with industry. The area provides us with interesting problems, massive amounts of data and substantial funding streams. But along with these benefits, come challenges and responsibilities and further beyond lies the threat that if statisticians do not step up and offer solutions, then others will certainly do so. These ideas will be explored and some examples of rewarding industrially-driven projects given. In particular, I will discuss some ongoing work on a Bayesian approach to multi-objective optimization.

In many industrial situations, an essential task is to find settings of various input variables to achieve an optimal quality output, where 'quality' is a multivariate concept. Industrially this is often solved by a mixture of common sense, experience and acceptance of a 'good enough' solution. However, numerous statistical issues arise in even the simplest version of the problem. Often no formal modelling of the output-input relationship will be attempted; if there is modelling it will probably be through classical regression models for each response variable separately; there is likely to be a fundamental lack of appreciation that such models cannot simply be inverted to identify optimal inputs; and handling of the multiple objectives is likely to be relatively *ad hoc*, perhaps by reduction to a combined score or a primary/subsidiary specification. By adopting a Bayesian perspective, these difficulties can be overcome. The asymmetries inhibiting inversion are made explicit and a full modelling of the joint behaviour of inputs and outputs allows subsequent interrogation to identify relevant conditional, or even predictive, distributions. Extensions to multivariate inputs and outputs is natural and incorporation of (often substantial) practitioner experience can be handled in an explicit manner. The ideas will be illustrated with an example from 3D printing (additive manufacturing).

Key Words: *Bayesian Multi-Objective Optimization, Additive Manufacturing,*

An Overview of Methods for Social Network Monitoring

William H. Woodall
Department of Statistics
Virginia Tech
Blacksburg, VA 24061-0439
bwoodall@vt.edu

ABSTRACT

In this paper I will give an introductory overview of social networks and some of the statistical methods for detecting anomalies over time. Discussion will include the advantages and limitations of various methods and some relevant practical issues such as the aggregation of data over time. Relationships will be given between network monitoring methods and related monitoring methods in engineering statistics and public health surveillance. Results on the performance of a popular scan method for monitoring social networks will be briefly summarized. Work in this area is encouraged and a number of research ideas will be given.

Key Words: Degree-corrected stochastic block model; Scan methods; Social network analysis; Social network change detection; Statistical process monitoring.

REFERENCES:

- Woodall, W. H., Zhao, M., Paynabar, K., Sparks, R., and Wilson, J. D. (2017). "An Overview and Perspective on Social Network Monitoring", *IISE Transactions*, 49(3), 354-365.
- Yu, L, Woodall, W. H., and Tsui, K.-L. (2018). "Detecting Node Propensity Changes in the Degree-corrected Stochastic Block Model", To appear in *Social Networks*.

A Stochastic Fuzzy Time Series Forecasting Model

Cagdas Hakan Aladag

Department of Statistics, Faculty of Science, Hacettepe University, Turkey

aladag@hacettepe.edu.tr

ABSTRACT

J.A. Paulos, who is a professor of mathematics, said: "Uncertainty is the only certainty...". Almost every real world data has some uncertainty associated with it. Fuzzy time series method was proposed as an alternative method for conventional time series methods to analyze real world time series including uncertainty. There are various fuzzy time series forecasting models in the literature, and they aim to reach better forecasting results. A few of these models utilize probabilities. It would be wise to use probabilities in order to strengthen the fuzzy time series method. In other words, it can be a good way to use much more knowledge obtained from probabilities. In this study, a new seasonal fuzzy time series forecasting model based on Markov chain transition matrix is introduced. In this forecasting approach, fuzzy inference process is performed by using transition probabilities. Therefore, fuzzy time series approach introduced in this study can be considered as a stochastic seasonal fuzzy time series method. The method is used to model two real-world time series in the implementation part of the study. As a result of the application, it is observed that the forecasting method produces accurate forecasting results for both time series.

Key Words: Forecasting, Markov chain, Fuzzy time series, Seasonality, Transition matrix

Using Concomitants or Order and Record Statistics in Directional Dependence Modeling

Engin A Sungur

University of Minnesota Morris
600 E. 4th Street
Morris, MN 56267 USA
E-mail: sungurea@umn.edu

ABSTRACT

Understanding and modeling dependence structures depending upon the direction are challenging but an interest of theoretical and applied researchers. In this article, we introduce a way of looking at directional dependence by using concomitants of order and record statistics in two-dimensional case. This approach allows us to measure the strength of direction of dependence and provides tools for deeper understanding of the hidden information in data.

Within this framework, we will introduce distributional and dependence cycles and present their use in various set ups such as the gold prices in Chinese, Japanese, European, and United States of America markets.

Key Words: Concomitants, Correlation, Dependence cycle, Directional dependence, Distributional cycle, Order statistics, Record statistics

7. ABSTRACTS OF ALL PARTICIPANTS

A Bayesian Generalized Linear Model for Crimean–Congo Hemorrhagic Fever Incidents

Devrim Bilgili

ABSTRACT

The global spread of the Crimean–Congo hemorrhagic fever (CCHF) is a fatal viral infection disease found in parts of Africa, Asia, Eastern Europe and the Middle East, with a fatality rate of up to 30%. A timely prediction of the prevalence of CCHF incidents is highly desirable, while CCHF incidents often exhibit nonlinearity in both temporal and spatial features. However, the modelling of discrete incidents is not trivial. Moreover, the CCHF incidents are monthly observed over a long period and take a nonlinear pattern over a region at each time point. Hence, the estimation and the data assimilation for incidents require extensive computations. In this paper, using the data augmentation with latent variables, we propose to utilize a dynamically weighted particle filter to take advantage of its population controlling feature in data assimilation. We apply our approach in an analysis of monthly CCHF incidents data collected in Turkey between 2004 and 2012. The results indicate that CCHF incidents are higher at Northern Central Turkey during summer and that some beforehand interventions to stop the propagation are recommendable.

A Comparative Application of Multiple Correspondence and Joint Correspondence Analysis: A Case Study of Turkey

Elif Çoker
Mimar Sinan Fine Arts
University,
Faculty of Science and Letters,
Department of Statistics,
İstanbul, Turkey
elif.coker@msgsu.edu.tr

ABSTRACT

Correspondence Analysis is used as an exploratory multivariate technique when the data set is expressed as tables which are contained of frequencies or counts. Correspondence analysis both gives the numerical and mostly the graphical analysis of these type of data. It lets us understand and interpret the relationship between the variables which are placed in a contingency table. Simple correspondence analysis explores the relationship between two variables in a contingency table. But in practice, we are confronted with more than two variables in most cases. In this situation, multiple correspondence analysis is used to examine the relationship between these variables. Joint correspondence analysis is proposed as an improvement of the multiple correspondence analysis. The data set used in the application is taken from the latest available survey of "Life Satisfaction of Turkey" which was accomplished by the Turkish Statistical Institute in 2016. The aim of this study is the investigation of the general job satisfaction of Turkey through using alternative correspondence analysis methods such as multiple correspondence analysis and joint correspondence analysis.

Key Words: *Multiple Correspondence Analysis, Joint Correspondence Analysis, Job Satisfaction*

A Comparative Study of Internal Validity Indices for Correlated Data with K-means Algorithm

Derya Alkın

Necmettin Erbakan University,
Institute of Science,
Department of Statistics
derkonizm@gmail.com

Aydın Karakoca

Necmettin Erbakan University,
Faculty of Science,
Department of Statistics
akarakoca@konya.edu.tr

ABSTRACT

Clustering analysis is applied in many different areas. Despite having many application areas, many studies continue on determining the number of clusters. The number of clusters can be vary depend on the selected method for same data set. The possibility of forming different clusters for the same dataset poses the problem of selecting the number of clusters. Validation of the clustering results constitutes the most important part of the clustering studies. Internal validity indices are the most used approach for cluster validation.

In this study we compared the performance of 27 internal validity indices according to different level of relationships among the variables. An extensive simulation study is conducted using the k-means method to put forth the indices performance. The results show in depth the differences in performance among the indices.

Key Words: *K-means, Clustering, Internal Validity Indices*

A Comparative Study on the Robust Estimators for the Parameters of the General Linear Regression Model via Simulation

Hakan Savaş Sazak

Ege University,
Faculty of Science,
Department of Statistics,
İzmir, Turkey
hakan.savas.sazak@ege.edu.tr

Nalan Mutlu

Ege University,
Faculty of Science,
Department of Statistics,
İzmir, Turkey
nalanmutlu1992@gmail.com

ABSTRACT

The general linear regression (GLR) model is one of the most popular statistical models used in several areas in science. Its popularity has always been a motivation for the scientists to find effective estimation methods for the parameters belonging to it. Despite its full efficiency under normal distribution, because of the sensitivity of the least squares (LS) method to outliers and deviations from the assumed assumptions and models, many robust methods have been proposed so far for the estimation of the parameters of interest as alternatives to the LS method.

In this study, we compared the performance of the most popular robust estimators which are the Huber and Tukey M-estimators, the S-estimators and the MM-estimators w.r.t. the LS estimators for the parameters of the GLR model via simulation by using our own programs written in Matlab. The results verified the high efficiency and robustness of the MM-estimators. The results also revealed the bias of the robust estimators of the variance of the error term in many situations while the LS estimator is always unbiased. Finally, we give illustrations using real life data sets.

Key Words: *General Linear Regression Model, Least Squares, M-estimators, MM-estimators, Robust*

A Comparison of Stochastic and Deterministic Algorithms on a Ranking Task

Engin Taş
Afyon Kocatepe University,
Department of Statistics,
Faculty of Science and Literature
engintas@aku.edu.tr

ABSTRACT

Training a classification or a regression model involves a minimization of an error function using a convenient optimization algorithm. There are several approaches exist in these type of learning problems, but much of the successful approaches can be categorized as gradient-based learning methods. In the batch learning, one iteration requires a complete pass through the entire dataset in order to compute the average or true gradient. On the other hand, stochastic (online) learning only works with one example at each iteration. This extremely reduces the computational cost of one iteration. Furthermore, carefully tuned stochastic algorithms converged to the optimal solution of the problem in one pass. As available datasets grow ever larger, stochastic gradient-based methods have a major advantage in large and redundant data sets. In this study, we try to create a framework for comparison of stochastic and deterministic learning algorithms with respect to prediction performance and convergence characteristics on a ranking task.

Key Words: *Ranking, Regression, Least Squares*

A Comparison of Various Parameter Estimation Methods for Mixtures of Von Mises Distributions

Muhammet Burak Kılıç
Department of Business Administration,
Mehmet Akif Ersoy University,
Burdur, Turkey
muhammet.b.kilic@gmail.com

ABSTRACT

In this research article, we consider five methods for estimating the parameters of the mixture of two von Mises distributions namely, the methods of the maximum likelihood estimation using the Newton-type optimization, the Expectation and Maximization, the proposed genetic algorithm, the method of the modified moment estimation and the method of the minimum distance estimation. These methods are investigated by Monte Carlo simulations and two real data examples from molecular biology and ecology areas. We also propose a novel approach for constructing the confidence interval of the parameters of the mixture of two von Mises distributions based on bootstrap methodology. This confidence interval is to identify the problem of a search space in genetic algorithm for circular data. Finally, the results of the simulations and two real data reveal that the method of the maximum likelihood estimation using the proposed genetic algorithm is an effective estimation method compared to the presented other methods.

Key Words: *Circular Data, Mixture of Two Von Mises Distributions, Maximum Likelihood Estimation, Genetic Algorithms*

A Comparison Study of Optimal Bandwidth Selection with Density Functional Estimation for Contaminated Data

Celal Aydın
Gazi University,
Department Of Statistics,
Ankara, Turkey
aydinc@gazi.edu.tr

Necla Gündüz
Gazi University,
Department Of Statistics,
Ankara, Turkey
ngunduz@gazi.edu.tr

ABSTRACT

The bandwidth selection is the most important part of the kernel density estimation. Various bandwidth selection methods have been proposed in past decades. In the nonparametric cases density functionals $\int \eta(u)f^2(u)du$ are very common in use for estimating location and scale parameter. The purpose of this study is mainly directed to evaluate the performance of optimal bandwidth selection based on the mean square error (*MSE*) of density functionals $\int \eta(u)f^2(u)du$ for contaminated data with outliers.

A simulation study is constructed for the random samples from normal distributions with different parameters and with different contamination levels. In addition the performance of bandwidth selections is compared in terms of *MSE* of density functionals $\int \eta(u)f^2(u)du$.

Key Words: *Bandwidth, Density Functionals, Kernel Smoothing, Contamination Data*

A Computer Application for a Decision Making Algorithm Which is Based on the Soft Set Theory

Gözde Yaylalı Umul
Muğla Sıtkı Koçman
Üniversitesi,
Fen Fakültesi,
Matematik Bölümü,
Muğla, Türkiye
gozdeyaylali@mu.edu.tr

Nazan Çakmak Polat
Muğla Sıtkı Koçman
Üniversitesi,
Fen Fakültesi,
Matematik Bölümü,
Muğla, Türkiye
ncakmak@mu.edu.tr

Bekir Tanay
Muğla Sıtkı Koçman
Üniversitesi
Fen Fakültesi
Matematik Bölümü,
Muğla, Türkiye
btanay@mu.edu.tr

ABSTRACT

Molodtsov introduced soft set theory in 1999 to deal with uncertainties in the real world problems. Now a day there are increasingly many studies about the soft set theory. The soft set relation and orderings on soft sets were introduced by Babitha and Sunil.

Moreover decision making has an important place in the real life. Zhang introduced the interval soft set and applied it as a decision making method to solve uncertainties. The difference between the soft interval and the interval soft set was analyzed. A Decision Making algorithm based on the soft intervals was introduced to solve real social problems. To apply this algorithm, we used tabular form of soft intervals. Moreover, the computer application of our algorithm is written to obtain the results faster by the computer and applied to the examples. According to this decision making method, priority of the ranking is very important. In this study, we showed that if chooses are same with a different order, choice object will be changed.

Key Words: *Decision Making, Soft Set, Computer Application*

A Dual Response Problem Based on Median and Median Absolute Deviation

Gözde Özçırpan
Ankara University,
Faculty of Science,
Mathematics Department,
Ankara, Turkey
gozcirpan@ankara.edu.tr

Ahmet Kocatürk
Gazi University,
Faculty of Science,
Statistics Department,
Ankara, Turkey
ahmetkocaturk@gazi.edu.tr

Elif Şen
TURKSTAT Ankara
Regional Office,
Economic Research
Group, Statistician
Ankara, Turkey
elif.sen@tuik.gov.tr

Özlem Türkşen
Ankara University,
Faculty of Science,
Statistics Department,
Ankara, Turkey
turksen@ankara.edu.tr

ABSTRACT

Dual response surface methodology is one of the most commonly used approach for analysis of the simultaneous processes in some areas, e.g. quality engineering. In many of the studies, it is aimed to obtain the target value of central tendency and the measure of spread at zero value for replicated response measured data set. The mean and standard deviation statistics are used as measures of central tendency and spread, respectively, in most of the dual response problems in the literature. However, if there are outliers in replicated measures, the mean statistic will not be suitable to use as a central measure. For this reason, it is thought that the use of different descriptive statistics can be more appropriate to convert the replicated response measured data set to a dual response problem.

In this study, median and absolute median deviation (MAD) statistics are preferred to use to describe the replicated response measures as a single central and spread quantities. The suggested approach is applied on a data set, which is about to increase printer quality, defined in the literature. Statistical polynomial modeling assumptions are checked for both median and MAD responses. It is seen from the results that normality assumption on errors is not satisfied for MAD response model in which the errors are distributed log-logistic distribution with three parameters. Therefore, the MAD response model parameters are estimated by using maximum likelihood estimation method. In this study, Genetic Algorithm is used as an optimization tool to maximize the nonlinear likelihood function of MAD predicted response. When the performances of predicted median and MAD models are compared with the performances of mean and standard deviation predicted models, it can be said that the prediction performances of the models formed by median and MAD are better than the mean and standard deviation predicted models for dual response problem.

Key Words: *Dual Response Problem, Median, Median Absolute Deviation, Genetic Algorithm*

A Load Exchange Mechanism for Less-than-Truckload Operations

Başak Altan
Ozyegin University,
Istanbul
basak.altan@ozyegin.edu.tr

Okan Örsan Özener
Ozyegin University,
Istanbul
orsan.ozener@ozyegin.edu.tr

ABSTRACT

Faced with increased pressure to operated more efficiently, companies in trucking industry are seeking opportunities to increase their operational efficiencies. One such means is to collaborate with other companies in the market, which is classified as horizontal collaboration in the literature, and exchange loads to maximize the utilization of the trucks, hence the profitability of their overall operations.

In this paper, we study such a collaborative setting and propose a solution methodology that first maximizes the overall benefit from such a collaboration and second provides incentives for each participant of the collaboration. A key point in such a collaborative setting is the trust issues between participants, hence we devise a mechanism requires minimal information transfer between participants. A detailed computational analysis shows that our proposed collaborative solution provides benefits to the participants significantly higher than their non-collaborative solutions.

Key Words: *Less-than-truckload Shipping, Non-cooperative Game Theory, Exchange Mechanism*

A Macro Level Comparison of Hospital Efficiency in Turkey by Using Malmquist Total Factor Productivity Index

Ceren Dirik Serap Şahin
Kırıkkale University Kırıkkale University
cerendirik@kku.edu.tr serapsahin@kku.edu.tr

ABSTRACT

Due to rapidly changing dynamic structure of healthcare industry and increasing cost of healthcare services, analysis of hospital efficiency is becoming increasingly paramount. Efficient functioning of hospitals, which are the most important elements of the health system, is an inevitable necessity to increase the quality of health services. In Turkey, especially since 2003, health reform efforts have been accelerated. A number of reasons, including observing the results of these reforms, planning for scarce resources and focusing on the potential improvement areas necessitate the evaluation of hospital performance. We aimed to carry out efficiency analysis of public, private and university hospitals on the basis of 81 cities in Turkey for the period 2012-2016. In this macro level study, data envelopment analysis (DEA) is used to measure the hospital efficiencies of cities. Moreover, Malmquist Total Factor Productivity (TFP) Index, which is a DEA-based approach, is applied to monitor the change in efficiency for the relevant period and to determine the source of the change. The data used in the study is obtained from the annual health statistics published by the Ministry of Health. Input-oriented model is selected, taking into account that it is relatively difficult to intervene in the hospitals' outputs and the potential improvement suggestions about the outputs may cause some ethical problems. The input and output variables are decided as a result of a comprehensive literature review. In this direction, the number of specialists, number of practitioners, number of nurses, number of other personnel and number of beds are chosen as inputs while number of surgical operations, number of inpatients, bed occupancy rate, average length of stay and bed turnover rate are determined as outputs.

Under the constant return to scale assumption, Gaziantep, Kayseri, Bayburt, Kilis and Osmaniye are found efficient in all periods. Based on the Malmquist TFP Index, it is determined that the total factor productivity has decreased by 1% on average because of the failure to achieve technological change. Despite this technological decline, there is a technical efficiency change increase of 0.8% which emerges from the rise in pure technical efficiency change, suggesting an efficient usage of inputs. For the 5-year period, the cities with the greatest improvement in productivity are Kırıkkale (8.1%), Bitlis (7.7%) and Siirt (7.7%) and the cities with highest decline in productivity are Kars and Şırnak. It is also a finding that the highest productivity change is took place between 2013 and 2014.

Key Words: *Efficiency, Data Envelopment Analysis, Malmquist Total Factor Productivity Index*

A Model on the Total Number of Defective Items

J.B.G Frenk

Sabancı University,
Faculty of Engineering and
Natural Sciences, Istanbul
frenk@sabanciuniv.edu

Ayda Amniattalab

Sabancı University,
Faculty of Engineering and
Natural Sciences, Istanbul
ayda@sabanciuniv.edu

ABSTRACT

Based on the practical problem of estimating the total number of spare parts needed to repair defective products we propose in this talk a model describing the total number of defective items under warranty returned to a manufacturer over time. This model represents the sales process by means of a general point process (continuous time version) or time series (discrete time version) and it also describes the total number of repairs applied to an arbitrary item of this product under warranty. Combining these two stochastic processes yields an exact representation of the stochastic process counting the total number of defective items to be repaired by the manufacturer. Within the sub-model describing the total number of repairs, we also consider different repair strategies which can be used by the manufacturer. Due to the detailed description of both the sales process and the repair process of a particular item and the used repair strategy we are able to derive all kinds of properties of the stochastic process counting the number of defective items over time. In particular, detailed expressions of moments, variances and covariances can be obtained under the simplifying assumption that the sales process is represented by a non-homogeneous Poisson process. Also, for the discrete time version of the model in which the sales process is represented by a time series we will derive some of these expressions. Although this talk mainly deals with the theory of stochastic point processes and the analysis of general counting processes the developed theoretical model and its derived properties will then be used as a parametric model to fit a large data set. The statistical analysis of the proposed model will be discussed in a subsequent talk. As such the approach serves as an alternative to the approach of using time series in case item tracking information is available.

Key Words: *Stochastic Point Processes, Time Series, Repair Strategies*

A New Approximation for Fuzzy Directional Clustering(FDC)

| | | | |
|---|---|---|--|
| <u>Orhan Kesemen</u> Department of Statistics and Computer Sciences, Karadeniz Technical University, 61080, Trabzon, Turkey okesemen@gmail.com | Özge Tezel Department of Statistics and Computer Sciences, Karadeniz Technical University, 61080, Trabzon, Turkey ozge_tzl@hotmail.com | Eda Özkul Department of Statistics and Computer Sciences, Karadeniz Technical University, 61080, Trabzon, Turkey eda.ozkul.gs@gmail.com | Buğra Kaan Tiryaki Department of Statistics and Computer Sciences, Karadeniz Technical University, 61080, Trabzon, Turkey bugrakaantiryaki@gmail.com |
|---|---|---|--|

ABSTRACT

Cluster analysis is a useful tool used commonly in data analysis. The purpose of cluster analysis is to separate data sets into subsets according to their similarities and dissimilarities. Generally, statistical data analysis is used for data on the linear axis. However, classical statistical methods used for these data cannot be applied inherently to directional data. This is because; directional data have a modular structure. Generally, angular-based data is called directional data. The directions of the winds; the directions of migrating birds or animals; the orientation of objects in the plane can be held up as directional data.

In the literature, the FCD and FCM4DD algorithms are proposed to cluster directional data. This study aims to improve a new directional clustering algorithm which is based on trigonometric approximation. The trigonometric approximation is used for both descriptive statistics and clustering of directional data. In this study, some existing clustering algorithms (FCD and FCM4DD) and the proposed method were applied on various numerical and real examples, and their results were compared. As a result, these comparisons show that the superiority of the FDC algorithm in terms of consistency, accuracy and computational time. Fuzzy clustering algorithms for directional data (FDC, FCM4DD and FCD) were compared according to membership functions and the FDC and FCM4DD algorithms obtained more acceptable results than the FCD algorithm. While membership degrees must be in the interval $[0,1]$, the membership degrees obtained by the FCD algorithm are out of this interval.

Key Words: *Directional Data, Fuzzy Clustering, Trigonometric Mean, Angular Distance*

A New Extended Lindley Distribution

Mustafa Çağatay Korkmaz
Artvin Çoruh University,
Department of Measurement and
Evaluation, Artvin, Turkey
mcagatay@artvin.edu.tr
mustafacagataykorkmaz@gmail.com

ABSTRACT

Lindley distribution has been applied for life time data and reliability studies in the literature. Its probability density function is given by (for $x > 0$ and $\beta > 0$)

$$f(x; \beta) = \frac{\beta^2}{1 + \beta} (x + 1) e^{-\beta x},$$

where $\beta > 0$ is the scale parameter of the distribution. In fact, it is a mixture distribution of exponential and gamma distributions with mixing proportions $\frac{\beta}{\beta+1}$ and $\frac{1}{\beta+1}$ respectively. Its probability density function is decreasing on through its support and has thin tail property. Its hazard rate function is an increasing function. Its ability of data modeling is bounded with monotonic increasing failure rate and decreasing density shape. So, these properties cannot be sufficient for data modeling.

With this aim, we introduce a new extended Lindley distribution with the following cumulative distribution function

$$F(x; \alpha, \beta, \lambda) = 1 - \frac{1 + \beta x^\alpha + \beta \exp[\beta x^\alpha (e^{-\lambda/x} - 1)]}{1 + \beta} e^{-\beta x^\alpha},$$

where $x > 0$, $\alpha, \lambda > 0$ are the shape parameters and $\beta > 0$ is the scale parameter of the distribution. Its some distributional properties such as density shapes, special sub models, hazard rate function, moments have been obtained. Different estimation procedures have been used to estimate the unknown parameters and their performances are compared using Monte Carlo simulations. The flexibility and importance of the proposed model are illustrated by a real data set.

Key Words: *Lindley Distribution, Extended Weibull Distribution, Generalized Power Lindley Distribution, Generalized Distribution*

A New Method for Measuring Service Quality: An Application in Public Transportation Sector

Ümit Kuvvetli
Dokuz Eylül University,
Izmir, Turkey
umit.kuvvetli@gmail.com

Ali Rıza Firuzan
Dokuz Eylül University,
Izmir, Turkey
aliriza.firuzan@deu.edu.tr

Atakan Gerger
Dokuz Eylül University,
Izmir, Turkey
atakangerger@gmail.com

ABSTRACT

The quality of public transportation services is one of the most important performance indicators of modern urban policies for both planning and implementation aspects. Service performance of public transportation has direct impact on the future policies of local governments. Therefore, all the big cities, especially the metropolitan areas, have to directly deal with transportation issues and related public feedback. On the other hand, as in most service industries, it is very difficult to measure and assess the quality of service in public transportation, due to the intangible aspects of the service and the subjective methods used in quality measurement. Moreover, in the public transport sector where the potential problems associated with service quality should be determined and solved quickly, the current methods are insufficient to meet this need of public transport sector. In this project, it is aimed to fill this gap and a statistical model that measure service quality by using smart card boarding data and allows to measure service quality in detail such as route, time interval, passenger type and so on has been accordingly developed.

The main purpose of this project is to develop a model measuring quality of service for rubber-wheeled urban public transport firms have smart card systems. The model uses smart card data which is an objective data source as opposed to the subjective methods commonly used nowadays to measure service quality. The model measures service quality based on quality dimensions such as comfort, information, passenger density in the bus, type of bus stop etc. The weights of the dimensions in the model have been determined by statistical analysis of the data from passenger surveys. The results obtained from this model allow various detailed analyses for passenger types, routes and regions both on a general perspective with weighted criteria and on specific service dimensions requested. It is thought that the model results will guide the political decisions to provide the development of urban public transport systems, ensure standard service quality level and help to provide rapid intervention in problematic areas. Additionally, the project will contribute to the sector by measuring and monitoring service passenger satisfaction and comparing service quality offered by different cities.

Within the scope of the project, five routes with different passenger densities in Izmir/Turkey was selected as an example and the service quality for each passenger for a week (total 349.359 boarding) was measured and the results obtained were analyzed.

Key Words: *Statistics, Analytic Hierarchy Process, Urban Public Transportation, Service Quality, Smart Card Boarding Data*

A New Quantile Penalty Estimation

Bahadır Yüzbaşı
Inonu University,
Malatya, Turkey
byzb@hotmail.com

Yasin Asar
Necmettin Erbakan
University,
Konya, Turkey
yasar@konya.edu.tr
yasinasar@hotmail.com

Zühal Küçükarslan Yüzbaşı
Fırat University,
Elazığ, Turkey
zuhal2387@yahoo.com.tr

ABSTRACT

The ordinary least squares (OLS) estimator is the best linear unbiased estimator (BLUE) under the Gaussian assumptions. On the other hand, the OLS has the following limitations: (1) it has larger variance when there exists the problem of multicollinearity among predictors, (2) it has lack of interpretation in the presence of large number of variables, (3) it does not make variable selection compare to penalized estimation methods such as the Lasso. In order to overcome these limitations, there are a number of studies in literature. Regularization is a method to overcome some problems related to the lack of efficiency of OLS. We suggest a new estimation to combat the above three limitations when the Gaussian assumptions is not valid by adding two penalty factor which are Lasso and Correlation Based Penalty (CP) in quantile regression models. Hence, we obtain its theoretical result and conduct a detailed Monte Carlo simulation study along with a real data example. Our numerical results show that the new estimator performs well compare to its competitors.

Key Words: *Correlation Based Penalty, L1 and L2 regularizations, Quantile Regression*

A Production Inventory Model with Phase-Type Service Time

Serife Özkar
Balıkesir University,
Department of International
Trade and Logistics
serife.ozkar@balikesir.edu.tr

Umay Uzunoğlu Koçer
Dokuz Eylül University,
Department of Statistics
umay.uzunoglu@deu.edu.tr

ABSTRACT

The production inventory model with positive service time is a popular in production and manufacture. The inventory model studied is governed by (s, S) policy. That is, the production is switched on when inventory level depletes to s . Then, the production process remains in the on mode until the inventory level reaches S . As soon as the inventory level reaches S , the process is switched off. On the other hand, the inventory level may drop to zero. In this case, a local purchase is made to prevent customer loss. A local purchase of one unit at a time is made at a higher cost. Also, it is assumed that supply of items is instantaneous in local purchase. There is only one server in the system. Customers arrive according to Markovian arrival process. The produced item requires a processing time before it is served to the customer. The processing time follows phase-type distribution.

In this study, the above mentioned production inventory model is structured as continuous time Markov chain. We obtain the steady-state probabilities of the system and ~~then~~ the performance measures. Moreover, we perform some numerical analysis for optimal control variables.

Keywords: *Production Inventory Model, Local Purchase, Markovian Arrival Process, Phase-type Distribution*

A Queuing-Inventory System with Interruption and Two-Demand Classes

Umay Uzunoğlu Koçer
Dokuz Eylül University
Department of Statistics
umay.uzunoglu@deu.edu.tr

Şerife Özkar
Balıkesir University
Department of International
Trade and Logistics
serife.ozkar@balikesir.edu.tr

ABSTRACT

Queueing systems with inventory have been challenging problem for the last decades. The queueing –inventory system is different from classical queueing systems in that customers are served from the inventory.

In this study, we examine a queueing-inventory system with two classes of customers. Demand from each customer class arrives according to an independent Poisson process with different rate. Service time for each customer class is also independent from each other and follows exponential distribution. Each customer in the queue requires one item for service from the inventory. A served customer leaves the system and inventory level decreases by one unit. The demands occur during stock out period is backordered. Continuous review (s,S) inventory policy is implemented. That is when inventory level reaches the level s , the production decision is made. When the inventory level reached S , then the production is switched off.

The system is formulated by five-dimensional continuous time Markov process. The steady-state probabilities are found by using Bright and Taylor algorithm and performance measures are defined. Finally some numerical analysis are presented.

Keywords: *Queueing-inventory System, Markov Process, (s,S) Inventory, Two-customer Classes*

A Shiny App for Subgroup Discovery Using Activity Region Finder (ARF)

| | | | |
|--|---|--|---|
| <u>Serpil Kılıç Depren</u> Dept. of Statistics, Yıldız Technical University, Davutpaşa, İstanbul, Turkey serkilig@yildiz.edu.tr | Praneeth Vepakomma Dept. of Computer Science, MIT, Cambridge, MA, USA praneeth.in@gmail.com | Birol Emir Statistical Research and Data Science Center, Pfizer Inc., New York, NY, USA birol.emir@pfizer.com | Javier Cabrera Dept. of Statistics and Cardiovascular Institute, Rutgers University, New Brunswick, NJ, USA Cabrera@stat.rutgers.edu |
|--|---|--|---|

ABSTRACT

One of the most important questions in any data analysis is to be able to discover subgroups of observations where the mean or the proportion of responses is significantly higher than in other subgroups or in a comparative population. For example in medicine, we are interested in finding out who benefit the most from a drug compared to competitive drugs. Standard data mining techniques like recursive partitioning (CART, C4.5) optimize a criterion for partitions of a single population, but they do not address situations with comparative populations. In this paper, we introduce a new algorithm for recursive partitioning multiple populations. The ARF algorithm (Amaratunga and Cabrera (2004), Alvir et al (2009)) was introduced to resolve this question by new robust criteria that are suitable for different types of response variable.

In this presentation, we will introduce a shiny application that allows us to run ARF in a very user friendly environment but at the same time keeping access to all important capabilities. We will illustrate the use of this application with an example from a drug trial.

Key Words: *Subgroups Discovery, ARF, Shiny, CART*

A Statistical Characterization of a Simulated Earthquake Fault Plane Area by Using Spatial Statistics

Özlem Türkşen
Ankara University,
Faculty of Science,
Statistics Department
Ankara, Turkey
turksen@ankara.edu.tr

ABSTRACT

The field of spatial statistics is formed by applying statistical analysis to spatial data. Understanding the basic characteristics of the spatial data is possible with the use of spatial descriptive statistics. The spatial descriptive statistics deals with observations recorded as a pair of coordinates which represents a point. The analysis of point distributions provides fundamental clues about the underlying spatial processes and relationships. In this study, geodetic points of a simulated earthquake fault plane area is analyzed by using spatial statistics. In order to understand the basic spatial structure of fault plane, summaries of spatial distribution are obtained. For this purpose, the spatial measures of central tendency and the spatial measures of dispersion of geodetic points are calculated over the operation region. The distributional pattern of geodetic points, which can be random, clustered, or dispersed, is analyzed through quadrat analysis, nearest neighbor and Ripley's K function. The obtained results are discussed for generating simulated geodetic points which have important roles on the surface displacements in earthquake prediction studies.

Key Words: *Spatial Descriptive Statistics, Spatial Pattern Analysis, Earthquake Fault Plane*

A Structural Equation Model for Success Measurement of Development Projects: The Case of ISTKA Projects

Seda Yanık Ozbay
Istanbul Technical University,
Industrial Engineering,
Maçka Kampüsü, 34367,
Istanbul, Turkey
sedayanik@itu.edu.tr

Meltem Yontar
Istanbul Technical University,
Industrial Engineering,
Maçka Kampüsü, 34367,
Istanbul, Turkey
yontarmeltem@yahoo.com

ABSTRACT

Recent studies on determination of success criteria and critical success factors have gained a large share in project management research. However, it is not possible to define a standard project success. It is also difficult to determine the success criteria and critical success factors to cover each type of project.

This study analyzes the results of the survey that aims to explore the success criteria of the projects financed by Istanbul Development Agency, one of 26 development agencies in Turkey, and to determine the project success as perceived by project team members. This research presents the development and investigation of the attributes of the success criteria and factors of development projects, as well as an analysis of the relationship between the criteria and factors.

By examining the possible relationship between success criteria and critical success factors, it is aimed to find the most important factors affecting project success. In order to achieve this goal, by employing the partial least squares structural equation modelling (PLS-SEM) technique, a model has been established that includes variables of training, project team, project design, risk, sponsor, monitoring and project success. It is aimed to test whether the links established with this model, i.e. hypotheses, have been verified in terms of the target group. The model was validated through a survey returning 148 completed questionnaires. According to results, project managers should strengthen project design, determine risks and take precautions accordingly and select proper project team thus improve project implementation as well as the chances for project success.

In the literature, no empirical study was conducted to examine the relationship between the performance and critical success factors of regional development projects. For that reason, this study is thought to provide noteworthy contributions by filling this gap in the literature.

Key Words: *Project Management, Success Criteria, Critical Success Factors, PLS-SEM*

A Study on the Robust Bootstrap Procedure by Hampel Weighting Function

Uğur Binzat

Dokuz Eylül University,
Department of Statistics,
Buca, Izmir, Turkey
ugur.binzat@deu.edu.tr

Engin Yıldıztepe

Dokuz Eylül University,
Department of Statistics,
Buca, Izmir, Turkey
engin.yildiztepe@deu.edu.tr

ABSTRACT

Resampling methods become more important in the last decades and bootstrap is the most well-known resampling method. Efron's bootstrap is based on giving equal selection probabilities to the observations in the sample. In the presence of the extreme values, bootstrap (re)samples may contain more extreme values than the original sample. Therefore, this yields unsatisfactory results while constructing confidence intervals or estimating bias and standard error of an estimator. Using robust methods with bootstrap may be a sensible approach here. But, discarding extreme values may not be appropriate in some cases. Alternatively, robustness can be provided in the bootstrap process by giving fewer probabilities to outlying observations rather than giving equal probabilities.

The weighted bootstrap technique (WBT) aims to protect bootstrapping process against a given number of extreme values. The WBT consists of two phases, the probabilities are determined in the first phase and it is followed by the bootstrap procedure based on the probabilities. In this study, a modified WBT is investigated. Hampel's weighting psi function is used to determine the weight belonged to each observation. Then these weights are transformed to the selection probabilities to control resampling probabilities. In the study, a simulation study is conducted to compare WBT's performance with ordinary bootstrap in the presence of extreme values, and the results are discussed.

Key Words: *Resampling, Weighted Bootstrap, Hampel's Weighting Function*

A Two-stages MCMC Approach to Tree-ring Dating in The Absence of Master Chronologies

Masoud Muhammed Hassan
University of Zakho, Iraq
Masoud.Hassan@uoz.edu.krd

ABSTRACT

Tree-ring dating involves matching undated tree-ring width sequence to dated 'master' chronologies and hence provides a calendar date estimate for the undated trees. However, master chronologies are not available for most locations around the world, and are at best only available for limited areas, thus many tree timbers remain undated.

In this paper, a hierarchical Bayesian modelling approach for matching tree-ring width sequences is introduced and demonstrated in a relatively small but computationally challenging exercise. The Bayesian inference and matching process via Markov chain Monte Carlo, MCMC, includes evaluating the likelihood of the model parameters at every possible date offset which allows the posterior distribution of the unknown date for the undated sequences to be estimated. Implementation involves two main tasks: generating dated pseudo-chronologies from a mechanistic forward model, such as VSLite, given dated climatic variables, and matching undated sequences to the generated pseudo-chronologies thus providing a posterior estimate of the match. These tasks have been implemented via a two-stage MCMC method, which enabled us to speed up our model implementation. Our focus is therefore on areas where historical climate records exist, but no master chronologies have been established. Although, this method would only work when we have climate data, it would be of great interest to know whether a mechanistic forward model as simple as VSLite has the potential to create tree-ring width chronologies from climate data in areas where we have undated tree-ring data but no master chronologies exist. Results show that the approach has successfully matched the undated sequences to the pseudo-master chronology and hence provided useful information about the most likely offsets.

Key Words: *Two-stages MCMC, Tree-ring Dating, Bayesian Inference, VSLite Model, Pseudo-chronologies*

Adaptive Principle Components Calibration in Survey Sampling

Shehzad Muhammad
Ahmed

Department of
Statistics,
Bahauddin Zakariya
University,
Multan, Pakistan
ahmad.shehzad@
bzu.edu.pk

Khurram Haris

Department of
Statistics, Bahauddin
Zakariya University,
Multan, Pakistan
hariskhurram2@gmail.
com

Sajjad Haider Bhatti

Department of
Statistics, Government
College University,
Faislabad, Pakistan
goodhaider@gmail.com

Muhammad Noor ul
Amin

Department of Statistics,
COMSATS Institute of
Information &
Technology,
Lahore, Pakistan
nooramin.stats@
gmail.com

ABSTRACT

An extension to the calibration estimation due to Deville and Sarndal (1992) was given by Cassel et al (2002) namely model-based calibration. Inefficient calibration estimates are found in presence of linearly correlated auxiliary variables in survey sampling. Calibration based on the principal components is discussed by Shehzad (2012) and Cardot et al (2017) when the number of auxiliary variables is large. In this paper, we consider the model-based calibration and emphasize on the case when the auxiliary variables are linearly correlated and at the same time, the error terms display some type of heteroscedasticity for a linear regression model. We used the Nadaraya–Watson estimates of variance function. We call it adaptive principal components calibration estimation. A simulation study is presented here to compare the mean square errors of the proposed estimator with principle component calibration for different level of heteroscedasticity. Our proposed adaptive principal components calibration estimator performed better than the existing principal components calibration estimator.

Key Words: *Multicollinearity, Heteroscedsticity, Calibration, Principal Components*

Adjusting of WABL Parameters for Decision Tree Algorithm with Fuzzy Numbers: A Case study

Suzan Kantarci Savaş
Econometrics Department,
Kırklareli University,
Kırklareli, Turkey
suzan.kantarci@klu.edu.tr

Efendi Nasibov
Computer Science Department,
Dokuz Eylül University, İzmir, Turkey.
Institute of Control Systems,
Azerbaijan National Academy of Sciences,
Baku, Azerbaijan
efendi.nasibov@deu.tr

ABSTRACT

In real life, words have an important manner for the communication. Yet, the words include imprecise information that can be used in statistical analysis. WABL-LR-FID3 algorithm helps to deal with the classification or decision-making problem for linguistic data sets based on LR (Left-Right) fuzzy numbers. In this study, it is aimed to observe the parameters' behavior on the results of WABL-LR-FID3 algorithm. This approach is an adaptation of fuzzy c-means and Fuzzy ID3 decision tree classification algorithm. Linguistic variables are defined by using triangular fuzzy numbers given as LR fuzzy numbers. And weighted averaging based on levels (WABL) method is used as a defuzzification method for each data in different parameters ($s=0.5$, $s=1.0$, $s=1.5$ and $s=2.0$). Then, fuzzy c-means algorithm is performed in order to handle the membership degrees for each variable given in each data set ($c=3$). Four datasets are used in the experimental study. Finally, Fuzzy ID3 algorithm is applied and reasoning is done. Also, adjusting of WABL defuzzification parameters that affect the results of WABL-LR-FID3 approach is theoretically investigated.

Key Words: *Fuzzy ID3 Algorithm, Fuzzy Classification, Linguistic Data, WABL Method*

Algorithms for Computing the Reliability of (r, s) -out-of- n Systems

Gökhan Gökdere
Department of Statistics,
Faculty of Science,
Fırat University,
Elazığ, Turkey
g.g.gokdere@gmail.com

Selahattin Güneş
Department of Software
Engineering,
Faculty of Technology,
Fırat University,
Elazığ, Turkey
selahattin.gunes.5@gmail.com

Duygu Tanay
Department of Statistics,
Faculty of Science,
Fırat University,
Elazığ, Turkey
duygutogtay@hotmail.com

ABSTRACT

In the binary context, a (r, s) -out-of- n system consists of n components such that each consisting of two dependent subcomponents (A_i, B_i) , $i = 1, 2, \dots, n$. The system functions if and only if at least r of n subcomponents A_1, A_2, \dots, A_n function and at least s of n subcomponents B_1, B_2, \dots, B_n function. The reliability and mean residual life functions of this system was first studied by Bairamov (2013). Later, Eryılmaz (2017) defined a linear consecutive- (k, k) -out-of- $n:F$ system having a general structure under the model when the components have two dependent subcomponents as in Bairamov (2013). A linear consecutive- (k, k) -out-of- $n:F$ system fails if and only if at least consecutive k of n subcomponents A_1, A_2, \dots, A_n fail and at least consecutive k of n subcomponents B_1, B_2, \dots, B_n fail. The aim of this study is to propose algorithms for to evaluate the reliability of the linear (r, s) -out-of- $n: F$ system, linear weighted (r, s) -out-of- $n: F$ system, linear consecutive (r, s) -out-of- $n: F$ system and linear consecutive weighted (r, s) -out-of- $n: F$ system, respectively. Algorithms are based on a logical approach and are obtained using the state table obtained by the number of components in the system.

Key Words: *Linear (r, s) -out-of- $n: F$ System, Linear Weighted (r, s) -out-of- $n: F$ system, Consecutive (r, s) -out-of- $n: F$ System, Consecutive Weighted (r, s) -out-of- $n: F$ System, System Reliability, Algorithm*

Analysis of penalized regression methods in a simple linear model on the high-dimensional data

Reza ARABI BELAGHI
Department of Statistics,
University of Tabriz, Tabriz, Iran
rezaarabi11@gmail.com

Zari FARHADI
Department of Statistics,
University of Tabriz, Tabriz, Iran
z.farhadi88@yahoo.com

Özlem GÜRÜNLÜ ALMA
Department of Statistics,
Muğla Sıtkı Koçman
Unv., Muğla, Turkey
ozlem.gurunlu@gmail.com

ABSTRACT

Shrinkage methods for linear regression were developed over the last ten years to reduce the weakness of ordinary least squares regression with respect to prediction accuracy. In this study, shrinkage methods were used to evaluate regression coefficients effectively for the high-dimensional multiple regression model, where there were fewer samples than predictors. We used three regulation methods based on penalized regression to select the appropriate model. Lasso, Ridge and Elastic Net have desirable features; they can simultaneously perform the regulation and selection of appropriate predictor variables and estimate their effects. Here, we compared the performance of three regular linear regression methods using cross-validation method to reach the optimal point. Prediction accuracy using the least squares error (MSE) was evaluated. Through conducting a simulation study and studying real data, we found that all three methods are capable to produce appropriate models. The Elastic Net has a better prediction accuracy than the rest. However, in the simulation study, the Elastic Net outperformed other two methods and showed a less value in terms of MSE.

Key Words: Shrinkage estimator, High dimension, Cross-validation, Ridge regression, Elastic Net.

An Alternative Testing Procedure for The Equality of Two Population Means Using Median Ranked Set Sampling

Meral Ebegil
Gazi University,
Department of Statistics
mdemirel@gazi.edu.tr

Yaprak Arzu Özdemir
Gazi University,
Department of Statistics
yaprak@gazi.edu.tr

Fikri Gökpınar
Gazi University,
Department of Statistics
fikri@gazi.edu.tr

ABSTRACT

Ranked set sampling (RSS) is a sampling technique when the measurement of sampling units is quite difficult or expensive in terms of cost, time, or other factors. However, a small set of units can easily be ranked according to the variable of interest, without requiring the actual measurement. Muttlak (1997) suggested the Median ranked set sampling (MRSS) to reduce the errors in ranking, and to increase the efficiency of estimator for unimodal symmetrical distributions. RSS and MRSS can also be used for statistical inference such as hypothesis testing and constructing confidence interval for population parameters.

Özdemir et al. (2017) investigated the hypothesis testing for the difference of means of two populations under RSS for normal distributions with unknown variances. In this study, we propose testing procedure based on MRSS for the equality of two normal population means. Since the theoretical distribution of any statistic based on MRSS is very complex and even impossible to obtain, we obtained the critical values of test statistic under homogenous variance assumption using Monte Carlo simulation. Simulation study shows that the proposed test is quite well.

Key Words: *Median Ranked Set Sampling, Hypothesis Test, Monte Carlo*

Application of Extreme Value Theory (EVT) to Inflation Rate Targeting: The Ghanaian case

Ezekiel Nii Noye
Nortey
Department of Statistics
and Actuarial Science,
College of Basic and
Applied Science,
University of Ghana,
Ghana
ennnortey@ug.edu.gh

Eric Ocran
Department of Statistics
and Actuarial Science,
College of Basic and
Applied Science,
University of Ghana,
Ghana
eocran@ashesi.edu.gh

Abeku Atta
Asare-Kumi
Department of
Statistics and
Actuarial Science,
College of Basic
and Applied
Science,
University of
Ghana, Ghana
aasare-kumi@
ashesi.edu.gh

Richard Minkah
Department of Statistics
and Actuarial Science,
College of Basic and
Applied Science,
University of Ghana,
Ghana
rminkah@ug.edu.gh

ABSTRACT

Inflation targeting has been practiced by many countries of which Ghana is no exception. Successive governments of Ghana have used inflation targeting as a tool in ensuring that the economic hardship on its citizenry is lessened. This provides these governments the bragging rights to score some political points. The aim of this study therefore is to investigate the tail behaviour of Ghana's inflation rates using the Generalized Pareto distribution via the Peak Over Threshold methodology. This enabled the estimation of lower as well as upper bound values of the rates of inflation and determined the exceedance probabilities for the very low inflation rates known as inflation targeting by the Bank of Ghana and also estimate very high inflation rates yet to be observed. The results were that the L-moment method of estimation yielded estimated of the parameters of the Generalized Pareto distribution of 12.237 for the scale parameter and 0.099 for the shape parameter with 95% confidence estimates (8.704, 17.362) and (- 0.180, 0.339) respectively for the scale and shape parameter. Furthermore, the probability of obtaining an inflation rate lesser than 8.39%, which is the lowest inflation rate so far to be recorded, is zero. Therefore, it is possible for Ghana to attain single digit inflation however, this rate cannot fall below a threshold of 8.39%.

Key Words: *Extreme Value Theory, Inflation, L-Moment, Generalised Pareto distribution, Peak Over Threshold*

An Application of Frailty Models in Motor Insurance

Uğur Karabey

Department of Actuarial Sciences,
Hacettepe University,
Ankara, Turkey
ukarabey@hacettepe.edu.tr

Nihal Ata Tutkun

Department of Statistics,
Hacettepe University,
Ankara, Turkey
nihalata@hacettepe.edu.tr

ABSTRACT

This study proposed a possible application of survival analysis: frailty models on motor insurance. Frailty provides a good way to model unobserved heterogeneity in survival analysis. We use a motor insurance data that belongs to a private Turkish insurance company. We employ age, sex, region and age of automobile as covariates in the model where we use non-claim discount level as survival time. The effects of covariates on survival time are analyzed.

Key Words: *Claim, Motor Insurance, Survival Models, Frailty*

An Application of Unrelated Parallel Machine Scheduling with Sequence-Dependent Setups

Merve Burcu Sarıkaya
Ozyegin University, Istanbul
merve.sarikaya@vestel.com.tr

Okan Örsan Özener
Ozyegin University, Istanbul
orsan.ozener@ozyegin.edu.tr

Ali Ekici
Ozyegin University, Istanbul
ali.ekici@ozyegin.edu.tr

ABSTRACT

In this study, we analyze the scheduling problem faced by a TV manufacturer. TV manufacturing is planned based on a make-to-order strategy, and the manufacturer utilizes multiple heterogeneous assembly lines to produce more than 100 different product groups and more than 3000 different models. Once the customer orders are received, production scheduling is performed at the beginning of each month, and the goal is to satisfy the demand on time as much as possible. Each order/job is processed on one of the compatible assembly lines, and the decision maker has to consider several factors including the job-assembly line compatibility, release and due dates of the jobs and workload balance among different assembly lines when forming the production schedule. The manufacturer wants to complete the orders as close as possible to the due date. Late deliveries result in a penalty cost due to contractual agreements, and early completion of an order results in high inventory holding cost because of the capital tied to stock and occupied warehouse space until delivery. The problem under consideration is a variant of the unrelated parallel machine scheduling problem where the objective is to minimize the total tardiness and earliness of the jobs. We propose a wide range of heuristics including (i) a sequential algorithm, and (ii) a random set partitioning approach. We test the heuristics on the real-life data and compare the solutions with the current practice. We observe that the proposed approaches not only outperform the current practice by more than 400% on average but also provide solutions with low optimality gaps.

Key Words: *Unrelated Parallel Machines; Sequence-Dependent Setups; Set Partitioning*

An Efficient Treatment for Best-Worst Method with Multiple Best and/or Worst Criteria

Hüseyin Koçak

Division of Quantitative
Methods, Department of
Business Administration
Pamukkale University,
20160, Denizli
hkocak@pau.edu.tr

Atalay Çağlar

Division of Operations
Research, Department of
Econometrics,
Pamukkale University,
20160, Denizli
acaglar@pau.edu.tr

Gülin Zeynep Öztaş

Division of Quantitative
Methods, Department of
Business Administration
Pamukkale University,
20160, Denizli
gzeynepa@pau.edu.tr

ABSTRACT

In the literature there are various multi-criteria decision making (MCDM) methods depending on different principles. However, having conflicting criteria in a decision problem requires easy to use, less computational and consistent methods that can be applied in all possible situations. Therefore, each existing method tries to obtain more reliable solutions, with less effort, for different kind of decision problems. Rezaei (2015, 2016) proposed a method called Best-Worst Method (BWM) that requires less pairwise comparisons and results in reliable weights compared to other MCDM methods, especially Analytic Hierarchy Process (AHP). Recently, Koçak *et. al* (2018) improved BWM by taking the advantage of Euclidean norm rather than using Chebyshev norm. Although Euclidean BWM provides less computation and more reliable results, none of those proposed BWM models can handle the problem that includes more than one best or worst criteria. For this reason, in this study, a model was developed for situations where there are more than one best and/or worst criteria. Moreover, in order to illustrate the efficiency of the modified method, we use a real world MCDM problem, which is the transport company selection, and compare the solutions with the existing methods in the literature (Zhang *et. al* (2017)).

Key Words: *Multi-criteria Decision Making Methods, Best-worst Method, Multiple Worst Criteria*

An Investigation on the Efficiency of Research Universities in Turkey

Aslı Günay

Social Sciences University
of Ankara,
Department of Economics,
Hükümet Meydanı No:2,
06050 Ulus, Altındağ, Ankara
asli.gunay@asbu.edu.tr

Ebru Yüksel Haliloğlu

TOBB University of Economics
and Technology,
Department of Management,
Söğütözü Mahallesi,
Söğütözü Cd. No:43, 06510
Çankaya, Ankara
e.yuksel@etu.edu.tr

ABSTRACT

This paper measures the research efficiency and productivity of officially announced 15 public research universities in Turkey over the period between 2013-2017. Data Envelopment Analysis is applied to assess the relative research efficiency of these universities, while Malmquist Total Factor Productivity Index is used to measure the total factor productivity change with respect to research inputs of universities. Therefore, the analyses made here differ from the other studies regarding this issue in terms of the extent and diversity of data set and the measurement of university research efficiency and productivity changes together. One other objective of this study is to determine the degree of impact of various external and/or internal critical factors like having technopark, medical faculty etc. on research activities in Turkish research universities. In this context, a Tobit model will be used. Hence, when the lack of studies within the literature both measuring the research efficiency and examining critical factors influencing research activities of Turkish research universities are taken into account, it is considered that this study can fill a gap in this area. It is thought that the results obtained will provide managerial information and act as a guide to public research university administrations in Turkey, in using their resources more effectively for their research activities.

Keywords: *Data Envelopment Analysis, Efficiency, Productivity, Research University*

An Optimization Model for Warehouse Load Balancing to Cover High Seasonal Demand in Supply Chain

Tufan Baydemir
Research and Development
Team Leader at OBASE,
İstanbul, Turkey
tufan.baydemir@obase.com

ABSTRACT

Supply chain is a system of organizations, people, information and resources involved moving a product or service from supplier to customer. Main purpose in supply chain is to increase overall profit while increasing efficiency. For an effective supply chain, it is important to predict the future demand in advance. On the other hand, it is much more important to make efficient shipment plan to satisfy the customer demand in real life.

It is known that demand characteristics may vary on time. During the season, there are sudden increases in demand especially on religious holidays or New Year's Eve. Companies generally tend to work overtime to cover this pick demand. Because demand is temporary and will not continue. Generally, companies think that hiring a new worker is more costly. On the other hand, working overtime increases the worker costs and reduces the satisfactory of workers. Depending on the increase in demand, it is also possible to lose demand. In particular, retailers have no chance to back order. This means that unsatisfied demand is sink. In the end, covering the pick demand by working overtime is a costly choice.

With this study, a two-tiered mathematical model proposed which will ensure a balanced distribution of the warehouse workload, especially during special periods when the demand is much higher than the warehouse picking capacity such as religious holidays and New Year's Eve.

Key Words: *Warehouse, Load Balance, Demand Planning, Optimization*

Analysis of PISA-2015 Performance of Turkish Students by Multilevel Structural Equation Modeling

Ertan Akgenç
Selçuk University,
Graduate School Of
Natural Sciences,
Campus, Konya, Türkiye
statistician.ertan@gmail.com

Nimet Yapıcı Pehlivan
Selçuk University,
Science Faculty,
Department of Statistics,
Campus, Konya, Türkiye
nimet@selcuk.edu.tr

ABSTRACT

Structural equation modeling (SEM) is a statistical modeling method, which is used to test hypotheses based on cause and effect. Today, it is observed that some of the obtained data are complex or hierarchical. In this case, a multi-level structural equation modeling (MSEM) have been used. The multi-level structural equation modeling analyzes the nested (hierarchical) data by dividing into within-group and between-groups components.

The aim of this study is to analyze the performance of Turkish students in Mathematics, Reading and Science, on the Program for International Student Assessment (PISA) 2015 using a multi-level structural equation modeling method. The population of the research consists of Turkish students who completed at least six years of formal education, whose ages between 15 and 16 as of the date of their participation. According to the PISA-2015 data, the educational performances of Turkish students are analyzed at the student level and school level by means of the Mplus package program.

Key Words: *Multilevel Structural Equation Modeling, PISA-2015, Mplus*

Analyzing Housing Market Dynamics Using Linear and non-Parametric Models

Bilgi Yilmaz
Institute of Applied
Mathematics, METU
ybilgi@metu.edu.tr

Fatma Yerlikaya-Özkurt
Department of
Industrial Engineering,
Atılım University
fatma.yerlikaya@atilim.edu.tr

A. Sevtap Selcuk-Kestel
Institute of Applied
Mathematics, METU
skestel@metu.edu.tr

ABSTRACT

This paper analyzes the dynamic effect of macroeconomic indicators, such as financial and commodity market indicators on national housing markets. Furthermore, it does not only focus on the impact of the variables but also introduces a variety of models that based on the generalized linear models and multivariate adaptive regression splines. The models help us to identify the macroeconomic drivers of housing markets. Since the US has an adequate housing market data, the empirical analysis within the paper focuses on the US national housing market. Namely, the illustration of the proposed models is done through the historical realizations of S&P/Case-Shiller National Home Price Index (HPI) and US macroeconomic indicators for the period 2000 to 2017. The empirical results indicate that although the efficiency of the models may be questionable for their variable selection and their accuracy, the directions of the indicators effect are consistent with the standard expectation, and a proposed model estimates the trend of the market successively.

Key Words: *Housing Market, Generalized Linear Models, Multivariate, Adaptive Regression Splines*

Anomaly Detection in Multivariate Data Using Histogram Based Outlier Score

Beyza Kızılkaya

Dokuz Eylul University,
Department of Statistics,
Buca, Izmir, Turkey

beyza.kizilkaya@ogr.deu.edu.tr

Engin Yıldıztepe

Dokuz Eylul University,
Department of Statistics,
Buca, Izmir, Turkey

engin.yildiztepe@deu.edu.tr

ABSTRACT

Anomaly detection (AD) covers a large spectrum of methods. These methods can be categorized into the following three approaches; Supervised AD, Semi-Supervised AD, and Unsupervised AD. Unsupervised Anomaly Detection (UAD) is the challenging and the most widely applicable anomaly detection approach because of its philosophy. UAD methods try to find anomalies in data sets without using labeled training data. UAD methods can be grouped into three main categories; statistical based, clustering based and nearest neighbor based. Nearest neighbor and clustering based methods are the very well-known algorithms and widely preferred in practical applications. Statistical based methods try to fit a statistical model to the data and use an inference technique to obtain anomalies. Parametric or non-parametric methods can be used to fit a statistical model. The non-parametric methods need fewer assumptions when compared to the parametric techniques. Histogram-based outlier score (HBOS) is one of the non-parametric statistical methods that can be used to assign an anomaly score to each instance. This method is also named as the frequency-based method. The HBOS method is based on traditional histogram construction. It can be applied to multivariate data and an overall anomaly score for each instance can be computed. But its performance highly depends on the technique that is used to determine bin-widths. In this study, we investigate HBOS algorithm and evaluate the effect of the techniques that are used to determine bin-widths. The receiver operator characteristic (ROC) and the area under the curve (AUC) were used for comparison. The results show that HBOs algorithm with dynamic bin width performs well to detect anomalies.

Key Words: *Anomaly, Anomaly Detection, Histogram Based Outlier Score*

Artificial Neural Network Approach on Type II Regression Analysis

Berkalp Tunca

Department of Statistics,
Institute of Science,
Afyon Kocatepe University,
Turkey
berkalp0606@gmail.com

Sinan Saraçlı

Department of Statistics,
Faculty of Science Literature,
Afyon Kocatepe University,
Turkey
ssaracli@aku.edu.tr

ABSTRACT

Artificial Neural Network (ANN) is developed by biologically inspiring from human beings brain cells and it is a computer program that mimetic the neural networks of human brain. One of the most important feature of neural network is being able to learn via the information given. it is a kind of program that learns and creates and explores new information like a human brain without requiring any help. In recent years, the superiority of thinking like a human being and find a solution in any situation let the Neural Networks to be preferred pretty much. Some of the common features of ANN can be listed as: estimation and forecasting, classification and clustering, image and voice recognition, failure analyzing and etc.

Regression analysis models the effects of independent variables on the dependent variable. The most common estimation technique in classical regression approach is also Ordinary Least Squares (OLS). One of the most important assumptions of OLS is that the error term occurs just because of the dependent variable and independent variables do not include any error. However in daily life it is possible for the independent variables to include any error. The alternative regression techniques at these situations are called as Type II regression and these models are called as measurement error models in general.

Considering this point, the purpose of this study is to estimate the parameters of Type II regression via ANN approach and to calculate Bisector Regression line, which is one of the most famous techniques of Type II regression analysis. With this purpose the data set which is simulated via a software at different sample sizes and distribution types are analyzed and the results are given in related tables and figures.

Key Words: *Artificial Neural Network, Type II Regression, Measurement Error Models*

Assembling of Structural Equation Model and Fuzzy Analytic Hierarchy Process Methods: An Application for Residence Selection

Özlem Akay

Department of Statistic,
Faculty of Arts and Sciences,
Çukurova University,
01380, Adana, TURKEY
oakay@cu.edu.tr

İlkay Altındağ

Department of Banking,
Faculty of Applied Sciences,
Necmettin Erbakan University,
42060, Konya, TURKEY
E-mail: ialtindag@konya.edu.tr

ABSTRACT

A selection is an important matter that is needed to imcomplete and uncertain data in decision making processes. Multiple criteria decision-making (MCDM) is considered as a complex decision-making (DM) tool involving both quantitative and qualitative factors. In recent years, several MCDM techniques and approaches have been suggested to choosing the optimal probable options. One of the most common methods for MCDM in fuzzy environments is the Fuzzy Analytic Hierarchy Process (FAHP). FAHP is adopted to measure the relative importance of attributes, which is converted into the deterministic weight vector by using the extent analysis technique. The aim of this paper is to develop a new composite model using Structural equation models (SEM) and FAHP techniques. SEM is a extensive statistical modeling tool that is used for analyzing multivariate data involving complex between directly and indirectly observed (latent) variables. SEM includes confirmatory factor analysis, path analysis, path modeling, and latent growth modeling. This method is widely used in many academic fields such as political science, economics, management, marketing, psychology, sociology, education and health. In this study, we performed an application for residence selection using integrated FAHP and SEM methods.

Key Words: *Structural Equation Modeling, Multi Criteria Decision Making, Fuzzy Analytic Hierarchy Process, Residence Selection*

Bayes Estimation of the Reliability Function and Hazard Rate of the Topp–Leone Distribution

İlhan Usta
Anadolu University,
Science Faculty,
Department of Statistics,
Eskişehir, 26470, Turkey
iusta@anadolu.edu.tr

Merve Akdede
Usak University,
Faculty of Science and Art,
Department of Statistics
Uşak, 64200, Turkey
merve.akdede@usak.edu.tr

ABSTRACT

The Topp–Leone (TL) distribution is a univariate continuous distribution with a finite support and is an alternative distribution to the beta and Kumaraswamy distributions which have the same support set $(0,1)$. However, the estimations of reliability characteristics, reliability function and hazard rate function, of the TL distribution have not yet received much attention in the literature. In this study, the Bayes estimators of the reliability characteristics of the TL distribution are derived under squared error (symmetric) and linex, general entropy, precautionary (asymmetric) loss functions for complete sample case. Based on a Monte Carlo study, the performances of the proposed Bayes estimators are also compared with the corresponding maximum likelihood estimators in terms of mean squared error (MSE) and Pitman nearness (PN) probability. The simulation results show that the Bayesian estimators using asymmetric loss functions show good performance in terms of MSE and PN for most of the considered cases.

Key Words: *Topp-Leone Distribution, Reliability Function, Hazard Rate, Bayes Estimation, Pitman Nearness Probability*

Bias Types in Genome-Wide Association Studies

| | | | |
|---|--|---|---|
| Didem DERICI YILDIRIM Mersin University Department of Biostatistics and Medical Informatics didemderici@hotmail. com | Havva Didem CELIKCAN Mersin University Department of Biostatistics and Medical Informatics didemcelikcan@gmail. com | Merve TURKEGUN Mersin University Department of Biostatistics and Medical Informatics merveturkegun@gmail. com | Bahar TASDELEN Mersin University Department of Biostatistics and Medical Informatics bahartasdelen@gmail. com |
|---|--|---|---|

ABSTRACT

Recently, ability to obtain genetic data in healthcare studies, genetic data analysis has an increasing trend. In particular, the number of genetic-based studies to determine etiology of the complex diseases has been increasing day by day. The methods used to identify the relationships between diseases and genetic structure, determine risk group have been improving as well. But sufficient and balanced sample size is very important to make the results available. Moreover, there are several sources of bias in genome-wide association studies except common bias types as information and selection bias. The most important bias type is ascertainment bias and the other types are Hardy-Weinberg disequilibrium, population stratification, latent case and super control and super case biases for genome-wide association studies, respectively. This study is aimed to introduce bias types in genome wide association studies, suggestions to deal with these biases. By this way the qualities of these studies increases and offer a perspective to the problem of bias.

Key Words: *Bias, Genome-Wide Association Studies, Genetic Data*

Bland Altman and Deming Regression Methods from Method Comparison Methods and An Application in Prostate Cancer*

Seyma YAŞAR
Inonu University Faculty of
Medicine Department of
Biostatistics and Medical
Informatics
Malatya Turkey
seyma.yasar@inonu.edu.tr

Bulut DURAL
Inonu University
Faculty of Medicine
Turgut Ozal Medicine
Center Department of
Urology
Malatya Turkey
bulutdural@hotmail.com

Leyla KARACA
Inonu University
Faculty of Medicine
Turgut Ozal Medicine
Center Department of
Radiology
Malatya Turkey
karacaleyla4@gmail.com

Saim YOLOĞLU
Inonu University
Faculty of Medicine
Department of
Biostatistics and
Medical Informatics
Malatya Turkey
saim.yologlu@inonu
.edu.tr

ABSTRACT

In the method comparison studies, the degree of agreement between measurements obtained by different methods on the same subject is wondered. Bland-Altman and Deming Regression are methods developed for method comparison studies. There is a growing need for imaging methods that will clinically identify prostate cancer that is the most common type of cancer after skin cancer in men, avoid unnecessary biopsies, and provide active follow-up if necessary. In this study, "Computed Tomography(CT)" and "Magnetic Resonance(MR)" are that two of the imaging methods used to identify prostate cancer was examined by Bland-Altman analysis and Deming Regression method. In this study, the agreement between the CT and MR results of 43 patients diagnosed with prostate cancer diagnosed at the Inonu University Turgut Ozal Medicine Center Department of Urology was determined by both methods comparison techniques. When the confidence intervals obtained from the Deming Regression method were examined, neither systematic nor proportional bias was observed between CT and MR outcomes. Therefore, it may be preferable to use the Bland-Altman analysis to determine the agreement between measurements. In this case, according to the Bland-Altman statistical method, the results obtained by the MR may be found to be as large as 91,17 cc and to be as small as 47,04 cc from the results obtained from the CT. Considering these results, it would be sufficient to use only one imaging technique for patients who are new to the hospital and suspected of having cancer or who have prostate cancer and need to follow up.

Key Words: *Bland-Altman methods, Deming regression, method comparison, prostate cancers*

* This study was supported by Inonu University Research Found (BAP: TYL-2017-790).

Bus Driver Scheduling in Interurban Transportation System

Ali Rıza Güller Mehmet Ulaş Çakır Gizem Erol
STM, Ankara, Türkiye STM, Ankara, Türkiye STM, Ankara, Türkiye
aguller@stm.com.tr ucakir@stm.com.tr gizem.erol@stm.com.tr

ABSTRACT

Scheduling problem is one of the major topics in operational research that interests finding optimum resource allocation given time period by taking constraints into account. These resources can be machines, nurses, doctors, bus drivers, pilots while they are assigned to the jobs, rooms, trips, flights; respectively. Bus driver scheduling optimization is also being an important topic in transportation area. Efficiency helps minimizing labor cost in public or interurban transportation.

It's well-known that there are three components of the optimization problems: objective function, constraints and decision variables. In this work, the problem objective is finding minimum number of bus driver scheduling for certain trips weekly. On the other hand, constraints are minimum and maximum work, break, layover duration, legal rights, maximum number of different trip numbers, and location constraints in system. To find decision variables as third component, minimum number of the bus drivers.

The given problem is computationally demanding and therefore, cannot be solved by complex methods such as Branch and Bound. Furthermore, using only the metaheuristics methods possibly produce infeasible solutions, due to search space is not known clearly on account of number of constraints is numerous. Therefore, we attempt to solve both these problems in a two-stage approach. First, Set Covering Problem (SCP), a method applied to many combinatorial problems in literature, is utilized to locate feasible region. SCP results with a feasible region by obtaining all feasible duties. Duties consist of work units that are the parts of trips where driver changes occur. In the second part where the optimization algorithm runs, feasible duties will be assigned to minimize the bus driver numbers, by applying a metaheuristic approach. Underlying reason for selecting metaheuristics is based on the fact that real-life optimization problems usually bring a huge computational burden which cannot be handled by linear and integer programming. Among multiple options such as Tabu Search (TS), Simulated Annealing (SA), Ant Colony Optimization (ACO), Genetic Algorithm (GA) is chosen due to its strong diversity ability due to mutation and cross-over strategies with a simple and intuitive implementation.

In this study, we propose a GA-based bus driver scheduling framework to solve this problem in efficient way. We present the efficiency of our approach by testing it on a pilot data set provided by one of leading transportation company. The results clearly indicate the success of our framework.

Key Words: *Bus Driver Scheduling, Optimization, Genetic Algorithm, Set Covering Problem*

Calculating Happiness Index with Fuzzy DSC TOPSIS Method

Semra Erpolat Taşabat

Mimar Sinan Fine Art University
semra.erpolat@msgsu.edu.tr

Nalan Cinemre

Mimar Sinan Fine Art University
nalan.cinemre@msgsu.edu.tr

ABSTRACT

The happiness expressed in the most general definition as "a state of well-being and contentment" is a relative concept that can be changed from person to person or from community to community. It is an emotional state that can be both personal and social. Although the measurement of the personal and the social happiness are difficult, various approaches and variables have been determined to measure them as a result of the studies performed by the researchers. In order to measure personal happiness, five different approaches were applied. These are: biological, behavioral, implicit measures, other reports (asking others to rate a person's happiness), and self-reports (asking people about their level of happiness by using multiple-item scales or a single question). On the other hand for measuring social happiness six variables were provided. These can be ordered from the most important to the less as follows: GDP per capita, social support, healthy life expectancy, social freedom, generosity, and absence of corruption. One of the most important works carried out to measure social happiness is the World Happiness Report which is first published in 2012. In this report, the happiness of the countries was measured with the help of the above mentioned variables. But then, with the changing world, different variables including social progress and public policies have been added to these variables.

In this study, the happiness index of world countries was tried to be calculated with a different approach by using fuzzy DSC TOPSIS (Distance-Similarity-Correlation Technique for Order of Preference by Similarity to Ideal Solution method). For this purpose, the data published by Gallup Inc. ("Gallup") and used in the world happiness index report were utilized. The results are compared with the results of the world happiness index report.

Key Words: *Fuzzy DSC TOPSIS, Happiness Index*

Calculating Value at Risk: Comparing of Different Estimators

Uğur Karabey

Department of Actuarial Sciences,
Hacettepe University,
Ankara, Turkey
ukarabey@hacettepe.edu.tr

Derya Karagöz

Department of Statistics,
Hacettepe University,
Ankara, Turkey
deryacal@hacettepe.edu.tr

ABSTRACT

Risk measurement is perhaps the most important procedure for insurance companies. Value at Risk (VaR) becomes popular in the last decades as a risk measure for this purposes. On the other hand Weibull distribution is one of the mostly used probability model for modelling the losses.

In this paper we study robust estimation of VaR by employing many estimators such as: Maximum likelihood, Least Squares, Median, Quantile and Quantile least squares. We compare performance of different estimators with a data set that is simulated from Weibull distribution.

Key Words: *Insurance, Robust Estimation, Risk Measures*

Calibration of Internal Clustering Validation Indices

Serhat Emre Akhanlı
Department of Statistical Science,
University College London,
Gower St., London, WC1E 6BT,
United Kingdom
serhat.akhanli.14@ucl.ac.uk

ABSTRACT

In cluster analysis, an important question is how to choose a suitable clustering method and how to determine the best number of clusters. To decide about appropriate cluster analysis technique and the number of clusters, researchers should consider what data analytic characteristics the clusters they are aiming at are supposed to have. For this aim, different internal clustering validation index values (e.g., average within-cluster distances or average between-cluster separation) can be evaluated, which is crucially dependent on the aim of clustering.

Hennig (2017) introduced several internal validation criteria that refer to different desirable characteristics of a clustering, and stated that the user can be interested in some of these criteria rather than just one of them. In this respect, he proposed two random clustering algorithms that are meant to generate clusters in order to standardise the different characteristics so that users can aggregate them in a suitable way specifying weights for the various criteria that are relevant in the clustering application at hand.

As a continuation of Hennig (2017)'s paper, some new additional random clustering algorithms are introduced and the calibration of indexes are more scrutinized with simulation studies. In a final phase, the idea of random clustering is performed on the football player performance data from Akhanli and Hennig (2017)'s paper.

Key Words: *Cluster Analysis, Random Clustering, Calibration, Internal Clustering Validation Indices*

Cardinality Constrained Mean-Variance-Skewness Model by the Artificial Bee Colony Algorithm: An Application in Istanbul Stock Exchange

İlhan Usta
Anadolu University,
Science Faculty,
Department of Statistics,
Eskişehir, 26470, Turkey
iusta@anadolu.edu.tr

Volkan Yeşildere
Anadolu University,
Science Faculty,
Department of Statistics,
Eskişehir, 26470, Turkey
volkanyesildere@anadolu.edu.tr

Mustafa Sarıkavak
Anadolu University,
Science Faculty,
Department of Statistics,
Eskişehir, 26470, Turkey
mustafasarikavak@anadolu.edu.tr

ABSTRACT

The portfolio selection problem is one of the important research topics in modern risk management and Markowitz mean-variance model (MVM) has been accepted as a pioneer model for this problem. However, the MVM is generally inadequate for modeling real-world problem features, such as non-normal returns and a limit on the number of assets in the portfolio. In this study, we present a cardinality constrained mean-variance-skewness model (CCMVSM), extended the MVM to include the skewness of returns and two sets of constraints: cardinality and bounds on holdings. In addition, we design artificial bee colony (ABC) algorithm for the solution of the CCMVSM since it is very difficult to solve the CCMVSM with traditional methods. Finally, the CCMVSM is tested on an empirical dataset consists of the stocks incorporated in the Istanbul Stock Exchange 30. The empirical results show that the CCMVSM is more effective than others, and also the ABC algorithm perform well when dealing with the portfolio selection problem.

Key Words: *Markowitz Mean-variance Model, Mean-variance-skewness Model, Cardinality Constrain, ABC Algorithm*

Cardiovascular Disease Population Risk Tool (CVDPoRT): The Development and Application of a Risk Prediction Model Using Population Health Surveys

Meltem Tuna
Program Manager / Lead Analyst
Ottawa Hospital Research Institute Ottawa,
Ontario, Canada /
Institute for Clinical Evaluative Sciences,
Ottawa and Toronto, Ontario, Canada
The Ottawa Hospital, Civic Campus
1053 Carling Avenue, Box 684
Administrative Services Building - 1st Floor, Room 1014
Ottawa, ON K1Y 4E9
mtuna@ohri.ca

ABSTRACT

Institute for clinical evaluative sciences enables the discovery that leads to better health with the vast health admin data repository and routinely collected population health surveys. Population health surveys complements administrative data with patient characteristics and creates an opportunity to develop more robust prediction algorithms.

We derived Cardiovascular Disease Population Risk Tool (CVDPoRT) using Canadian Community health surveys linked to administrative databases. Predictors included demographic and general health risk factors along with existing diseases. The primary outcome was the first major cardiovascular event resulting in hospitalization or death. Death from a non-cardiovascular cause was considered a competing risk. We developed a website where you can enter information about your lifestyle and it calculates your risk of developing cardiovascular disease in five years according CVDPoRT.

We included 104 219 respondents aged 20 to 105, representing 98% of the Ontario population. There were 3 709 cardiovascular events and over 800 thousand person-years follow-up in the combined derivation and validation cohorts. The final CVDPoRT algorithm was well discriminating (male: C-statistic 0.82, 95% CI [0.81–0.83]; female: 0.86 [0.85–0.87]) and was well-calibrated in the overall population. The algorithm then implemented to an online calculator to make science accessible and to motivate people towards healthier life choices by calculating their risk of having a cardiovascular disease in the next five years.

Key Words: *Cardiovascular Disease, Predictive Algorithms, Competing Risk Analysis*

Characterization Parameters of Domestic Wastewater Samples Collected from Muğla District in Turkey by Decision Tree Method

Sait Güller

Muğla Metro. Municip.
Waste Water Treat. Plant,
Muğla, Turkey
stguller@gmail.com

Gökhan Silahtaroglu

Medipol University
School of Bus. and Man.
Sci.
Management Inf. Syst.
Dep.
Istanbul, Turkey
gsilahtaroglu@medipol.edu.tr

Oğuz Akpolat

Muğla Sıtkı Koçman Uni.
Chemistry Department,
Muğla, Turkey
(Corresponding Author)
oakpolat@mu.edu.

ABSTRACT

This study is related on the characterization parameters of domestic wastewater or treated domestic wastewater samples from municipal wastewater treatment plants located in the provinces of Muğla in Turkey, which are pH, temperature, conductivity, dissolved oxygen, oxygen saturation, salinity, electrical conductivity, chemical oxygen demand, total phosphorus and biological oxygen demand. As 11 of these parameters can be measured in a day-to-day run only in the laboratory the measurement of the BOD₅ parameter lasts 5 days. Statistical evaluation for the laboratory study a data set was created by measuring 12 parameters from 334 samples. Then the effects of parameters of dataset on that of BOD₅ were examined by decision tree method using KNIME data mining package. As a result, the BOD₅ (biological oxygen demand) value distribution was found to be lower than 100 by 53% BOD₅. The ratio of those with BOD₅ values between 100 and 200 is 15.3% and the ratio between those with 450-550 is 12.6%. The parameter which affects BOD₅ value most is chemical oxygen demand (COD). If the chemical oxygen requirement is less than or equal to 214.93, then the value of BIO₅ reaches a value between 0 and 100 and, it has frequency of occurrence of 98.6%. As COD is greater than 214.93, BOD₅ never exceeds 200 value. The probability of BOD₅ being between 100 and 200 is only 1.4%. In cases the chemical oxygen requirement is between 214.93 and 390.445, the most important parameter in predicting the value of BOD₅ is the total phosphorus value.

Key Words: *Experimental Design, Statistics, Optimization, Waste Water*

Classification of Author's Education Degree by Using Three Learning Methods

Betül Kan Kılıç
Eskişehir Technical University,
Faculty of Science,
Department of Statistics,
Eskişehir, 26470
bkan@anadolu.edu.tr

Yonca Yazırlı
Eskişehir Technical University,
Graduate School of Science,
Department of Statistics,
Eskişehir, 26470
yoncayazirli@gmail.com

ABSTRACT

The data that are publicly available in web may help researchers to work easily on their field of study more than generating a data for their aim. Sometimes the unstructured format of the data may prevent to extract the useful information in raw data. To achieve the desired information in data, the pre-processing methods can be used alternatively. One of the important research area in such applications is known as text mining. This paper provides an application of text mining obtained from a text that is not proper for analysis. After a pre-processing phase, then the performance of different learning methods such as random forest, support vector machine and k-nearest neighbour are applied to the structured data. In this paper, a publicly available information is pre-processed by using text mining algorithms in R studio and used to compare the performance in classifying the author's education. The accuracy values obtained show that the best performance is achieved by random forest algorithm.

Key Words: *Web Data, Text Mining, Classification*

Classification of Data by Using Machine Learning Methods

Bünyamin Sarıbacak

University of Ondokuz Mayıs,
Faculty of Education,
Department of Computer
Education and Instructional
Technology,
Samsun, Turkey
bunyamin@omu.edu.tr

Erol Terzi

University of Ondokuz Mayıs,
Faculty of Science and Art,
Department of Statistic,
Samsun, Turkey
eroltrzi@omu.edu.tr

ABSTRACT

Our aim in this study is to explain the concept of classification, which is one of the methods used in the data mining of computer science and other disciplines, using machine learning techniques. In classification problems, the aim is to create an education from the information given using a vector with more than one feature, and to classify the new data correctly in this training. The existing dataset is divided into two to be able to measure the system's success. First one is used as a train set and the second as a test set. This study focuses on the main causes of the discovery of machine learning in statistical data analysis. It tried to explain what the machine learning is. Machine learning is a technology design to build intelligent systems. These systems also have the ability to analyze data by learning from past experiences. The process steps applied in the data mining process are given. Recently, data is rapidly increasing in all area. Very large amounts of data are difficult to process and analyze with old methods. Machine learning methods have been developed to solve is kind of problems. Naive Bayes, Decision Trees, K Nearest Neighbor, Artificial Neural Networks, Support Vector Machines are mentioned most commonly used techniques for classification of data.

Key Words: *Machine Learning, Data Mining, Classification, Classification Algorithms*

Clustering Active Users and Potential Users of a Mobile Payment Application through Self Organizing Maps

Gökhan Aydın

Istanbul Medipol University,
Kavacık Mah. Beykoz, Istanbul
gaydin@medipol.edu.tr

Pakize Yiğit

Istanbul Medipol University,
Kavacık Mah. Beykoz, Istanbul
pyigit@medipol.edu.tr

Şebnem Burnaz

Istanbul Technical University,
Maçka Kampüsü, İstanbul
burnaz@itu.edu.tr

ABSTRACT

The present study aims to arrive at meaningful clusters for both users and potential users of a mobile payment application offered by a major mobile network operator active in Turkey. Attitudes, use intentions and several antecedents (ease of use, usefulness, perceived security, personal innovativeness, compatibility) based on technology acceptance behavior were considered as variables in the study. The data were collected through a field survey using stratified random sampling, where 1395 questionnaires were collected via CATI method. The data were analyzed using SPSS Clementine 12.0 software, by applying Kohonen Self Organizing Maps (SOM), one of the artificial neural network techniques for clustering. The results of the analysis with 22 input nodes and 8 output nodes giving the most meaningful structure, generated four distinct clusters. Non-users or potential users are divided into Cluster 1 and 2 whereas users into Cluster 3 and 4.

Cluster-1 and 2 have 481 and 250 members respectively. Gender distribution is more balanced (46% vs. 54%) in Cluster-1 compared to other clusters. The members of Cluster-1 have the least favorable attitudes and use intentions towards mobile payment application; they are relatively less innovative, have highest security concerns, find the application least useful and least compatible with their lifestyles compared to other clusters. Cluster-2 is mostly composed of men (85%), 93% is aged between 33-47 years old and the members have more favorable attitudes and use intentions than the ones of Cluster-1. Cluster-2 is more similar to Cluster-4 in terms of their members' dispositions towards mobile payment application than Cluster-1. As for the potential users; Cluster-3 and Cluster-4 have 170 and 494 members respectively, most of which are men (%94 and %92). Cluster-3 have the most favorable attitudes and use intentions among all groups and all are aged between 33-47. It's the most innovative group that perceives the application useful and compatible with their lifestyles. The members of Cluster-3 have the lowest security concerns among all clusters. Cluster-4 is the youngest group, with 95% aged between 18-32. The attitudes and use intentions of the members of Cluster-4 are not as favorable as the ones of Cluster-3.

It is clearly observed that Kohonen SOM methodology was able to distinct between users and non-users of a mobile payment application in the context of this study. Differences were detected in terms of dispositions towards application usage and demographics within users and non-user groups, which can provide meaningful implications for the company in further marketing strategies to be directed to each cluster.

Key Words: *Kohonen, SOM, Self Organizing Maps, Neural Network, Mobile Payment*

Clustering Analysis in Spatio-temporal Data Mining

Turgut Özaltındis

Mimar Sinan University

turgut.ozaltindis@msgsu.edu.tr

Elif Özge Özdamar

Mimar Sinan University

ozge.ozdamar@msgsu.edu.tr

ABSTRACT

Nowadays, with the technological developments, the data that is produced and stored has grown and diversified as never before. Now, data mining is used to discover information from large-sized signals obtained by observing photographs, video or space. As the result of the growth and diversification of the data, the structure of the data is different and the techniques used in the data mining have to be adapted according to the changing dimensions and data structures.

One of these different data structures is the spatial data. Spatial data sets are created by including the spatial information of the observations into the data set as latitude and longitude. As spatial (geographical) data is frequently used in today's technology, data mining has been applied to this area and the concept of Spatial Data Mining has emerged.

If a spatial data set contains time variable data, the structure of the data set changes and the techniques used in Spatial Data Mining need to be adapted to this structure. With this requirement, the increase in the number of institutions and scientific researches that produce/store spatial-temporal data has led to the emergence of Spatio-Temporal Data Mining in the near future.

In this study, clustering algorithm used in Spatial and Spatio-Temporal Data Mining were introduced and between the years 1970-2017 ST-DBSCAN spatio-temporal clustering algorithm was performed on the average temperature and precipitation covering all the provinces of Turkey. In the coming years, it is predicted that the literature study in this area will be increased and adapted to the spatio-temporal data structure of many algorithms.

Key Words: *Spatio-Temporal Clustering, Spatial Clustering, Spatio-Temporal Data Mining, Spatial Data Mining*

Clustering Bank Customers Using Self-Organizing Maps

Seda Yanık

Istanbul Technical University,
Industrial Engineering Dept.,
34367, Macka, Istanbul
sedayanik@itu.edu.tr

Abdelrahman Elmorsy

Istanbul Technical University,
Industrial Engineering Dept.,
34367, Macka, Istanbul
abddohossam@hotmail.com

Meltem Yontar

Istanbul Technical
University, Industrial
Engineering,
Maçka Kampüsü, 34367,
Istanbul, Turkey
yontarmeltem@yahoo.com

ABSTRACT

The process of marketing segmentation is vital for deriving the right marketing strategy and any form operations to be successful. It is of the utmost importance to be as precise and this will not be possible without thorough and careful analysis of the data generated by the market. In this paper, an attempt of such analysis was carried out based on two approaches K-means and Self-Organizing Maps (SOM) clustering technique. The main goal was to derive eight distinct market groups, which would require further analysis of their individual demographical and consumption features. A data set of initially more than 4 million credit card transaction for a Turkish bank was used, though it was of an extremely large size. A decision was made to use only a portion of the existing data, as the processing power in hand was incapable of performing an analysis of the full data set. After preparing the chosen data, the result was 38887 different customers with 37 different demographical and expenditure features. In the following step, a PCA was performed in order to remove any redundant data and simplify it used so it would not complicate the clustering process. After that, an initial K-mean clustering was performed in order to obtain a reference point and to find a proper number of clusters. A Self-Organizing Map of 5*5-25 neuron clustered the data set twice, the first time all available features were included in the process, an ANOVA analysis was also performed to assess the distinguishability of the clusters by comparing all clusters centroids or means and determine which of the 37 features were significant for clustering process. As next, SOM-clustering excluding expenditure behavior features and including only the demographical feature was carried out. It obtained eight clusters for which ANOVA analysis was also carried out. This was important for assessing the validity of including criterion variables (expenditure behavior) in differentiating the obtained clusters by to the very same features, or criterion validity.

As expected, the ANOVA results concluded that only demographical features were significant in the clustering on that clusters' centroids were only different for demographical features. Only three features of the expenditure behavior feature were valid as clustering criteria, which are transaction behaviors (sum of transactions, number of transactions and average value of a transaction). In the ANOVA analysis both P-value for the three features showed their validity and significance in the analysis. A complete analysis of the obtained clusters by all three processes was also provided for a further recommended study and analysis to derive the banking market strategy.

Key Words: *Credit Card Transactions, Segmentation, Self-organizing Maps, PCA, ANOVA*

Combination of Forecasting Methods based on Fuzzy C-Means Clustering

Nihat Tak
Kırklareli
Universitesi,
Kayali Kampusu,
Kırklareli
nihattak@gmail.com

Erol Eğrioğlu
Giresun Üniversitesi,
Güre Yerleşkesi,
Giresun
erole1977@yahoo.com

Eren Baş
Giresun Üniversitesi,
Güre Yerleşkesi,
Giresun
eren6061@hotmail.com

Ufuk Yolcu
Giresun Üniversitesi,
Güre Yerleşkesi,
Giresun
varyansx@hotmail.com

ABSTRACT

Forecasting future values of a time series is a popular topic that is studied by researchers. There many forecasting methods that are recently introduced in the literature. While some of these studies aim to solve probabilistic time series, some of them aim to solve non-probabilistic time series. Because the real-world time series contain both probabilistic and non-probabilistic structure in them, the proposed method is introduced for dealing both probabilistic and non-probabilistic part of a time series. The proposed method aims to aggregate as many forecasting methods as we can collect in functions by using fuzzy c-means clustering technique. The aggregation is conducted based on forecasting performances of the methods for a given dataset. Ankara air pollution and Australian beer consumption datasets and numerous forecasting methods for both datasets are selected as applications. The results show that it is possible to get better forecasting results by using the power of many forecasting methods in a function.

Key Words: *Forecasting, Fuzzy C-Means, Meta-Analysis*

Comparing Performances of Artificial Intelligence Techniques based Wavelet Transforms and Feature Selection for Detection of Epileptic Seizures

Ozan Kocadağlı

Department of Statistics,
Mimar Sinan Fine Arts University,
Istanbul, Turkey.
ozankocadagli@gmail.com

Ezgi Özer

Department of Electrical Engineering,
Nova University of Lisbon,
Portugal.
ezgiozerhs@gmail.com

ABSTRACT

Epilepsy is a central nervous system disorder, which is temporary abnormal electric discharges in the nerve cells, so the patients suffer from recurrent seizures that occur at unpredictable times and usually without warning. This study presents an efficient procedure that provides an accurate classification of Electroencephalogram (EEG) signals for detection of epileptic seizures. Essentially, this procedure can be adapted to various artificial intelligence (AI) techniques based on the discrete wavelet transforms (DWT) and feature selection methods.

To bring out the systematic behaviors of epileptic seizures, the proposed automated multi-resolution signal processing technique splits EEG signals into the detailed partitions with different bandwidths, and then establishes an efficient feature extraction framework using DWT. By means of the feature selection methods, the feature matrix is reduced into the significant components where they are utilized as the inputs in the methodologies of the model estimation. In analysis, to classify the epileptic behaviors in EEG signals, the performances of various AI techniques such as deep neural networks, support vector machines, decision trees, nearest neighbor classifiers and ensemble learning as well as the classical statistical techniques are compared with each other over the benchmark and clinical datasets.

According to the analysis results, the proposed procedure not only allows making an efficient analysis of EEG signals for detection of epilepsy, but also provides the best model configurations for all the approaches in the context of reliability and complexity.

Key Words: *Epileptic Seizure Detection, Signal Processing, EEG Signals, AI Techniques, Model Complexity*

Comparison of Bayes and OLS based Type II Regression

Ece Özgören

Department of Statistics,
Institute of Science Afyonkarahisar,
Afyon Kocatepe Universtiy, Turkey
ecem_ozgoren@hotmail.com

Sinan Saraçlı

Department of Statistics,
Faculty of Science Literature,
Afyon Kocatepe Universtiy, Turkey
ssaracli@aku.edu.tr

ABSTRACT

While Ordinary Least Squares (OLS) is the commonly used parameter estimation technique in regression analysis, accepting a subjective definition by considering the parameters having a probability distribution, Bayesian estimation technique may give better results in some conditions. Bayesian approach calculates posterior distribution via combining the prior distribution information the earlier results of the data set.

Bisector regression technique is a kind of Type II regression technique that considers the error terms of both dependent and independent variables. It is also a commonly used technique among Type II techniques. When there are one dependent and one independent variable in regression analysis, it calculates the regression line by bisecting the two regression lines, which are calculated by getting two of the variables dependent respectively.

There is no study in the literature that calculates the parameters of Type II regression via Bayesian approach. In this study, the bisector regression line is calculated via Bayesian approach. The data set is simulated by MATLAB software at different distribution types and sample sizes. The results of OLS based Type II regression and Bayesian based Type II regression are compared with Mean Square Error (MSE) approach and the results are given in related tables and figures.

Key Words: *Bayesian Estimation Method, Type II Regression, Bisector Regression*

Comparison of Classification Performances for Predictiveness Curve and ROC Curve

Merve Türkegün
Department of
Biostatistics and Medical
Informatics,
Faculty of Medicine,
Mersin University,
Turkey
merveturkegun@gmail.com

Bahar Taşdelen
Department of
Biostatistics and
Medical Informatics,
Faculty of Medicine,
Mersin University,
Turkey
bahartasdelen@gmail.com

Selma Unal
Department of Child
Health and Diseases,
Faculty of Medicine,
Mersin University,
Turkey
unaselma@hotmail.com

Damla Hazal Saral
Department of
Biostatistics and
Medical
Informatics,
Faculty of Medicine,
Mersin
University,
Turkey
dmlhzl5@gmail.com

ABSTRACT

Biomarkers are used in personalized medicine to provide a link between disease diagnosis and prognosis and to estimate patient response to treatment for each diagnosis. Novel biomarkers have been developed in order to estimate disease risk at chronic diseases. For this reason, an appropriate threshold must be calculated to determine high or low risk of disease. The Predictiveness Curve (PC) predicts diseases risks for one or more biomarkers. The PC is a plot of cumulative percentage of individuals whose predicted risks are less than or equal to risk value. PC can predict disease risk for each individual. While ROC curve distinguishes between diseased and healthy populations, the PC yields an estimate of the likelihood of being diseased given a particular test value. Thus, the PC helps the clinician to decide treatment plan for each patient. Our purpose in this study is to compare classification performance of Predictiveness Curve and ROC curve and to evaluate biomarkers of sickle cell diseases using R software in archive-based clinical data set.

Key Words: *Risk, Predictionness Curve, ROC Curve, Biomarker*

Comparison of Classification Performance of the Galaxies in the Shapley Concentration Region

| | |
|---|---|
| Nida Gökçe Narin Department of Statistics, Faculty of Science, Muğla Sıtkı Koçman University, Muğla, 4800, Turkey <i>nidagokce@yahoo.com</i> | Nazlı Deniz Ergüç Department of Statistics, Faculty of Science, Muğla Sıtkı Koçman University, Muğla, 4800, Turkey <i>nazldenizerguc@gmail.com</i> |
|---|---|

ABSTRACT:

Galaxies are systems that come together in the form of stars, gas, dust, and dark matter combined with mass gravity. They are made up of star clusters of different sizes, different degrees of brightness and different coordinates. The first classification of galaxies by Edwin Hubble in 1926 has an important place in astronomical data analysis. Galaxies are classified in three main groups according to a common morphological classification called Hubble order, Elliptic, Spiral and Irregular.

The classification of galaxies in the literature is addressed by two different approaches: morphological and spectral classification. In the morphology-based classification approach, while the appearance of galaxies is defined, the locations of galaxy-forming stars in the spectral classification approach are considered.

In this study, the morphological classification of galaxies was discussed, and in particular, 4215 galaxies in the Shapley Concentration region were investigated taking into account 5 variables (deviation in latitude, longitude, brightness, velocity and velocity).

Natural groups of the galaxies in this region were identified by IDL programming from the Abell catalog. The galaxies, whose natural groups were identified, were classified by Weka programming. For classification, Naive Bayes and Bayes Net algorithms from Bayes classifiers, J48, LMT, Random Forest algorithms from Decision tree algorithms, MLP for classification by artificial neural networks and finally Support Vector Machine algorithm were used. The performances of the methods examined were compared with the Abell catalog and the best classification method was determined in the classification of the galaxies.

Keywords: *Shapley Concentration Area, Data Mining, Classification, Galaxies, Abell Catalog*

General area of research: *Statistics, Astronomy*

Acknowledgement: *This paper has been granted by the Mugla Sıtkı Kocman University Research Projects Coordination Office. Project Grant Number: 18/037 and title, "Shapley Konsanstrasyon Bölgesindeki Galaksilerin İstatistiksel Öğrenme Yöntemleriyle Sınıflandırması"*

Comparison of Least Squares Estimator with M-Type Regression Estimators

Ahmet Toy

Department of Statistics,
Ondokuz Mayıs University,
Samsun, Turkey
toyahmet01@gmail.com.tr

Erol Terzi

Department of Statistics,
Ondokuz Mayıs University,
Samsun, Turkey
eroltrzi@omu.edu.tr

Cemil Çolak

Biostatistics and Medical
Informatics, Inonu University,
Malatya, Turkey
cemil.colak@inonu.edu.tr

ABSTRACT

The least squares estimator is an easy and common method for understanding the relationship between variables. The purpose of this research is to minimize the sum of the squares of the error terms. The least squares estimator gives good results if there are no outliers in the data set. However, since this condition is difficult to reach in practice, alternative methods offered are not affected by outliers. These methods are called as robust regression estimators because they are not affected by outliers.

M-type robust regression estimators were introduced in 1964. This estimator has taken its name from the Maximum Likelihood-ML estimator. The purpose in the least squares estimator is to minimize the sum of the squares of the error terms while the goal is to minimize the function $p(e)$ of the error terms in the M-type estimators.

In this study, M-type estimators and least squares estimator are compared on the dataset. For this purpose, methods of determining outliers and robustness criteria are also discussed.

Key Words: *Least Squares, Robust Regression, M-Type Estimators, Outliers*

Comparison of Optimal Reinsurance Studies on One Side and on Both Sides

Betül Zehra Karagül

Department of Actuarial Sciences,
Hacettepe University,
06800 Beytepe, Ankara
betul.zehra@hacettepe.edu.tr

Murat Büyükyazıcı

Department of Actuarial Sciences,
Hacettepe University,
06800 Beytepe, Ankara
muratby@hacettepe.edu.tr

ABSTRACT

Reinsurance is an effective risk management tool for an insurance company to transfer risk to a reinsurance company. When an insurer seeks reinsurance protection, the insurer is faced with the classic trade-off between the retained loss and the reinsurance premium. There have been many studies addressing the optimality of reinsurance depending on the optimality criterion and the chosen premium principle. Most of the papers on optimal reinsurance have only considered the insurer point of view and then try to maximize/minimize some measure of the risk of the insurer. However, there are two parties to a reinsurance contract. In recent years we find some optimal reinsurance studies which have considered the joint point of view insurer and reinsurer. The aim of this study is to compare the optimal reinsurance studies that deal with the both sides and that deal with only one side.

Key Words: *Optimal Reinsurance, Optimal Joint Reinsurance, Optimal Retention Limit*

Comparison of Some Estimators for Parameters of log-Dagum Distribution

Caner Tanış Merve Çokbarlı Buğra Saraçoğlu
Selcuk University Selcuk University Selcuk University
ctanis@selcuk.edu.tr mervecokbarli@gmail.com bugrasarac@selcuk.edu.tr

ABSTRACT

In this study, we have studied about maximum likelihood and approximate bayes estimation for the log-Dagum distribution with (β, λ, δ) parameters. The maximum likelihood estimators and asymptotic confidence intervals for unknown parameters of log-Dagum distribution are obtained. Bayes estimators for unknown parameters of this distribution are obtained using Tierney-kadane approximation under squared error loss function. Finally, a Monte-Carlo simulation study is performed to compare performances of these estimators in terms of mean square errors and bias.

Key Words: *Log-Dagum Distribution, Maximum Likelihood Estimator, Asymptotic Confidence Interval, Approximate Bayesian Estimation, Tierney-Kadane Approximation*

Comparison of Supervised Disease Diagnosis Methods for Unbalanced Classes: The Case of Diabetes Diagnosis

| | | | | |
|---|--|---|---|---|
| <u>Sultan Turhan</u> Muğla Sıtkı Koçman University, Faculty of Science, Department of Statistics, Muğla, Turkey sltnturhan@hotmail.com | Yüksel Özkan Ege University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, İzmir, Turkey yukseozkan03@gmail.com | B. Sarer Yurekli Ege University, Faculty of Medicine, Department of Endocrinology, İzmir, Turkey bsareryurekli@yahoo.com | Aslı Suner Ege University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, İzmir, Turkey asli.suner@ege.edu.tr | Eralp Doğu Muğla Sıtkı Koçman University, Faculty of Science, Department of Statistics, Muğla, Turkey eralp.dogu@mu.edu.tr |
|---|--|---|---|---|

ABSTRACT

As the volume of knowledge produced in health research has been rapidly increasing, it is critical to pre-process research outcomes to be able to learn from large-scale and complicated disease diagnosis dataset. When classification of healthy and disease groups is considered unbalanced data problem is one of the major problems. Unbalanced data occurs when a certain class (for example, disease class) is proportionally less represented compared to the others. Thus, a learning bias towards the majority class is expected. The bias is generally defined as the problem of misclassification resulting from the tendency to ignore the minority class. Because of the natural complexity of these datasets, it should be subjected in the pre-processing phase for balancing data set and classification is recommended to be done afterwards.

In this study, under-sampling and over-sampling methods were used to eliminate the unbalanced class problem for the classification of diabetes. Effects on the classification performance were compared by integrating balanced data into bagging and boosting based ensemble learning methods (Random Forest, Weighted Sub-Space Random Forest, Adaboost and Logitboost).

Key Words: *Ensemble Learning, Classification, Unbalanced Data, Disease Diagnosis, Diabetes*

Comparison of the Claims Development Results Using the Chain Ladder and Extended Complementary Loss Ratio Methods for Turkey

Müge Yeldan
mugeyeldan@hacettepe.edu.tr

Yasemin Gençtürk
yasemins@hacettepe.edu.tr

ABSTRACT

Although the insurance companies have information from different sources such as paid and incurred claims, generally only one source of information is used in order to estimate reserve in many classical reserve methods. When an estimation is calculated based only on paid or incurred claims, the ultimate claims will be different. As the incurred claims are expected to be paid at the end of the development year for each accident year, the ultimate claim estimates obtained using these two sources of information are also expected to be equal. Therefore, reserve methods, one of which is the Extended Complementary Loss Ratio method, based on both paid and incurred claims have been developed in order to obtain equal or almost equal estimations.

The claims development result which is the difference between the estimates of the outstanding claims liabilities in successive years measures the change in the claim reserve for one year. As this result has a direct impact on the profit and loss statement and on the financial strength of the insurance company, it has an effect on the solvency of the company. The aim of this study is to calculate and compare the claims development results when the reserve is estimated using both the Chain Ladder and Extended Complementary Loss Ratio methods for compulsory motor third party liability insurance in Turkey.

Key Words: *Reserve Estimation, Chain Ladder Method, Extended Complementary Loss Ratio, Claims Development Result*

Comprehensive Meta Analysis

İrem Yanık
Hacettepe University,
Department Of Statistics
06800 BEYTEPE ANKARA
TURKEY
irem.yanik@hacettepe.edu.tr

Prof.Dr.Hüseyin Tatlıdil
Hacettepe University,
Department Of Statistics
06800 BEYTEPE ANKARA
TURKEY
htatlidil@yahoo.com

ABSTRACT

Comprehensive Meta Analysis is a package program developed to perform meta analysis. Meta analysis provides a general judgment on the subject by combining work done independently on the same subject. A general judgment on these studies is reached by analyzing independent studies on similar topics statistically. It is easier and more reliable to make a new study by taking advantage of the previous studies because the number of analyses performed on the same subject increases every day and every analysis is forced by the researcher in terms of material and time. In addition, the researcher may not have sufficient sample for the subject to be studied; since we combine independent studies in meta analysis, the sample of the subject to be studied has also been expanded, so the study will have more consistent results. Because multiple studies are combined in metadata analysis, the measure that is needed to combine these studies is the size of the impact. The concept of impact size is very important for meta analysis. To do meta analysis, the researcher must first select the work that he / she will include in the analysis. Then, each study should calculate the impact sizes and perform statistical analysis for impact sizes and interpret the results.

In this study, the package program developed to perform meta analysis was introduced and the use of the program was explained. An original application was resolved with the help of Comprehensive Meta Analysis V3.

Key Words: *Comprehensive meta analysis, meta analysis, effect size*

Conditional Value-at Risk Modelling on Oil Prices for Emerging Markets

| | |
|--|---|
| <u>Tolga Yamut</u> Dokuz Eylul University, Faculty of Sciences, Department of Statistics, Izmir yamuttolga12@yahoo.com | <u>Burcu Hudaverdi Ucer</u> Dokuz Eylul University, Faculty of Sciences, Department of Statistics, Izmir burcu.hudaverdi@deu.edu.tr |
|--|---|

ABSTRACT

Concept of systemic risk is crucial to identify risk environment of a financial instrument. A market can become more risky by an influence of an element within the system or another market which is highly associated. There exist some methods for the measurement of a systemic risk. One of them is Conditional Value-at-Risk (CoVaR) which captures the risk systematically. CoVaR involves a crisis scenario which influencer market is in distress and measures the risk under this condition. In this study, we use proper copula distributions to identify the CoVaR of emerging markets when Brent oil market is in trouble and risk participations of Brent oil are calculated. Two different CoVaR approaches are applied. Also, backtesting is used to test the CoVaR models by the notion of joint violations through simulation.

Key Words: *Systemic Risk, Conditional Value at Risk, Copula*

Conditional VaR Forecast for Dependent Losses

Risti Nur'aini
Institut Teknologi Bandung
Indonesia

Khreshna Syuhada
Institut Teknologi Bandung
Indonesia
khreshna@math.itb.ac.id

Mahfudhotin
Institut Teknologi Bandung
Indonesia

ABSTRACT

Value-at-Risk (VaR) is a risk measure defined as a maximum loss that will occur in a certain probability. Forecasting VaR is aimed to prevent higher losses. VaR forecast may be calculated for the case of random loss alone and/or of a random loss that depends on another random loss. In both cases, VaR forecast is obtained via its (conditional) probability distribution of loss data, specifically the quantile of loss distribution.

Modelling random losses may be carried out by employing certain stochastic processes such as ARCH/GARCH class of model. This paper aims at investigating quantile-based VaR calculation of dependent random loss, including such heteroscedastic processes. To obtain its VaR forecast, the use of Copula may be unavoidable since loss distributions are neither normal nor identical. Copula is a multivariate distribution model that provide more flexibility in their marginal distribution. We specifically show the effect of dependence, e.g. Pearson correlation and Kendall's Tau, in forecasting VaR. Furthermore, we consider on two random losses that have a strong dependence in their extreme losses or their tail distribution, measured by tail dependence coefficient. We determine VaR forecast accuracy when two losses have tail dependence coefficient.

Key Words: *ARCH, Copula, Tail Dependence, Stochastic Prediction, Volatility*

Confidence Intervals and Region for Kumaraswamy Distribution Using Wang Pivotal Quantities under Progressive Censoring

Kadir Karakaya
Selcuk University,
Konya, Turkey
kkarakaya@selcuk.edu.tr

İsmail Kınacı
Selcuk University,
Konya, Turkey
ikinaci@selcuk.edu.tr

Coşkun Kuş
Selcuk University,
Konya, Turkey
coskun@selcuk.edu.tr

Yunus Akdoğan
Selcuk University,
Konya, Turkey
yakdogan@selcuk.edu.tr

ABSTRACT

The Kumaraswamy (Kw) distribution which has $(0,1)$ domain has been introduced by Kumaraswamy (1980). In this study, progressive Type-II censored sample is considered and an exact confidence interval and confidence region for the parameters of Kumaraswamy distribution are discussed through Wang pivotal(Wang, 2008) quantities. A simulation study is performed to compare the coverage probability and the mean volume of exact and asymptotic intervals and regions. A numerical example is also given to illustrate the derived methodology.

Key Words: *Confidence Interval, Confidence Region, Simulation, Progressive Type-II Censored Sample*

Conflicts in the Calculation of the Power of a Statistical Test, by Psychology Students. A Semiotic Point of View.

Osmar Darío Vera

Universidad Nacional de Quilmes.
Roque Sáenz Peña 352. Bernal.
Buenos Aires. Argentina
osmar.vera@unq.edu.ar

ABSTRACT

For a statistical test, we ask as an open question how to calculate its power, given a value of the parameter under the alternative hypothesis and a level of significance. We analyze the answers in that question in the mathematical practices of University of Huelva's psychology student (N=224). We use as theoretical framework the Onto Semiotic Approach of mathematical cognition. As part of the framework, we consider the mathematical objects, processes involved, and results with the aim of discover semiotic conflicts that lead to institutionally inadequate responses. We present as a result a detailed classification of semiotic conflicts related to objects that are part of the determination of the power statistical hypothesis testing.

Key Words: *Power of a statistical test, university students, Psychology students, onto semiotic conflicts. Onto Semiotic Approach (OSA).*

Constructing Discrete Bayesian Networks from Mixed Data in R

Günel Bilek

Department of Statistics,
Bitlis Eren University,
Bitlis, Turkey
gunalbilek@gmail.com

Efehan Ulaş

Department of Statistics,
Çankırı Karatekin University,
Çankırı, Turkey
ef_ulas@hotmail.com

ABSTRACT

Bayesian networks (BNs) are graphical structures used to reason and represent knowledge in an uncertain field. While constructing BNs, the data may need to be discretized for several reasons.

The aim of this study is to show how to discretize continuous data by using two approaches, among many others, and construct discrete BNs in R. Firstly, three discrete and seven continuous variables are simulated in statistical software R. Next, the first discretization process is conducted using the domain knowledge, while the latter is conducted using a statistical approach called Information-Preserving Discretization. After the discretization process, the BNs are built, the results are interpreted and queries are made via plots and figures using mainly the **bnlearn** package in R. Finally, the two BNs are compared using network scores. The comparison of the network scores indicate that the Information-Preserving Discretization method leads to the BN that is more likely to produce better estimates.

Key Words: *Bayesian Networks, Data Discretization, Bnlearn*

Continuum Regression and Comparisons of Some Related Estimators

Berrin Gültay
Çanakkale Onsekiz
Mart University,
Faculty of Art and Science,
Department of Statistics
berringultay@comu.edu.tr

Hüseyin Güler
Çukurova University,
Faculty of Economics
Administrative Sciences,
Department of Econometrics
hguler@cu.edu.tr

Selahattin Kaçıranlar
Çukurova University,
Faculty of Art and Science,
Department of Statistics
skacir@cu.edu.tr

ABSTRACT

For multiple linear regression with non-orthogonal regressors there are some alternatives to the ordinary least squares (OLS) method which are advantageous when the regressors are near collinear. Several so-called shrinkage or regularized regression methods can be highly preferable to OLS, by trading bias for variance. Well-known methods are Ridge Regression (Hoerl and Kennard, 1970), Two Parameter Ridge Regression (Lipovetsky and Conklin, 2005) and Shrinkage Estimation (Mayer and Wilke, 1973).

Stone and Brooks (1990) introduced the method of 'Continuum Regression' (CR), in which they considered a spectrum of possible regressors, each associated with a value of a parameter γ . Continuum regression is a regularized regression estimation method, and being so, it is also particularly intended for dealing with the multicollinearity problem. The aim of this study is to evaluate the performances of the mentioned methods according to Mean Squared Error of Prediction (MSEP) criteria. For this purpose the data set in Fearn (1983) is analyzed and a simulation study is conducted.

Key Words: *Continuum Regression, Multicollinearity, Ridge Regression, Biased Estimators*

Control Charts Based on Com-Poisson Distribution

Canan Hamurkaroğlu
Karabük University,
Faculty of Business,
Department of Actuary
and Risk Management
cananhamurkaroglu@
karabuk.edu.tr

Ayten Yiğiter
Hacettepe University,
Faculty of Science,
Department of Statistics
yigiter@hacettepe.edu.tr

Yasemin Gençtürk
Hacettepe Üniversitesi
Faculty of Science,
Department of Actuarial Science
yasemins@hacettepe.edu.tr

ABSTRACT

The classical Shewhart c and u control charts developed via the Poisson assumption are respectively used to examine the number of defective per unit and the average number of defective in a process. One important characteristic of Poisson distributed data is that the mean and the variance should be close to the same. However, this equi-dispersion assumption of the Poisson distribution is limiting in a real data. Therefore, the use of the Shewhart c and u control charts based on Poisson distribution when data displays over-dispersion will result in falsely detecting many observations actually in-control as out-of-control. In the meantime when data displays under-dispersion, many observations actually out-of-control are falsely interpreted as in-control because they fall within the associated bounds as a result of the limit bounds determined with a Poisson assumption being too broad. Hence, in the literature the control charts based on Conway-Maxwell Poisson distribution for over-dispersed or under-dispersed data are proposed. In this study, the performance of the control chart based on Con-Maxwell Poisson and the Shewhart c and u charts is compared by simulation.

Key Words: *Statistical Process Control, Shewhart c and u Control Charts, Com-Poisson Distribution*

Copula Based Regression Models: An Application

Rümeysa Karataş

Hacettepe University
rmys@hacettepe.edu.tr

Uğur Karabey

Hacettepe University
ukarabey@hacettepe.edu.tr

ABSTRACT

Estimating total loss of an insurance portfolio is crucial for pricing, calculation of premiums and risk assessment. Total loss is defined as product of two components which are average claim size and number of claims. Traditionally, claim sizes and counts are assumed to be independent in non-life insurance. In practice, claim frequency and severity are often dependent and ignoring this dependency can cause an over or underestimate for total loss. In this study, we model average claim sizes and number of claims if both quantities are dependent. We use a joint copula-based regression model for insurance claims and sizes. The dependence between claim sizes and number of claims is expressed by a bivariate copula. To use regression models three covariates are specified. These are age, gender and type of car. The copula regression model is applied to simulated data set.

Keywords: *Dependence Modeling, Regression Models, Copula, Simulation*

Correlation Based Penalty Estimator in Vector Autoregressive Model

| | | |
|---|---|--|
| Bahadır Yüzbaşı Inonu University, Malatya, Turkey byzb@hotmail.com | <u>Ahmet Demiralp</u> Inonu University, Malatya, Turkey ahmt.dmrlp@gmail.com | Mehmet Güngör Inonu University, Malatya, Turkey mgungor44@gmail.com |
|---|---|--|

ABSTRACT

The vector autoregressive (VAR) model is very popular for modeling multiple time series. Also, VAR model is a very suitable tool for the analysis of the behavior of economic time series, for forecasting the future values, data description. And VAR model produces more accurate result than the other model such as Conditional Model (CM). But, as many economic time series are concerned with their own past values and present and past values of other economic time series, the increasing number of variables and lags may cause multicollinearity. In literature, there are a number of methods in order to beat with multicollinearity. Some popular of them are Ridge Estimation, LASSO, Elastic net, Liu type estimation and Principal Component Analysis (PCA). Correlation Based Penalty (CP) estimation is also an efficient method to combat with multicollinearity. Because it aims at the selection of groups of correlated variables. In this study, we incorporate VAR model to CP estimation. In order to compare this estimator with its competitive, we conduct a Monte Carlo simulation study and, a real data example to illustrate the usefulness of the suggested methods.

Key Words: *Correlation Based Penalty, VAR Model, Multicollinearity*

Acknowledgement: *This work is supported by Research Fund of the Inonu University with project number : SDK-2018-991*

Cox Regression Model with Covariates Subject to Measurement Error

Hatice Işık

Department of Statistics,
Hacettepe University,
06800, Beytepe, Ankara, Turkey
haticeyeniay@hacettepe.edu.tr

Duru Karasoy

Department of Statistics,
Hacettepe University,
06800, Beytepe, Ankara, Turkey
Duru@hacettepe.edu.tr

ABSTRACT

Cox regression model is the most preferred model for the survival data. It is important to evaluate the true effect of the covariates in the model. Measurement error is a problem in the studies with those especially using biomarkers for the covariates. The error arises for different reasons such as laboratory error, self-reporting, limitations of measurement. The existence of some covariates measured with error causes biased parameter estimation in the Cox regression model.

With the increasing advantage of the flexible model structure and the available software programs in Bayesian computation, Bayesian methods have been favored recently in the analysis. Lately, one of the suggested methods coping with the measurement error problem is corrected Bayesian method.

We compare the results for the Cox's partial likelihood method ignoring measurement error and its Bayesian approach and the corrected Bayesian method with measurement error accounted for. We have conducted simulation studies with different scenarios to evaluate the performance of the estimators obtained with these methods.

Key Words: *Cox Regression, Measurement Error, Bayesian Analysis*

Creation of a University League Based on the Contribution of their Economics Departments to Economics Education

S. Baha Yıldız
Manisa Celal Bayar University
baha.yildiz@cbu.edu.tr

Volkan Alptekin
İzmir Katip Çelebi University
volkan.alptekin@ikc.edu.tr

Sibel Selim
Manisa Celal Bayar University
sibel.selim@cbu.edu.tr

ABSTRACT

This study aims to measure the added value created by the economics departments of the universities in Turkey for students throughout their undergraduate education. For the analysis section, the minimum admission scores of the universities' economics departments for the years from 2000 to 2011 were used as input and the net average scores obtained in the economics tests of Public Personnel Selection Examination (PPSE) for the years from 2004 to 2015 were used as output. The data were normalized using the Min-Max method and the universities were ranked using the "Borda Count" method. According to the results, Ankara University ranked the 1st, Hacettepe University the 2nd and Middle East Technical University the 3rd. The first seven positions in the top ten ranking are held by the universities located in the capital Ankara. This is attributed to various factors such as the long-established character of the universities in the capital city, availability and diversity of training courses and materials for PPSE, permanency of the academic staff, prevalence of standard daytime education in these universities, and the role of the capital as the seat of public institutions, resulting in greater motivation among students as they have more contact with senior government officials.

Key Words: Added Value, Normalization, University Ranking, Public Personnel Selection Examination (PPSE)

Cryptocurrencies, Currencies and Gold Fractal Volatility Analysis

Mária Bohdalová
Faculty of Management,
Comenius University in Bratislava,
Odbojárov 10, 82005 Bratislava,
Slovakia
maria.bohdalova@fm.uniba.sk

Michal Greguš
Faculty of Management,
Comenius University in
Bratislava, Odbojárov 10, 82005
Bratislava, Slovakia
michal.gregus@fm.uniba.sk

ABSTRACT

Bitcoin is breaking down the barriers of digital currency transactions and the technology behind it has the potential to transform the way we pay for goods and services. This is the reasons why a few innovations in the money markets have brought more attention by regulators and policy makers than the digital currency Bitcoin. The Bitcoin is continuously changing its volatility. The unusual level of Bitcoin's volatility can be attributable to speculative trading. This paper investigates fractal properties of selected cryptocurrencies namely Bitcoin, Ethereum and Litecoin. We have discovered periodical and non-periodical cycles during existence of these cryptocurrencies and dynamics of the Hurst coefficient during analyzed period. Finally, we compare fractal structures of cryptocurrencies, gold and selected currencies.

Key Words: *Hurst Coefficient, Fractal Analysis, Cryptocurrencies, Cycles*

CT-Visible Doxorubicin-Eluting Clay Particles for Transarterial Catheter-Directed Drug Delivery

Maide Gökçe Bekaroğlu
Istanbul Technical University,
Faculty of Science and Letters,
Dept. of Physics
bekaroglum@itu.edu.tr

Fuad Nurili
Department of Radiology,
Memorial Sloan Kettering
Cancer Center
fuatnurili@gmail.com

Sevim İşçi
Istanbul Technical University
Faculty of Science and
Letters, Dept. of Physics
iscisev@itu.edu.tr

ABSTRACT

Transcatheter arterial chemoembolization (TACE) is an effective palliative treatment of cancerous lesions, which was developed to minimize systemic toxicity of the chemotherapeutic agent and simultaneously provide embolization. After embolization, these embolic microparticles loaded with chemotherapeutic agents permit a sustained drug release into the tumor blood supply. Multi-functional embolic microparticules with efficient drug carriers' abilities and traceability afterward may be significant in order to improve therapeutic outcomes of TACE procedures. A clay mineral, montmorillonite, is highly applicable and unique for transcatheter arterial embolization procedure due to the size, adsorption capability, and biocompatibility of the particles. In the technique, the arterial vessels around tumors are blocked by drug loaded micro-sized montmorillonite particles so reduced blood flow to tumor causes hypoxia and nutrient deprivation while anticancer drugs are released by the particles into the tumor site. Both devascularization (achieved by cutting off the oxygen and the nutrition) and drug release around tumor site can induce progressive shrinkage of the tumor size and cell growth reduction.

In this study, CT-visible doxorubicin-eluting montmorillonite microparticles were developed for TACE therapy. Doxorubicin (DOX) was attached to Mt particles with electrostatic interactions by adsorption method. DOX loaded Mt was dispersed in Urografin in order to obtain computed tomography (CT) images. The biocompatibility, tomography imaging properties and doxorubicin release kinetics of these microparticles were investigated during *in vitro* studies. A rabbit kidney model was used to demonstrate a) feasibility to deliver these microparticles to targeted region using transcatheter intra-arterial microcatheters, b) efficacy of embolization and c) potential to visualize the delivery of these microparticles to the desired region with CT after infusion.

Key Words: *Montmorillonite, adsorption, micro particles, drug delivery, chemoembolization*

Customer Satisfaction Survey for Old Age Asylum Using SERVQUEL

Suna Çetin
Kirikkale University,
Department of Industrial
Engineering
sunacetin@kku.edu.tr

Kezban Bulut
Kirikkale University,
Department of Industrial
Engineering
kezbanbulut@yahoo.com

Hamza Kuzucu
Kirikkale University,
Department of Industrial
Engineering
Hdk968@gmail.com

ABSTRACT

Customer satisfaction is defined as the ratio of satisfaction after meeting customer needs and expectations. Customer satisfaction is necessary for all businesses. If your customers are satisfied, your customers become loyal and this is one of the most important criteria in the success and continuity of your business. New age customer type waits for more special attention and service, desires to be understood of feelings and thoughts. In order to keep up with the changing conditions of today's competitive environment, "product and profit" oriented classical management approach has been replaced by "quality and customer" oriented modern management approach. In other words, the understanding of the product or service can find any kind of buyer in the market has changed and left the place to the understanding of product or service for the customer. In this study, it was aimed to measure the customer satisfaction of the elderly residing in an old age asylum operating in Ankara using SERVQUEL. The information obtained from the SERVQUEL questionnaires will be used to provide feedback to old age asylum manager and evaluate customer requests.

Key Words: *Servquel Method, Customer Satisfaction, Old Age Asylum*

Detection of Lines on Images with Multimodal Directional Artificial Bee Colony Algorithm

Orhan Kesemen
Department of
Statistics and
Computer Sciences,
Karadeniz Technical
University, 61080,
Trabzon, Turkey
okesemen@gmail.com

Eda Özkul
Department of
Statistics and
Computer Sciences,
Karadeniz Technical
University, 61080,
Trabzon, Turkey
eda.ozkul.gs@
gmail.com

Buğra Kaan Tiryaki
Department of
Statistics and
Computer Sciences,
Karadeniz Technical
University, 61080,
Trabzon, Turkey
bugrakaantiryaki@
gmail.com

Özge Tezel
Department of
Statistics and
Computer Sciences,
Karadeniz Technical
University, 61080,
Trabzon, Turkey
ozge_tzl@hotmail.com

ABSTRACT

Detection of shapes in digital images has many applications in image processing. Determination of the coins, finding the license plate and the iris in the image can be exemplified for these applications. Hough transform is the most common method used to detection of the shapes in digital images. However, there are usually noises and may be missing points and pixels in digital images. This situation poses an obstacle to exactly detect of shapes. However, shapes can be detected using the Hough transform as well in noisy images. The Hough transform associate pixels to parameters by matching the pixels in the image with parameter space.

The Hough transform is the most basic approach, but the computation time also increases when the image size increases. For this reason, in this paper, Artificial Bee Colony Algorithm (ABC), which is a meta-heuristic optimization method, is used to reduce the computation time for detection of lines in the digital images. When the Hough transform is used, a line in the image space represents a point in the parameter space. If the line in the image space is a perpendicular line, the values in the parameter space approach infinity. Therefore, polar coordinates are used to define lines. Polar coordinates contain angular changes. Because of this, the ABC algorithm must be adapted to directional system. Therefore, in this study, multimodal directional ABC algorithm is proposed for detection of the lines in the digital images and applications of this algorithm are performed on the cylindrical surface.

Key Words: *Line Detection, Hough Transform, Artificial Bee Colony Algorithm, Directional Optimization*

Determining Anomalies of BIST 100 Index by BURR XII Distribution

| | | | |
|--|--|---|--|
| <u>Esra Betül Kınacı</u> Gazi University, Faculty of Science, Department of Statistics esrakinaci@gazi.edu.tr | H. Hasan Örkücü Gazi University, Faculty of Science, Department of Statistics hhorkcu@gazi.edu.tr | Hasan Bal Gazi University, Faculty of Science, Department of Statistics hasanbal@gazi.edu.tr | Harun Kınacı Erciyes University Faculty of Economics and Administrative Sciences hkinaci@erciyes.edu.tr |
|--|--|---|--|

ABSTRACT

According to the efficient market hypothesis, investors, whatever they do, can not get a return on normal in an efficient market. However, in some of the empirical studies, different results have been encountered rejecting this proposal, which is called anomaly (Turaboğlu,2017). In this study, periodical anomalies which behave unusually in various time periods depending on time have been discussed. The BIST 100 index data for the detection of anomaly was modeled by Burr XII distribution. The Burr XII distribution is often used in many different areas such as reliability analysis, risk analysis and estimation of wind power capacities but its parameters can not be estimated analytically. In such cases, iterative methods are used for estimation. An important subset of the iterative search methods are heuristic search methods. In this study, Simulated Annealing (SA) algorithm which is a heuristic search method is used for parameter estimation. SA is a probabilistic single-solution-based heuristic search method based on logic of the annealing process in metallurgy. The existence of anomalies in a certain period of the modelled BIST 100 index by Burr XII was investigated by SA. The results are given in tables and graphs. In addition probability values were obtained based on different return or loss values

Key Words: *Anomalies, Burr XII Distribution, Simulated Annealing*

Diagonal Parameter Symmetry Model Based on Euclidian Distance in Square Contingency Tables

Serpil Aktaş Altunay
Department of Statistics
Hacettepe University
Ankara, Turkey
spxl@hacettepe.edu.tr

Gökçen Altun
Department of Statistics
Bartın University
Bartın, Turkey
gokcenefendioglu@gmail.com

ABSTRACT

Symmetry models are applied to square contingency tables in which the row and column variables have same categories. If the variables of the square contingency tables have the ordinal scale, Diagonal Parameter Symmetry model (DPS) tests whether the two diagonals that are below and above the main diagonal have equidistant from the main diagonal. The DPS model assumes that odds depend only on the distance between the diagonal containing the cell and the main diagonal. This model uses the distances for each category and is equivalent to the null association model over the sub-tables for each distance. This paper proposes a new model based on Euclidian distance of the categories. Father's and son's occupational mobility data in Britain is used as a real data example.

Key Words: *Square contingency tables, symmetry models, diagonal parameter symmetry model*

Differential Network Analysis with Partial Least Square Regression

Aylin Alin
Faculty of Science,
Department of
Statistics,
Dokuz Eylül
University,
İzmir, Turkey
aylin.alin@deu.edu.tr

Ayça Ölmez
The Graduate School
of Natural and
Applied Science,
Department of
Statistics,
Dokuz Eylül
University,
İzmir, Turkey
olmezayca@gmail.com

Gökhan Karakulah
Izmir International
Biomedicine and
Genome Institute,
Dokuz Eylül
University,
İzmir, Turkey
gokhan.karakulah@de
u.edu.tr

Aslı Suner
Medical Faculty,
Department of
Biostatistics and
Medical Informatic,
Ege University,
İzmir, Turkey
asli.suner@ege.edu.tr

ABSTRACT

Differential Network Analysis (DiNA) is a popular bioinformatics method, which provides information about complex biological processes, genomic associations, and the topological structure of high throughput gene expression datasets. DiNA aims to detect significant differences in network's topologies under different biological conditions. The first step is to build a network for each condition. For this purpose, there are commonly used statistical methods for calculation of connectivity scores such as; Correlation, Principal Components, Ridge Regression and Partial Least Square Regression (PLSR).

PLSR is highly recommended method for high throughput gene expression datasets that generally have large number of genes and relatively small number of samples. Furthermore, it can handle noise, missing values and collinearity problems. Once we calculate the connectivity scores, networks are compared in three different ways: 1) a test for connectivity of a single gene in two network, 2) a test for connectivity of a group of "interesting" genes, and 3) a test for connectivity in the overall modular structure.

Herein, we will use PLSR to obtain the connectivity scores. DiNA will be performed for examining the connectivity of a single gene, a set of genes and for overall network structures on fore-, mid- and hindbrain regions of a developing mouse brain in the embryonic period. Statistical analysis will be performed using R statistical software.

Key Words: *High Throughput Gene Expression Data, Gene Network, Differential Network Analysis, Partial Least Square Regression*

Directional k-means Algorithm Using Signed Angular Distance

Orhan KESEMEN

Buğra Kaan TIRYAKI

Özge TEZEL

Eda ÖZKUL

Department of Statistics and Computer Sciences,
Karadeniz Technical University, 61080, Trabzon, Turkey

okesemen@gmail.com bugrakaantiryaki@gmail.com ozge_tzl@hotmail.com eda.ozkul.gs@gmail.com

ABSTRACT

Clustering methods are widely used in many fields. Clustering is to group objects in a dataset according to similarities. In the literature, several clustering algorithms were improved. The k-means algorithm is one of the most widely used algorithm in the cluster analysis. The assignment mechanism of the k-means clustering algorithm, which is one of the unsupervised learning methods, allows each data point to belong to only one cluster. This is because, it is a crisp clustering algorithm and it is based on the idea that the centroid represents the cluster. The k-means clustering algorithm has been improved for clustering linear data set and it cannot be used directly to cluster directional data sets. Directional data are defined on a p-dimensional sphere. Angles, orientations, months can be exemplified as directional data.

The clustering algorithms calculate the similarity measure based on the distances between the data. In the literature, several different distance measures have been defined in the analysis of directional data. These distance measures are calculated with the help of trigonometric functions and for this reason it gives approximate results. In this study, the k-means clustering algorithm is adapted to the directional data using the angular difference. The k-means clustering algorithm has many disadvantages. The most important of these is that it is not stable. The most important of these is that it is not stable. The proposed circular k-means clustering algorithm prevents the stability problem. In addition, the k-means clustering algorithm is a superior algorithm in terms of the average computation time, when it is compared with the other clustering algorithms in the literature. As a result, the proposed directional k-means clustering algorithm is an effective method both in terms of the stability and the average computation time.

Key Words: *Clustering, Directional k-means Clustering Algorithm, Angular Distance.*

Discriminating between the Lognormal and Weibull Distributions under Progressive Censoring

Coşkun Kuş
Selçuk University
coskun@selcuk.edu.tr

Ahmet Pekkör
Necmettin Erbakan University
istatistik@hotmail.com

İsmail Kınacı
Selçuk University
ikinaci@selcuk.edu.tr

ABSTRACT

In this paper, the discriminating is studied for testing lognormality to Weibull alternative. The progressive censored sample is considered. Some critical values are simulated and power analysis is performed to investigate the performance of introduced procedure. The power analysis indicates that the new test has reasonable performances.

Key Words: *Discrimination, lognormal distribution, power analysis, simulation, progressive censoring*

Dynamic Measurement of Achieving Europe 2020 Goals in Smart Development – Turkey and European Union Countries

Andrzej Sokołowski

Cracow University of Economics,
Poland

sokolows@uek.krakow.pl

Małgorzata Markowska

Wrocław University of Economics,
Poland

malgorzata.markowska@ue.wroc.pl

ABSTRACT

The relative composite measure for assessing the degree of realization of Europe 2020 strategic goals in the field of smart development has been proposed in the paper. Global trimmed standardization, weighing and distance matrix calculation leads to the application of Multidimensional Scaling approach. Global standardization means that the average and standard deviation are calculated for the whole studied period, and not separately for each year. Trimming set the maximum values at goals, so variables cannot “borrow” the achievements from the others. Two points are added to the data: a starting point, and a final goal based on specific goals. The proposed measure is calculated as percentage ratio of a distance to a starting point and a distance between starting point and a goal. The empirical example deals with EU countries and Turkey, covering 2005-2016 period. Time series of composite measure make possible to study the disparities between countries. Forecasts based on trend extrapolation predict the year in which particular country is going to achieve the final goal, assuming the now observed speed of development.

Europe 2020 goals in smart development have been established in two regimes – one common for the whole European Union and the other, specific for each country, established by this country authorities. Country goals can be changed in time. The current study compares Turkish data with EU countries individual goals as well as joint EU goals.

Key Words: *European Union, Turkey, Multidimensional Scaling, Composite Indicator*

Effect of ionic polymers in the targeted drug delivery applications as the coating materials on superparamagnetic iron oxide nanoparticles

Maide Gökçe Bekaroğlu
Istanbul Technical University,
Faculty of Science and Letters,
Dept. of Physics
bekaroglum@itu.edu.tr

Sevim İşçi
Istanbul Technical University,
Faculty of Science and Letters,
Dept. of Physics
iscisev@itu.edu.tr

ABSTRACT

Magnetic nanoparticles such as iron oxide nanoparticles, are suitable for targeted drug delivery applications because of their biocompatibility in moderate doses, ease of surface modification, known metabolic pathways, variety of their sizes, magnetic properties and their ability to be manipulated upon application of external magnetic field. Targeted delivery can significantly reduce systemic toxicity and increase local concentration of chemotherapeutic drugs. Even though, iron oxide nanoparticles are highly applicable for targeted delivery, there are several handicaps to use iron oxide nanoparticles as drug delivery nanoparticles in their pure form as they could be toxic to healthy cells, colloiddally instable and low drug loading capacities. In order to eliminate these handicaps IONPs can be coated using biocompatible biopolymers to form core-shell structures. Biocompatible biopolymers are commonly used for functionalization of nanoparticles due to their non-toxic nature and ability to modulate physical and chemical properties (surface charge, etc.). Biopolymers also improve stability, enable high amounts of drug loading and provide protection of drugs for core-shell structures. In this study, three different biopolymers, hydroxyl ethylene cellulose (HEC), nano crystalline cellulose (NCC), and polyvinyl pyrrolidone (PVP) were studied in order to modify and cover the surfaces of superparamagnetic Fe_3O_4 nanoparticles. HEC, NCC and PVP were chosen based on their ionic charges which are cationic, anionic and nonionic, respectively. Magnetization, magnetorheological effects, toxicity to healthy cells, colloidal instabilities, drug loading capacities and also drug delivery to cancerous cells of the biopolymer covered Fe_3O_4 nanoparticles were compared in terms of the ionic character of the coating polymers. The results showed that cationic charges reduced cell viability for both normal cells and cancerous cells but reduction in normal cell viability was not significant. Anionic polymer coatings of the nanoparticles resulted in increased drug adsorption values, but anionic polymers could not completely cover the surfaces of the Fe_3O_4 nanoparticles which showed toxic effects on healthy cells. Moreover, even though nonionic polymers could completely cover the surfaces of the Fe_3O_4 , cancer drug precipitated on nonionic polymer covered surfaces when loading and this precipitation behavior decreased the drug delivery efficiency of the particles.

Key Words: *Superparamagnetic nanoparticles, ionic polymers, surface modification, drug delivery,*

Effect of Sample Size on Learning Performance in Sentiment Analysis

Adem Doğaner
Kahramanmaraş Sutcu Imam
University,
Faculty of Medicine,
Department of Biostatistics and
Medical Informatics
adem_doganer@hotmail.com

ABSTRACT

One precondition for a scientific research is a well-planned research. Optimal size of sample size is one of the basic requirements of a scientific research. Optimal sample size can be determined by power analysis in many research designs, such as prevalence studies, group comparisons or clinical trials. These methods can not be used to determine sample size for some research design and data types. For instance, in text mining methods, sample sizes can not be calculated with these methods. Text mining methods are data mining methods that classify texts according to their features and extract information from these texts. Basic data source is words, in text mining. For this reason, sample size determination methods which can be calculated with quantitative criteria can not produce successful results in text mining studies.

Sentiment analysis methods analyzes words and classifies content according to text. In sentiment analysis, the model needs to learn how to classify decisions according to words and texts. Sentiments or decisions are trained according to the texts and words for learning the model.

In this study, In order to develop a model that can perform directly sentiment analysis from texts, effect of sample size on learning were researched. New approaches for optimal sample size determination in sentiment analysis were studied. In this study, the opinions of persons about skin cancer were used as a data set.

As a result in this study, Alternative method has been suggested in which optimal sample size determination for sentiment analysis. On the other hand, the importance of the sample size has been revealed in model learning.

Key Words: *Learning Curve, Sentiment Analysis, Sample Size*

Efficiency Analysis of Ankara Rail and Light Rail System Stations

Harun Kınacı
Erciyes University,
Faculty of Economic and
Administrative Sciences,
Department of Business
hkinaci@erciyes.edu.tr

Abdullah Orman
Ankara Yıldırım
Beyazıt University,
Vocational School,
Department of
Computer Programming
aorman@ybu.edu.tr

ABSTRACT

For cities with high population density, the most convenient form of public transportation is based on rail systems. Furthermore, these systems operate with zero emissions so there are no direct harm to the ambient environment. Besides, underground rail systems improve visual environment around station areas and positively affect traffic circulation. However, installation of these systems is very costly. In this context, importance of preliminary feasibility studies increases and they should be done with utmost care. Besides, improving operational productivity should also be employed considering different elements of rail transit supply and demand.

In this study, an efficiency analysis of subway and light rail (Ankaray) system stations in Ankara is carried out by Data Envelopment Analysis (DEA). It is the most common method used in performance or efficiency analysis. DEA can be applied to almost all cases where the input and output variables are. DEA is a non-parametric method based on mathematical modeling. For this analysis, input variable is population of station area and output variables are total number of passengers alighting and total number of passengers boarding. The data used for the analysis are obtained under the Ankara Transportation Master Plan. By using DEA method, a model is set up to maximize outputs by keeping input constant. However, the efficiency scores obtained from the DEA are bias. When the number of decision-making unit is small, higher bias values are obtained. Therefore, both DEA efficiency scores and Bootstrap DEA efficiency scores were estimated accordingly. In the end, we obtained unbiased efficiency scores by Bootstrap DEA method.

Finally, the results obtained by both methods are shown in tables and graphs and the results are compared and interpreted.

Key Words: *Data Envelopment Analysis, Bootstrap, Efficiency, Rail-Light Rail System*

Efficiency Evaluation of Ankara Public Transport Bus Lines Using DEA

| | |
|--|--|
| <u>Harun Kinacı</u> Erciyes University, Faculty of Economic and Administrative Sciences, Department of Business hkinaci@erciyes.edu.tr. | <u>Hayri Ulvi</u> Gazi University, Faculty of Architecture, Department of City and Regional Planning hayriulvi@gazi.edu.tr. |
|--|--|

ABSTRACT

Public transport is an important service supplied by municipalities. Majority of public transport system in metropolitan areas is based on bus systems in Turkey. Maintaining public transport supply based on bus system in large metropolitan areas has important economic consequences as cost of providing the system increases in sprawling urban areas. For this reason, it is very important for the local administrators to effectively manage urban bus transportation. In this study, bus transit performance of the Ankara EGO bus lines was examined by Data Envelopment Analysis (DEA). The most common method used to measure the effective utilization performance of the available resources in Decision-Making Units (DMU) with certain input-output is DEA. It is a nonparametric method based on mathematical modeling. Performance of the Ankara EGO bus lines were analyzed by two main models of DEA: Charnes, Cooper, Rhodes (CCR) Model and Banker, Charnes, Cooper (BCC) Model. Both models use data collected during Ankara Transportation Master Plan Project.

The analysis phase consists of two stages. In the first stage, bus lines divided by EGO into five regions examined individually; relative efficiency scores were obtained for each bus line. In the second stage, Ankara was taken as a whole and relative efficiency scores were obtained for all bus lines in five regions. The output variables are the average number of passengers and the average occupancy rate of passengers. Service frequency, line length, and time of services are input variables. The efficiency scores obtained by the DEA are compared and the results are interpreted accordingly.

Finally, the obtained activity scores are put into the geographic information system (GIS) and the efficiency performances are shown on the routes of the lines.

Key Words: *Data Envelopment Analysis, Urban Transportation, Efficiency*

Efficient Algorithms to Compute Signatures of Binary Systems

Gökhan Gökdere

Department of Statistics,
Faculty of Science,
Fırat University,
Elazığ, Turkey

g.g.gokdere@gmail.com

Fatih Aydın

Department of
Computer Engineering,
Faculty of Engineering,
Fırat University,
Elazığ, Turkey

fataycomputers@gmail.com

Ayşenur Sönmez

Department of Statistics,
Faculty of Science,
Fırat University,
Elazığ, Turkey

sonmezaysenur23@gmail.com

ABSTRACT

The signature of the system is a vector that is defined based on the failure of the components in the system and depends only structure function of the system when the component lifetimes are independent and identically distributed with a common continuous distribution. The concept of system signature can be used to calculate the system reliability and expected lifetime, as well as to compare different systems. Many different binary systems are presented in the literature. Some of them are coherent system, k -out-of- n : F system, weighted k -out-of- n F system, consecutive k -out-of- n : F system, consecutive weighted k -out-of- n : F system, m -consecutive- k -out-of- n : F system, consecutive- k -within- m -out-of- n : F system and combined m -consecutive- k -out-of- n : F & consecutive- k_c -out-of- n : F system. Our aim in this study is to get algorithms to calculate the signatures of the abovementioned systems. Algorithms are based on a logical approach and provide an easy way to obtain the system's signature when given the system's structure function or the parameters required for the system.

Key Words: *System Signature, Logical Approach, Algorithm*

Empirical estimator of the Conditional Tail Moment

Laidı Mohamed
National High School
of Technology,
Algiers, Algeria
LRDSI Laboratory,
Blida University,
Blida, Algeria
mohamed.laidi@enst.dz

Rassoul Abdelaziz
National High School
of Hydraulics,
Blida, Algeria
a.rassoul@ensh.dz

Ould Rouis Hamid
University of Blida
1, Blida, Algeria
ouldrouis@hotmail.com

ABSTRACT

The Conditional Tail Moment (CTM) is defined as the moment of order $a \geq 0$ of a distribution above the upper α -quantile where $\alpha \in (0,1)$. In this paper, we propose an empirical estimator of the CTM, this estimator it permits to give an empirical estimator for all risk measures based on conditional moments such as Value-at-Risk, Conditional Tail Expectation, Conditional Value-at-Risk or Conditional Tail Variance. We explore the asymptotic properties of this estimator, in particular his consistency and we construct a confidence interval for it. We finish our work by a simulation studies that illustrate the performance of this estimator and we discuss the obtained results.

Key Words: *Statistic Order, Empirical Quantile, Conditional Tail Moment, Consistency, Confidence Interval*

Ensemble Learning Methods as In-silico Model for Prediction of Mutagenicity

Enis Gümüştas Ayça Çakmak Pehlivanlı
Mimar Sinan FA University, Mimar Sinan FA Universtiy,
Statistics Dept., Statistics Dept.,
Bomonti, İstanbul Bomonti, İstanbul
20172107001@ogr.msgsu.edu.tr ayca.pehlivanli@msgsu.edu.tr

ABSTRACT

Along with developing technologies, in-vitro experiments that are conducted on outside of living organisms and in-vivo experiments that are conducted on the living organism have begun to leave their place to statistical and computational methods developed in the computer environment without requiring laboratory experiments. Generally, these methods, called in-silico, are capable of predicting and predicting candidate drug molecules prior to in-vitro and / or in-vitro testing. Obtaining information about drug molecules by using an accurate in-silico approach can guide to decide whether the laboratory experiments should be conducted. Foreseeing the design of the tests to be performed can provide advantages such as using less experimental animals, predicting the chemical concentration to be used, reducing time and cost.

Today, various toxicity tests are used in the legal and ethical regulation of chemicals (drug molecules, food additives, cosmetics, etc.). Among the toxicity tests, mutagenicity defined as a genetic change that can occur due to an agent, has an important place. In particular, it is a prerequisite for candidate drug molecules to have no mutagenic effects in order to continue their clinical trials. Mutagenicity screening studies must be maintained in multiple steps which are consisted of in-vivo and in-vitro tests that are not enough alone.

The in-silico approach has been used with statistical learning algorithms as the first step in order to improve the mutagenicity determination process in general. This approach has been applied to the set of molecules containing mutagenicity information obtained by experiments and promising classification successes have been obtained.

Within the scope of the study, data sets of molecules known as Bursi and Benchmark were merged in the literature and the properties of the molecules in the data set were calculated via the Molecular Operating Environment (MOE) program. Statistical learning algorithms (AdaBoost, ExtraTrees and Random Forest) were applied on the obtained data set of 10835 x 193 and parameter selection was performed by grid search approach. All selections and applications were performed with 10-fold cross validation. As a result of the models established with the best parameters obtained, the selection of the variables was made according to the predicted levels of mutagenicity and the most effective 72 variables were obtained. The new database consisting of the selected variables has been applied to 19 different statistical learning algorithms and it has been decided to use seven (AdaBoost, Bagging, ExtraTrees, GradientBoosting, Random Forest, XGBoost, LGBM) ensemble classification algorithms that yield the best results. Using these algorithms, which increase the model performances by parameter optimization, accurate classification rates of mutagenicity ranging from 78% to 90% are obtained.

Key Words: Ensemble Learning, Statistical Learning, Feature Selection, Mutagenicity Prediction, In-Silico

Acknowledgement: This study was supported by Mimar Sinan FA University Research Foundation BAP2018-30.

Estimation of the Parameters of Two-Fold Weibull Model with Right Censored Reliability Data Using Genetic Algorithm

Erkut Tekeli

Department of
Computer Technologies,
Kozan Vocational School,
Çukurova University,
Kozan, Adana, Turkey
etekeli@cu.edu.tr

Güzin Yüksel

Department of Statistics,
Faculty of Science
and Letters,
Çukurova University,
Adana, Turkey
yguzin@cu.edu.tr

ABSTRACT

Mixture of distributions is often used for the analysis of reliability data in systems with different components. In this article, a new method has been developed in order to construct a model using two-fold Weibull mixture distribution with right-censored reliability data. The method is based on optimizing the parameters of the right-censored two-fold Weibull mixture distribution using genetic algorithm techniques. Parameters of the distributions to be used in the mixture model were determined by software developed in R language. An Aircraft component (Windshield) failure data, which are frequently used in the literature, were analysed using the proposed method, and the results are compared with other methods in the literature. Data contains failure and censor times. We showed that proposed method develops a more appropriate model than the other methods. In addition to real data analyse, a simulation study was conducted to compare the performance of the method for different sample sizes and different censor ratios.

Key Words: *Reliability, Two-Fold Weibull, Mixture Distribution, Genetic Algorithm, Right-Censored Data*

Estimating First Order Inclusion Probability Based on Percentile Ranked Set Sampling

Yusuf Can Sevil Tuğba Özkal Yıldız
Department of Statistics, Department of Statistics,
Dokuz Eylül University, Dokuz Eylül University,
İzmir, Turkey İzmir, Turkey
yusuf.sevil@ogr.deu.edu.tr tugba.ozkal@deu.edu.tr

ABSTRACT

Simple random sampling (SRS) is the most commonly used sampling method. Each of population units, u_k $k = 1, \dots, N$, appears in the sample of size n with the same probability n/N via SRS. This probability is called the inclusion probability and it is denoted by π_k , $k = 1, \dots, N$. If sample is selected from finite population, this inclusion probability is needed.

Ranked set sampling is proposed by McIntyre (1952). In finite population, some authors have discussed the performance of RSS method. Takahasi and Futatsuya (1988) showed that mean estimator based on RSS is unbiased. By using large set size, Patil et al. (1995) generalized the results of Takahasi and Futatsuya (1988). Ozturk et al. (2005) studied estimators based on RSS for population mean and variance. Three different sampling designs called Level-0, Level-1 and Level-2 are described by Deshpande et al. (2006). In each sampling design, sets are selected without replacement. If the all units in the set are replaced back into the population before drawing the next set, we obtain Level-0 sampling. Level-1 sampling is obtained if the measured unit in the set are not replaced back into the population before drawing the next set. Finally, if none of the units in the set are replaced back into the population before drawing next set, then we obtain Level-2 sampling.

Percentile ranked set sampling (PRSS) is suggested by Muttlak (2003) to estimate the population mean. In PRSS procedure, p th and q th percentile of each set are selected for full measurement, $0 < p < 1$ and $q = 1 - p$.

In RSS, inclusion probabilities for Level-0, Level-1 and Level-2 sampling designs have been investigated by some authors. Both first-order and second-order inclusion probabilities are given by Jafari Jozani and Johnson (2010) for Level-0 sampling design. Here, first-order and second-order inclusion probabilities are denoted by π_i and π_{ii} . Inclusion probabilities for Level-1 are investigated by Al-Saleh and Samawi (2007), Ozdemir and Gokpinar (2007, 2008), Gokpinar and Ozdemir (2010) and Frey (2011). Patil et al. (1995) studied these inclusion probabilities for Level-2 sampling design.

Ozturk (2014) estimated these inclusion probabilities under imperfect ranking. In our study, first order inclusion probability for Level-0, Level-1 and Level-2 sampling designs are investigated for PRSS. Also, a single auxiliary variable is used as ranking criteria. Thus, the ranking error can be controlled by correlation coefficient between interested variable and auxiliary variable. We illustrate the inclusion probabilities when the ranking mechanism is perfect and imperfect.

Key Words: *Inclusion Probability, Sampling Design, Percentile Ranked Set Sampling, Perfect and Imperfect Ranking*

Estimating the Total Number of Defective Items: A Case Study

| | |
|------------------------|------------------------|
| Ayda Amniattalab | J.B.G Frenk |
| Sabanci University, | Sabanci University, |
| Faculty of Engineering | Faculty of Engineering |
| and Natural Sciences, | and Natural Sciences, |
| Istanbul | Istanbul |
| ayda@sabanciuniv.edu | frenk@sabanciuniv.edu |

ABSTRACT

In this talk, we present the statistical analysis of two large item-level data sets obtained from a household manufacturing company in Turkey. Each data set contains some identifiers of two different household products as well as their sales and failure times. Based on the model presented in the talk on the total number of defective items, the statistical analysis of our data sets can be divided into two separate parts. We first discuss how the future sales of the two separate products can be estimated by means of linear regression techniques using some well-known parametric functions. Secondly, under the assumption that the repair does not change the age of the product, we estimate the failure times by means of the Maximum Likelihood Estimation method for censored data. In this method, we assume that the parametric class is given by the often-used class of Weibull distributions. We also present a non-linear optimization procedure designed for the class of Weibull distributions which guaranties that the optimal solution of the Maximum Likelihood function will be identified. Possible extensions of the above approach to different parametric classes are part of an ongoing Ph.D. research project.

Key Words: *Linear Regression, Maximum Likelihood, Censored Data*

Estimation of Bone Age from Radiological Images with Soft Computing Methods

| | | |
|--|--|--|
| <u>Nida Gökçe Narin</u> Department of Statistics, Faculty of Science, Muğla Sıtkı Kocman University nidagokce@yahoo.com | <u>İbrahim Önder Yeniçeri</u> Department of Radiology, Faculty of Medicine, Muğla Sıtkı Kocman University onderyeniceri@mu.edu.tr | <u>Gamze Yüksel</u> Department of Mathematics, Faculty of Science, Muğla Sıtkı Kocman University ngamze@mu.edu.tr |
|--|--|--|

ABSTRACT

Bone age refers to the degree of development of bones, and bone development varies from individual to individual. This is particularly problematic in judicial cases when determining the individual's criminal capacity or identifying the person. In pediatric clinics, early detection of diseases directly related to bone development, such as growth retardation or overgrowth, is important in order to improve future quality of life for children. For this reason, in this study, the performances of soft computing methods in estimation of bone age in children aged 1-9 years whose age is known are examined. The 1-9 age group includes the period when bones appeared and the developmental process was the fastest.

Soft Computing methods propose approximate solutions for problems that are encountered in everyday life, called real world problems, and which can not be solved mathematically. Because of the uncertainty inherent in the nature of bone age development, individual and individual variation, soft computing methods have been preferred in this study. In the scope of the study, radiographs of left hand wrists of 360 male and 389 female children in the radiology department of Muğla Sıtkı Koçman University Training and Research Hospital between the ages of 12-108 months were examined. Formations observed in this region during the age range from distal to proximal; 1. metacarpal epiphysis, Trapezium, Trapezoid, Capitatum, Hamatum, Scaphoideum, Lunatum, Triquetrum, Pisiform bones, radius and ulna head epiphyses.

The data obtained from the radiological images will be analyzed with soft computing methods and the results will be compared with the age of identity. By using fuzzy c-means, natural groups in the data will be determined and determined as a degree percentage of the fit between the estimated bone ages and the identity ages of the individuals. Feedback neural network models will generate a neural network that can predict the age of identity. This network model will then be used to estimate age in patients whose age is unknown. Finally, bone age prediction performances of all methods used will be compared with the age of identity and the Greulich-Pyle method.

Key Words: *Soft Computing, Bone Age Estimation, Clustering*

Acknowledgement: *This paper has been granted by the Mugla Sıtkı Kocman University Research Projects Coordination Office. Project Grant Number: 17/217 and title, "Esnek Hesaplama Yöntemleri İle Radyolojik Görüntülerden Kemik Yaşı Tahmini"*

Estimation of Demand for Motor Own Damage Insurance with Neural Networks

| | |
|--|---|
| Betül Zehra Karagül Department of Actuarial Sciences, Hacettepe University, 06800 Beytepe, Ankara betul.zehra@hacettepe.edu.tr | <u>Samet Gençgönül</u> Department of Mathematics, Atatürk University, 25240, Yakutiye, Erzurum sametgencgonul@atauni.edu.tr |
|--|---|

ABSTRACT

The fact that insurance demand can be estimated correctly will ensure that insurance companies make the right pricing. The accuracy of the estimations will also increase the validity of the planning. Motor own damage insurance is the line of business that Turkey's non-life insurance companies receive the highest premium level and maximum compensation payments. Artificial neural networks are parallel and distributed information processing structures that are inspired by the human brain, interconnected by weighted connections, and composed of processing elements, each having its own memory. In this study, the use of artificial neural networks was investigated in estimation of demand for motor own damage insurance.

Key Words: *Neural Networks, Motor Own Damage Insurance, Insurance Demand Estimate*

Estimation of The Reliability in Multicomponent Stress-Strength Model Based on Chen Lifetime Distribution Using Genetic Algorithm

Muhammet Burak Kılıç

Department of
Business Administration,
Mehmet Akif Ersoy University,
Burdur, Turkey
muhammet.b.kilic@gmail.com

Gökhan Gökdere

Department of Statistics,
University of Firat,
Elazig, Turkey
g.g.gokdere@gmail.com

ABSTRACT

A Chen lifetime distribution provides to fit real life data with bathtub-shaped or increasing failure rates. Some details on parametric probability distributions with bathtub-shaped failure rate function have received great attention in the literature. In this research article, we estimate multicomponent stress-strength reliability of a system when stress and strength variates are drawn from a Chen life distribution with different scale parameters and a common shape parameter. The reliability of such a system is estimated by maximum likelihood estimation using the genetic algorithm, which is a class of evolutionary algorithms and based on the principle of biological systems, Newton-Raphson, Nelder Mead and simulated annealing. The confidence interval for the parameters is constructed with bootstrap method for identifying search space in genetic algorithm approach. The comparisons of the reliability estimators are made by the bias and mean square error via Monte Carlo simulations. Simulation results show that the genetic algorithm approach seems to perform well in terms of the estimation of reliability.

Key Words: *Multicomponent Stress-strength Model, Chen Lifetime Distribution, Maximum Likelihood Estimation, Genetic Algorithms*

Estimation of the System Reliability for Generalized Inverse Lindley Distribution under Different Sampling Schemes

Fatma Gul AKGUL
Artvin Coruh University,
Department of Computer
Engineering, Artvin, TURKEY
ftm.gul.fuz@artvin.edu.tr

Keming YU
Brunel University,
Department of Mathematics,
London, UK
keming.yu@brunel.ac.uk

Birdal SENOGLU
Ankara University,
Department of Statistics,
Ankara, TURKEY
senoglu@science.ankara.edu.tr

ABSTRACT

The estimation of the stress-strength reliability $R = P(X < Y)$ is one of the most popular problems in statistical literature. Here, X represents the stress and Y represents the strength. Obviously, if $X < Y$, the system would fail. Otherwise, it continues to work. Therefore, R is the measure of the system reliability.

In this study, we consider the estimation of R when the stress and the strength are both independent generalized inverse Lindley (GIL) distribution based on different sampling schemes, namely, simple random sampling (SRS), ranked set sampling (RSS) and percentile ranked set sampling (PRSS). In the context of parameter estimation, we use maximum likelihood (ML) methodology. The performances of the ML estimators of R based on SRS, RSS and PRSS are compared via Monte-Carlo simulation study for different parameter settings and different sample sizes under both perfect and imperfect ranking assumptions. It is concluded from the simulation studies that the estimators of R based on RSS and PRSS demonstrate better performances than the corresponding estimator based on SRS. Furthermore, it is seen that the estimator of R based on PRSS is more efficient than the estimator of R based on RSS under the assumption of imperfect ranking. At the end of the study, a real data set is analyzed to implement the proposed methodologies.

Key Words: *Stress-strength reliability, Sampling procedures, Maximum likelihood, Imperfect ranking, Efficiency.*

Estimation of Turkish Court of Cassation Decisions via Machine Learning and Natural Language Processing Methods

Enes Dilber
University of Michigan,
Department of Statistics
enes@umich.edu

Melih Ağraz
METU, Department
of Statistics
melih.agraz@yahoo.com

ABSTRACT

Machine learning (ML) and natural language processing (NLP) methods, which are the fundamentals of artificial intelligence (AI), are advancing rapidly. Individuality and complexity in the field of law make technology transfer impossible in this field. However, the progress in ML and NLP has reached a level that will quite reduce this complexity and individuality. For this reason, various studies have been carried out in this field in recently. Although developing technology enables NLP to be more functional, it is not possible to work efficiently on NLP in Turkish which is an agglutinating language. In this study, we tried to develop a system to predict Turkish Court of Cassation decisions using ML and NLP algorithms. Although our system uses simple ML algorithms, it was successful in obtaining 90.70% accuracy in ML algorithms and 94.50% in artificial neural networks on the test set in the estimation of Turkish Court of Cassation decisions on a difficult level for NLP like Turkish.

Key Words: *Machine Learning, Neural Network, Natural Language Processing, Law, Court of Cassation*

Acknowledgement: *Our research is supported by TUBITAK (The Scientific and Technological Research Council of Turkey) with project number 2170347. We thank to TUBITAK for supporting to our research project.*

Evaluating The Gage Repeatability and Reproducibility for Industries

Ali Rıza Firuzan

Dokuz Eylül University

aliriza.firuzan@deu.edu.tr

Ümit Kuvvetli

ESHOT

umit.kuvvetli@gmail.com

Atakan Gerger

PFW Aerospace GmbH

atakangerger@gmail.com

ABSTRACT

The measurement system analysis is the determination of the efficiency of our present measurement system where measurement is required. Gage R&R identifies the variation, originated from the measurement system, in order to figure out the real duration variation, and separates it from the system variation. The fundamental reason for this separation is that the total variation arises both from the duration variation and from the measurement system variation. The ultimate purpose here is to develop a measurement system that best serves the purpose. Many manufacturers are presently using tools like statistical process control and the design of experiments to monitor and improve product quality and process productivity. However, if the data collected are not accurate or precise, they do not represent the true characteristics of the part or product being measured. Even organizations use the quality process; it is desired to reduce the variation in the measurement system that negatively affects the improvement process. Therefore, it is very important to have a valid quality measurement study beforehand to ensure the part or product data collected are both accurate and precise and the power of SPC and DoE are fully realized. Accuracy, in other words, no bias, the function of calibration is performed before a correct measurement study of the precisions of the gage and its operators. In order to reduce the variations in a process, it is necessary to identify the sources of the variation, quantify them and to have an understanding about the proper operation of the gage being used for collecting the measurements. In operating a gage, measurement error can be contributed to various sources like within-sample variation, measurement method, the gage/instrument used for measurement, operators, temperature, environment and other factors. Therefore, it is necessary to conduct a study on measurement system capability. This study is termed as Gage Repeatability and Reproducibility (GRR) study or Gage Capability Analysis. In this study, it was decided to examine the measurement system although the process is under control within a manufacturing company as various problems about quality emerge. Then, the measurement system was analyzed and results obtained were shared.

Evaluation of Epistasis Effect in Genome-Wide Association Studies with Ant Colony Optimization Technique

Havva Didem Celikcan Bahar Taşdelen
Biostatistics and Medical Biostatistics and Medical
Informatics Department, Informatics Department,
Mersin University, Turkey Mersin University, Turkey
didemcelikcan@gmail.com bahartasdelen@gmail.com

ABSTRACT

In recent years, novel approaches about the genetic causes of diseases are identified with genome-wide association studies (GWAS). The relationship between only one single nucleotide polymorphism (SNP) and disease was examined in many studies. However, it should also be considered that a SNP may be associated with another SNP. The interaction between SNPs is named epistasis. Classical methods are not suitable for epistasis analysis. So, it is necessary to use combinatorial optimization methods such as ant colony optimization. In this study, we investigate the success of ant colony optimization (ACO) method in epistasis. A simulation study is conducted to reveal how the inadequate sample size, minor allele frequency (MAF) and SNP numbers in the groups of polymorphism studies affect ant colony optimization results. Both the results obtained by the classical method and the results obtained by the optimization method were compared. The consistency of the results was assessed by the Kappa statistic.

Key Words: *Single Nucleotide Polymorphism, Genome-wide Association Studies, Ant Colony Optimization*

Examining the Efficiency of Hydroelectric Power Plants in Turkey

Metehan SONMEZ

Yıldız Technical University
Davutpasa Mah.
Davutpasa Caddesi
3420 Esenler-İstanbul
metehan.sonmez.03@gmail.com

Doğan YILDIZ

Yıldız Technical University
Davutpasa Mah.
Davutpasa Caddesi
3420 Esenler-İstanbul
dyildiz@yildiz.edu.tr

Mehmet Samil GUNES

Yıldız Technical University
Davutpasa Mah.
Davutpasa Caddesi
3420 Esenler-İstanbul
msgunes@yildizedu.tr

ABSTRACT

Hydroelectric energy has become an energy source preferred by many countries, owing to the development of technological conditions and its qualities as a renewable energy source. When hydroelectric energy outputs of countries are examined, China emerges as the country which has taken the first place. China is followed by India and Canada respectively. Turkey takes fourth place in hydroelectric energy output in Europe following Russia, Norway, and Sweden. Production from hydroelectric energy constitutes 20% of Turkey's total energy production. As it is the case worldwide, in our country the major part of energy production is met from energy sources such as petroleum and natural gas. However, we appear a foreign-dependent country in acquiring these two types of energy sources. This brings our country to a difficult situation in financial and environmental aspects. Consequently, the necessity for hydroelectric energy comes to the forefront. The raw material of hydroelectric energy is water. However, climate change causes the decrease in water sources in Turkey as well as in many parts of the world. As a matter of fact, water sources in our country must be used effectively and reasonably as much as possible. In this particular study, we are going to combine the different datasets which had taken from meteorology and stream gauging stations, and consequently, we will apply the multivariate statistical methods to create the right model for examining the efficiency of hydroelectric power plants in Turkey.

Key Words: *Statistical Hydrology, Multivariate methods, Hydroelectric energy, Statistics in water resources*

Examining the Matching Algorithms in Propensity Score Analysis

Meral Yay
Mimar Sinan Fine Arts
University
Faculty of Science and Letters
Department of Statistics
İstanbul/Turkey
meral.yay@msgsu.edu.tr

Elif Çoker
Mimar Sinan Fine Arts
University
Faculty of Science and Letters
Department of Statistics
İstanbul/Turkey
elif.coker@msgsu.edu.tr

Elif Özge Özdamar
Mimar Sinan Fine Arts
University
Faculty of Science and Letters
Department of Statistics
İstanbul/Turkey
ozge.ozdamar@msgsu.edu.tr

ABSTRACT

In observational studies when a binary outcome is present, the estimation of the treatment effects should be examined attentively. The reason for this consideration is about the assigning process of the treatments to the individuals which is not random in most cases. As a result, the absence of randomization leads to biased estimates. A conventional way to adjust this biasness is by using logistic regression models. In recent years, propensity score analysis is used as an alternative method to these models. A propensity score is the conditional probability of assigning an individual to a particular treatment or control group given a set of certain characteristics. Propensity score matching is a statistical method that a treatment group is matched with a control group based on each groups' propensity score. In this study, the aim is to investigate the effects of different types of matching algorithms in propensity score analysis. The comparison of these matching algorithms are analyzed using R Project.

Key Words: *Propensity Score, Propensity Score Matching, Propensity Score Analysis*

Experimental Design Applications in Chemistry

Oğuz Akpolat

Muğla Sıtkı Koçman University,
Faculty of Science,
Chemistry Department
Muğla, Turkey
oakpolat@mu.edu.tr

ABSTRACT

Chemometrics is described in general as to extract the more useful information from the chemical data obtained experimentally by various methods applying some mathematical techniques to these data. Today, studies in many fields such as chemistry, biochemistry, earth and environmental sciences are done experimentally and all analysis of semi-finished and finished products or raw materials for quality control are carried out in the laboratories. Establishment of test strategies for experimental design, analysis of experimental design models and optimization of the experimental factors of a reaction selected as a practical application for experimental design are the main subjects of this study.

Here it was summarized creation of experiment strategies in order to optimize the effective operation factors on chemical reactions, mathematical solution techniques of the response surface functions to define the relationships between the factors and the experimental result affected by the factors, computer programming for the optimization of the reaction parameters and usage of some related software, statistically evaluation for the experimental results and optimization of a selected reaction as an experimental design application.

Key Words: *Experimental Design, Statistics, Optimization*

Extreme Value Theory on Valuation of Actuarial Risks

Bükre Yıldırım Külekci
Middle East Technical
University, Institute of Applied
Mathematics, Ankara/Turkey
bukre@metu.edu.tr

Uğur Karabey
Hacettepe University, Actuarial
Sciences, Ankara/Turkey
ukarabey@hacettepe.edu.tr

A. Sevtap Selçuk-Kestel
Middle East Technical
University, Institute of Applied
Mathematics, Ankara/Turkey
skestel@metu.edu.tr

ABSTRACT

The low probability and high impact events are in the scope of insurance and should be analyzed throughly. In this study the dependence in the heavy tailed multivariate actuarial data is investigated. First we introduce the Extreme Value Theory (EVT) with Peaks over Threshold (POT) method. This method assumes that after a certain threshold exceedances follow a Generalized Pareto Distribution, and deals with the behavior specifically on the tail of a distribution without making any strong assumptions about the underlying distribution. Then from the perspective of risk, the inclusion of copula allows discovering the above threshold nonlinear dependence structure among multivariate data. Second by allowing the variability in the shape and scale parameters we incorporate a dynamic model with ARCH-GARCH time series process which captures the current risk. By using the threshold of POT method as an unknowns we obtain more realistic estimates for the capital risk measure Value at Risk. Finally using Monte Carlo simulation, VaR premiums are compared under the cases of dependency and independency.

Key Words: *Extreme Value Theory, Peaks Over Threshold, Time Series, Value-at-Risk, Copula*

Factors Affecting the Job Choice of Industrial Engineering Graduates

Damla Beril Özyazgan

APLUS Energy Investment Consulting
Technology & Trade Co. Ltd.
Süleyman Seba Caddesi No: 48
Beşiktaş 34357, Istanbul, Turkey
ozyazgand@itu.edu.tr

Erkan Işıklı

Istanbul Technical University
Faculty of Management
Industrial Engineering Department
Maçka 34367, Istanbul, Turkey
isiklie@itu.edu.tr

ABSTRACT

Industrial engineering undergraduate programs have been favored by prospective students in Turkey mainly since their newly graduates can find employment relatively easily across a wide range of industries and sectors, from energy and finance, to automotive and health. As getting closer to graduation, undergraduate students of this program need to make a choice among various employment opportunities. This is a hard decision that may have a long term effect on their career. Hence, they seek jobs that suit them best, but the abundance of alternatives may sometimes complicate their decision making process. On the other hand, employers always look for the most qualified candidates for jobs. Thus, they should be aware of the factors such quality candidates consider when making a job choice.

This exploratory study aims to reveal the factors affecting the job choices of industrial engineering graduates in Turkey. After a thorough review of the related literature, we conducted an unrestricted, self-selected web survey consisting of 50 questions in the period March 2018–June 2018 among 115 industrial engineers. A total of 27 questions measured – in two parts – the respondents' intention to leave their current job (10 items) and the factors they consider when making a job choice (17 items) on a 5-point Likert scale. The remaining questions gathered information on the demographics, education level, and current employment of the respondents.

An unweighted least squares factor analysis with varimax rotation was separately conducted on the scores of the two item sets of 115 respondents. Fit indices for the first item set suggested that four distinct factors, accounting for 16.4%, 15.3%, 11.7%, and 9.6% of the shared variance, provided an adequate representation of the data and 18% non-redundant residuals with absolute values greater than 0.05. Similarly, six distinct factors were extracted for the second item set, accounting for 10.4%, 9.4%, 9.1%, 7.2%, 6.7%, and 6.6% of the shared variance. This six-factor solution provided an adequate representation of the data and returned 15% non-redundant residuals with absolute values greater than 0.05. The results also indicate that *intrinsic motivation* and *social influence and extrinsic motivation* are significantly more important for women compared to men when making a job choice. In addition, two monthly income groups (2501–3500 TL and 3501–4500 TL) were also found to significantly differ in *intrinsic motivation* when compared to the monthly income group 4501–5500 TL.

Key Words: *Job Choice, Industrial Engineering, Exploratory Research, Factor Analysis*

Financial Stress Tests for Measuring Vulnerability of Turkish Economy

Tuğçe Leventoğlu

Dokuz Eylül University, İzmir
tugce.leventoglu@gmail.com

Esin Firuzan

Dokuz Eylül University, İzmir
esin.firuzan@deu.edu.tr

ABSTRACT

Financial stress tests are described as financial techniques to assess the vulnerability or instability of the monetary system. Financial institutions consult these tests to evaluate the risks of shocks or structural changes in the key macro economic variables for the Turkish financial system.

Financial stress index (FSI) is a global indicator variables which is used to aware of potential instability of financial markets. FSI contains three main stress sources: 1. Bank-related stress, 2. The securities-related stress and the exchange rate stress.

In this paper, we analyzed the linkage between structural breaks in economic activity and financial stress using Vector Autoregressive Regression (VAR) models. Performing VAR analysis is a relatively simple way to capture possible nonlinearities such as asymmetric reactions to shocks or the existence of structural breaks. According to the results, financial stress index is affected from the shocks to inflation and foreign exchange rate.

Key Words: *Financial Stress Tests, Instability, Inflation, Exchange Rate, Shocks*

Forecast Sales With Data Mining In The Public Sector

Vildan Yıldırım
Hacettepe University,
Ankara
vildansenel_@hotmail.com

Ümit Cengiz Uysal
Yıldırım Beyazıt University,
Ankara
ucuysal@gmail.com

ABSTRACT

In today's economic conditions, states must cooperate with the private sector in economic situations where it is not possible for them to offer public services. At this point where states and private sector meet, the concept of public procurement is emerging. Public procurement has changed day by day from 1925 four times and is in its final form under the "Public Procurement Law". Public procurement, which plays an important role in improving the quality of services provided by the state, indirect savings, encouraging innovation in the private sector, closing the current account deficit and reducing corruption, undoubtedly has a very important share for the country's economies.

In this study, it was aimed to forecast the sales of State Supply Office, which is authorized as central procurement institution within the scope of "exception" concept and to act as intermediary between private sector and public sector, in order to sustain public procurement effectively and efficiently with various parameters. These data, forecasted by data mining techniques, will be integrated into the decision support system and the strategic goals of the institution will be realized and followed.

In addition, it was planned to model the sales forecasts using analytical methods using the historical data of 2007-2017 and to present the results comparatively and determine the most appropriate method. Thus, it is considered that the key performance indicators included in the institutional strategic plan will be provided and academic contribution will be made to increase the efficiency of public procurement sector.

Key Words: *Applications of Data Mining Algorithms, Public Procurement, Forecasting, Modeling and Analytics, State Supply Office*

Full Automatic Evaluation of SeismicRefraction Data

Orhan Kesemen

Department of Statistics and
Computer Sciences,
Karadeniz Technical University,
61080, Trabzon, Turkey
okesemen@gmail.com

Sibel Ertürk

Department of Statistics and
Computer Sciences,
Karadeniz Technical University,
61080, Trabzon, Turkey
sibelerturkse@gmail.com

Ebru Naz

Department of Statistics and
Computer Sciences,
Karadeniz Technical University,
61080, Trabzon, Turkey
ebrunaz91@gmail.com

ABSTRACT

Seismic methods are based on the physical principles related to the propagation of elastic waves in a layer in analyzing the geological layers underground. Seismic methods are divided into two classes according to the signal pathways followed by seismic waves emitted from the source. There are Seismic Reflection and Seismic Refraction. Seismic refraction is a preferred method for determining structure of the shallow underground. Seismic refraction data should be evaluated by an expert engineer to determine primarily the first break and then the layer velocities after seismic data are collected from the ground.

The first break picking of seismic refraction data are usually determined manually by expert engineers. However, the technological developments in recent years have opened the way for automatically determining the first break picking by using the cross-correlation method via computers. The obtained data from the first break picking form the travel-time curve. With the help of travel-time curves, expert engineers manually determine the underground model by evaluating the layer structure of the underground and the change of wave velocities. These analyzes require time and cost as well as labor. It is also possible that human-induced errors will occur.

Fast and reliable methods are needed to remove human-based evaluations for analyzing seismic refraction data obtained from seismic fracture measurements. For this purpose, the analysis of seismic refraction data is automatically performed in this study using a piecewise linear regression analysis. It is aimed to use the fuzzy c-regression method which will remove the manual reading from the travel-time curve with the obtained first break. In this study, the intersection point of the layer was determined by using the fuzzy c-regression and the number of the layers was determined by using the fuzzy validity indices.

Key Words: *SeismicRefraction, Fuzzy c-Regression, Piecewiseregression*

Fuzzy ELECTRE I Method for Plant Location Selection: A Case Study

Kezban Bulut
Kırıkkale University
Department of
Industrial Engineering
kezbanbulut@yahoo.com

Suna Çetin
Kırıkkale University
Department of
Industrial
Engineering
sunacetin@kku.edu.tr

Metehan Zoğal
Kırıkkale University
Department of
Industrial
Engineering
metehanzogall@gmail.com

Kemalcan Bediz
Kırıkkale University
Department of
Industrial Engineering
kemalbediz@gmail.com

ABSTRACT

Businesses must make decisions under limited resources and infinite needs. Selection the most appropriate place for locating a plant directly affects the efficiency of a company and also it is a very important decision for businesses. When there are multiple alternatives and criteria in the decision making process, this situation called multi-criteria decision making. In this paper; facility location selection problem which is one of the multi criteria decision making problems has been addressed. As the conventional methods for facility location selection problems are inadequate for dealing with the imprecise or vague nature of linguistic assessment, by this reason Fuzzy ELECTRE I (Elimination Et Choix Traduisant la Réalité) method is applied. Fuzzy ELECTRE I is one of the effective fuzzy multi-criteria decision-making methods to solve the uncertainty of the concepts associated with decision-makers' decisions. In the case study; three alternatives and seven criteria that affect the decision are defined by expert opinions and literature. We used Fuzzy ELECTRE I method for selecting the most suitable facility location for an animal premix producer company which has been already located Ankara/Turkey.

Key Words: Fuzzy Set, Fuzzy ELECTRE, Plant Location Selection

Fuzzy Least Squares Regression Analysis and An Application

Derviş Topuz

Department of Health
Services Science,
Niğde Zübeyde Hanım
Vocational School of
Health Services,
Ömer Halisdemir University,
51200, Niğde, Turkey
topuz@ohu.edu.tr

İsmail Keskin

Department of Animal Science,
Faculty of Agriculture, Selçuk
University, 42075, Konya,
Turkey
ikeskin@selcuk.edu.tr

Suat Şahinler

Department of Medical
School, Faculty of
Department of Basic Medical
Sciences, Uşak University,
Uşak, Turkey
suat.sahinler@usak.edu.tr

ABSTRACT

Fuzzy least squares regression analysis method calculates the coefficient values of fuzzy models as fuzzy numbers in the symmetric triangular feature type represented by membership scores. The method calculates many uncertainties and errors that arise during the calculation of the estimator of each coefficient based on the minimum blur criterion. The approach assumes that deviations between the observed values and the predicted values are not due to measurement and observation errors, unlike the classical regression analysis method, that is, the variance of the coefficients of the fuzzy regression analysis model.

In this study, it is aimed to systematically show fuzzy least squares regression model fuzzy coefficient values and their deviations and the interpretation of the objective function (Z) value representing the blurring of these coefficients on a sample dataset.

Key Words: *Fuzzy Logic, Fuzzy Least Squares, Membership Level*

Fuzzy One-Way ANOVA and An Application

Derviş Topuz

Department of Health Services
Science, Niğde Zübeyde Hanım
Vocational School of Health
Services, Ömer Halisdemir
University, 51200, Niğde, Turkey
topuz@ohu.edu.tr

Suat Şahinler

Department of Medical School,
Faculty of Department of Basic
Medical Sciences, Uşak
University, Uşak, Turkey
suat.sahinler@usak.edu.tr

ABSTRACT

One way analysis of variance or completely randomized experimental design should be applied when there is homogeneous experiment material in terms of all other factors except effects to be compared. Because, in this experimental design, the variation is due to all the factors that disturb the homogeneity of the test material, except for the applications to be compared. So even small changes in the material can cause the experimental error to grow, perhaps causing significant differences between the groups, making it seem as if there is no difference. In other words, since the probability of the second type of error increases α , the higher the desired test power $(1-\alpha)$ will decrease.

In this study, it is aimed to explain the theoretical bases of the fuzzy completely randomized experimental design and to display the necessary fuzzy statistical values on a sample dataset systematically.

Key Words: *Fuzzy Variance, Fuzzy II Type Error, One- way ANOVA*

Fuzzy Regression Function Approach Based on Moving Block Bootstrap Method

Eren Bas

Giresun University, Faculty of Arts and Science, Department of Statistics, Forecast Research Laboratory, Giresun
eren.bas@giresun.edu.tr

Erol Egrioglu

Giresun University, Faculty of Arts and Science, Department of Statistics, Forecast Research Laboratory, Giresun
erol.egrioglu@giresun.edu.tr

Ufuk Yolcu

Giresun University, Faculty of Economics and Administrative Sciences, Department of Econometrics, Forecast Research Laboratory, Giresun
varyansx@hotmail.com

ABSTRACT

The classical methods used in the time series analysis are not easy to use because they have some strict assumptions defined on the number of observations, distributions and errors. In recent years, unlike these classical methods, fuzzy time series methods and so on based on the fuzzy set theory, which do not need assumptions that have classical methods and improve the forecasting performance, have been proposed. In the literature, fuzzy approaches for time series forecasting problem have been used frequently in recent years. Fuzzy inference systems, one of the fuzzy approaches, are able to solve complex real life problems as easily as the inference system of the human brain. In the literature, adaptive network fuzzy inference system (ANFIS) is the most commonly used fuzzy inference system. The fuzzy regression function approach, unlike the ANFIS method, is a fuzzy inference system that is as efficient as ANFIS and does not require the rule base in ANFIS. Although the fuzzy regression function approach produces successful forecasting results in the time series problem, it is not possible to obtain the confidence intervals for the forecasts and the distribution of forecasts. Performing probabilistic calculations from a nonlinear and data-based system can be accomplished by resampling methods. The "Bootstrap" method, one of the methods of re-sampling, is a commonly used method in the literature and provides the possibility of obtaining probabilistic results by drawing sub-samples from the current sample. In this study, unlike other fuzzy inference systems, the distribution of forecasts and confidence intervals for forecasts in fuzzy regression function is obtained by using by using moving block bootstrap method. The performance of the proposed method is investigated by using real world time series data sets.

Key Words: *Fuzzy Inference Systems, Fuzzy Function, Moving Block Bootstrap, Forecasting*

Acknowledgements: *This study is supported by "Giresun University Coordinatorship of Scientific Research Projects" with the number FEN-BAP-A-230218-35".*

Fuzzy Survival Analysis for Weibull and Gamma Distribution

Gözde Ulu Metin
Republic of Turkey,
Ministry of Industry and
Technology, Ankara
gozdeulu@hotmail.com

ABSTRACT

Survival analysis is generally defined as set of statistical methods used in analyzing time until the interested event occurs. The event in here can be death, occurrence of an illness, marriage, divorce etc. Analyses such as examining distributions of survival times obtained from n number of trials at T time interval, establishing the models relating to factors which effect to survival times, estimating the parameters of these models, predicting the recuperation times, constructing survival curves etc. constitutes the subject of survival analysis.

Survival data generally consists of censored data points that information about survival time is incomplete. For example, when considering that event examined is death, if death has not occurred during the study (during the experimental subject monitoring) or if experimental object has left the study because of any reason, data relating to these experimental subjects will be censored. So far, various approaches have been developed for models of survival analysis. However, in order to model survival data, non-parametric approaches are generally used since survival data are censored and thus information about the distribution of survival data is not complete.

The most known nonparametric approach is Kaplan Meier (KM) method. This method is based on estimating survival probabilities as the function of time, drawing graph of survival probabilities. However, the estimates obtained from KM estimator may be bias and the variance of estimates may be considerably smaller than actual variances when sample size is small or sample is highly censored. In order to overcome the disadvantage of KM estimator, KM method based on fuzzy sets (FKM) has been proposed.

The objective in here is to produce an estimate for survival curve when no reliable statistical methods are available. In this paper, it is aimed evaluating the efficiency of FKM methods proposed in literature via simulation studies of Weibull and Gamma distribution.

Key Words: *Fuzzy Logic, Survival Analysis, Fuzzy Kaplan Meier Method*

Generalized Linear Model with Hierarchical Likelihood for Sparse Estimation

Özlem Kaymaz

Department of Statistics, Faculty of
Science, University of Ankara,
Turkey
ozlem.gullu@gmail.com

Henry Wood

Leeds Institute of
Cancer and Pathology,
University of Leeds,
United Kingdom
h.m.wood@leeds.ac.uk

Arief Gusnanto

Department of Statistics,
School of Mathematics,
University of Leeds, United
Kingdom
a.gusnanto@leeds.ac.uk

ABSTRACT

Modelling on high-dimensional data is challenging because the number of variables generally far exceeds that of observation. Furthermore, the data usually exhibit some correlations between variables. With these characteristics, a sparse solution in the model fitting is usually preferred due to its simplicity in the interpretation. The term 'sparse' refers to some model parameters that are zero estimated and the other parameters are estimated away from zero. In effect, a variable selection is embedded in the modelling framework. A sparse solutions can be obtained using LASSO. However, this approach has some drawbacks in the context of correlated variables. In this talk, we propose to consider a sparse solution in a generalised linear mixed-effects model framework where the random effects are assumed to follow a gamma mixture distribution (Lee and Oh, 2009). This random effects distribution can be set to have a high concentration around zero with long and fatter tails than normal distribution to produce a sparse solution and makes ridge and lasso solution to be special cases. The results indicate that the modelling framework is able to produce sparse solution with sensible interpretation. We illustrate this methodology using a lung cancer genomic data, where a prediction on the disease subtypes is critical.

Key Words: *Generalized Linear Model, Hierarchical-likelihood, Sparse Estimation, Random Effects*

GIS-Based Approach for Defining Bioenergy Facilities Location: A Case Study in Bismil/Turkey Based on Marginal Delivery Costs and Resources Competition Between Facilities

Serap Ulusam Seçkiner

Faculty of Engineering, Department of
Industrial Engineering, Gaziantep
University, 27310 Gaziantep, Turkey
seckiner@gantep.edu.tr

Adem Akgül

Faculty of Engineering, Department of
Industrial Engineering, Gaziantep
University, 27310 Gaziantep, Turkey
Ademakgul41@hotmail.com

ABSTRACT

The development of new energy sources, especially renewable, has clearly emerged as a promising policy due to the environmental and economic problems. In this context, biomass utilization has come out as a viable alternative for energy production. The two of the most important factors in biomass-to-energy production are the raw material/endproduct (electricity and heat) transportation cost and appropriate capacity in terms of installed power. It is necessary to design the structure of the supply chain scientifically in order to minimize the transportation cost of the raw material, which will be collected in a particularly large area, and to make maximum use of the plant capacities. As the first step in this study, suitable raw material supply points on agricultural land in Bismil district of Diyarbakır, Turkey, suitable points for establishing a tri-generation power plant (demand point), electricity and hot / cold heat demand points and actual road length matrices between them were determined through Geographical Information System. Then, a solution, which maximizes the net present value of a possible investment by establishing an optimal number, capacity and place, was found by means of the 0-1 binary linear programming methods. In the net present value calculation, the installation of the tri-generation power plant, heat transportation and electricity transmission lines are taken into consideration as the investment costs. Annual cash inflows include electricity sales, hot and cold heat sales, and fertilizer sales, while cash outflows include transportation costs, endproduct transportation costs, losses due to distance in energy transmission, raw material purchasing costs and operating costs.

Key Words: *Renewable Energy, 0-1 Linear Programming, Supply Chain Optimization, Biomass*

HIV-1 Infection Prognosis: Evolutionary Game Theory Approach

Bilge BAŞER

Mimar Sinan Fine Arts University
Department of Statistics
bilge.baser@msgsu.edu.tr

Metin YANGIN

Mimar Sinan University
Department of Statistics
metin.yangin@msgsu.edu.tr

ABSTRACT

HIV-1 (Human Immunodeficiency Virus) is a virus that damages the immune system and leads to AIDS (Acquired Immune Deficiency Syndrome) by reducing the resistance of people to infections and diseases. According to World Health Organization (WHO), almost 21 million people were receiving antiretroviral treatment by mid-2017 (World Health Organization, 2018).

A variety of treatment methods called antiretroviral therapy are still being studied for the infection of the HIV-1 virus, which is a type of HIV virus that causes widespread vital dangers worldwide. These methods of treatment are based on drug designs that are developed using inhibitors that promote the dynamics which provide the development of the HIV-1 virus. However, there is a need for studies to improve the drugs against HIV-1 infection because the recommended medicines have serious side effects such as swelling in the throat and tongue, damage to the liver, life-threatening side effects, and interactions with other medicines.

In addition to these effects, HIV-1 is frequently undergoing mutation and mutant viruses develop resistance against to the treatment in use. During virus replication, about 10 billion viruses are originated every day. The tendency of such rapid reproduction and high-order mutation provides diversity and evolutionary success for the HIV-1 virus. For this reason, if the behavior of virus mutation can be modeled and findings can be obtained in the way of preventing mutation, very significant improvements can be provided for treatment of infections.

The aim of this study is to model the replication and developmental process of HIV-1 viruses by evolutionary game theory to obtain results that shed light on drug design against infection. In accordance with (Harada, 2013) the proposed model has four phenotypes of HIV-1. The calculations for describing HIV-1 infection, replication and mutation processes are built by referring to the in vivo role model (Perelson, Neumann, Markowitz, Leonard, & Ho, 1996). Since the nature of the virus and immune cells is dynamic, particle swarm optimization method is used to specify the common resource of viruses in a competitive environment to define the fitness function of the game.

Key Words: *Evolutionary game theory, particle swarm optimization, HIV-1 evolution*

Hierarchical Forecasting

Mehmet Gülşen

Başkent University, Engineering
Faculty, Industrial Engineering
Ankara, Turkey
mgulsen@baskent.edu.tr

Pelin Toktaş

Başkent University, Engineering
Faculty, Industrial Engineering
Ankara, Turkey
ptoktas@baskent.edu.tr

ABSTRACT

For complex data sets, it is difficult to find a single closed-form function that could represent the entire data set. This is particularly true if data includes categorical variables and/or there are missing values. Even for continuous variables, there could be multiple patterns in the data, each requiring a separate model for a better representation. To capture complex data in a closed form representation, a two-tier hierarchical approach is proposed. At the top level, there is a rule-based system that breaks the data into distinct components. Historical data is categorized and broken down into separate sections, each with a specific start and end points. At the lower level, an appropriate forecasting method is employed to each section. Regression models should satisfy a homogeneity criterion indicating that sliced data is sufficiently stable to be represented by a single function. The proposed approach is tested successfully on test data sets.

Key Words: *Hierarchical Forecasting, Regression, Bottom-up Approach*

Hybrid Differential Evolutionary Algorithms for Global Optimization Problems

Wali Khan Mashwani
Department of Mathematics,
Kohat University of Science & Technology,
KPK, Pakistan.
mashwanigr8@gmail.com

ABSTRACT

In the last two decades and so, global optimization has become an emerging area of research due to wide and broad spectrum applications in various fields of Science and Engineering. Evolutionary computation is the subfield of artificial intelligence and soft computing that groups different evolutionary algorithms inspired from the biological principles and theory of natural selection. Among them the nature inspired algorithms are extensively investigated in the last few years aiming at to develop such systems or techniques that might able to learn incrementally, adaptable to dynamic environment and tolerant to noise. In this paper, hybrid differential evolutionary strawberry (HDES) algorithm is developed for solving optimization problems with continuous search spaces. The suggested algorithm is also mainly inspired by the behavior of strawberry plant while utilizing runners and roots to perform search propagation. The simulation results of the proposed algorithm are much promising while converging toward the known optimal solution of each used test problem. The obtained results encourage us to apply the proposed method for dealing with Bin Packing Problems with multiple dimensions and some other practical problems like Cutting Stock problems, Scheduling problems, Loading and Resource Allocation problems, Sheet Metal Forming Processes problems and Cloud Computing Management problems.

Key Words: *Global Optimization, Evolutionary Computation, Strawberry Algorithm, Differential Evolution*

Hybrid Model Approach for Feature Selection in Content Analysis

Adem Doğaner

Kahramanmaraş Sutcu Imam
University Faculty of Medicine
Department of Biostatistics and
Medical Informatics
adem_doganer@hotmail.com

Cemil Çolak

Inonu University Faculty of
Medicine Department of
Biostatistics and Medical
Informatics
cemilcolak@yahoo.com

ABSTRACT

The rapid development of information technologies in recent years has allowed the creation of large datasets. The process of obtaining information from large datasets has ensured the development of different methods and algorithms. Data mining has become an area that is commonly used and developed in data classification and information extraction in recent years. However, many of the algorithms and statistical methods used in data mining provide high performance for structured data. Though there are unstructured data in the web environment, in written sources or in most of the visual sources.

Although text mining is a subdiscipline of data mining, the types of data that it is interested in are different. While data mining makes analysis on structured data, text mining makes on unstructured data. There are certain process steps for extracting information in text mining. One of these steps is the feature selection. With feature selection, unstructured text data reaches smaller data sizes and a more favorable dataset for classification is obtained. The performance of feature extraction methods used in text mining is limited in some cases.

There are different analytical methods that are examined within the context of text mining. Content analysis is a text mining method that systematically summarizes the data obtained from textual or verbal sources.

In the study, it was aimed to use the information obtained from the content analysis in the feature selection of the hybrid machine learning algorithms in order to increase the general representation of the texts. In the study, texts that express the reason of anxiety of 400 university students were used as datasets.

As a result, it has been observed that the use of hybrid methods in feature selection provides higher performance than classical feature selection methods.

Key Words: *Content Analysis, Machine Learning Algorithm, Hybrid Models*

Hybrid Type-1 Fuzzy Time Series Functions Approaches: The Application in Stock Exchanges

Ali Zafer Dalar

Department of Statistics, Giresun
University, Turkey
alizaferdalar@gmail.com

Erol Egrioglu

Department of Statistics, Giresun
University, Turkey
erole1977@yahoo.com

ABSTRACT

In recent years, alternative methods have been used commonly for time series forecasting besides probabilistic models. In the methods based on fuzzy set theory, unlike probabilistic models, there is a fuzzy approach to uncertainty. Fuzzy inference systems based on the fuzzy set theory work with linguistic variables similar to the extraction mechanism of the human brain. The aim of the methods based on the fuzzy set theory for time series forecasting is to produce better interval and point estimations than probabilistic models. Within the scope of this study, considering that the real-world time series are not completely linear, a hybrid method using autoregressive model, and type-1 fuzzy functions approaches is proposed. In addition, the proposed hybrid method has been modified by using the subsampling bootstrap approach, and thus a second forecasting method is proposed. In both approaches, particle swarm optimization algorithm is used to determine the parameters. The proposed second approach shows that it is possible to make statistical inferences in a fuzzy forecasting method. Four different stock exchange time series were used in order to show the forecasting performance of the proposed methods. The proposed methods are compared with some other forecasting methods in the literature in terms of forecasting performance and it is seen that the proposed methods have superior forecasting performance than the other methods.

Key Words: *AR, Forecasting, Fuzzy Functions Approach, Particle Swarm Optimization, Subsampling*

Improved RALS(2)-LM Cointegration Test in Presence of Structural Breaks and GARCH Innovations

Esin Firuzan
Dokuz Eylül Üniversitesi
Fen Fakültesi
İstatistik Bölümü
esin.firuzan@deu.edu.tr

Berhan Çoban
berhancoban@windowslive.com

ABSTRACT

In time series, data may experience a sudden change / break in its dynamics. Structural changes may occur in time series because of policy changes, financial crises and natural disasters. Most of the time series have some negative properties such as excess kurtosis, unexpected volatility periods and skewed or fat tailed distribution. These features lead to volatility of error term and non-normality in series.

In the cointegration framework, many studies demonstrated that the existence of the structural breaks and heteroscedastic error term may cause various problems such as biased, inconsistent estimations and poor predictions.

This study proposes a more efficient test among residual based test, the new test's name is the "RALS(2)-LM GARCH", which is suitable for presence of GARCH effects and level or/and slope structural breaks. The test formed using the residual based Lagrange Multiplier method with residual augmented least square (RALS) estimators. The newly suggested RALS-based cointegration test utilizing higher moment conditions exhibits substantial power gains when the innovations are GARCH distributed.

In this study, we investigate asymptotic and theoretical properties of RALS(2)-LM GARCH cointegration test that allows for aforementioned features in cointegration equation. Also, this study focuses on the existing empirical literature by further re-examining the efficiency gain and empirical size of Engle-Granger (EG), Gregory-Hansen (GH), Westerlund-Edgerton Lagrange Multiplier (WE-LM) and RALS(2)-LM GARCH tests.

When the results of the simulation are examined, it can be argued that when there is a structural break in the series, the RALS(2)-LM GARCH and WE-LM (2007) test yielded better results, the EG(1987) test has conservative results, and the GH(1996) tests have a spurious rejection issue.

Key Words: *Cointegration, Structural Breaks, Heteroscedasticity, GARCH, RALS(2)-LM*

Improvement of Vehicles' Production by Creating the Intelligent Information System to Verify Manufacturability of Design Documentation

Irina Makarova
Kazan Federal University
Syuyumbike prosp., 10a,
423822 Naberezhnye Chelny
kamivm@mail.ru

Ksenia Shubenkova
Kazan Federal University
Syuyumbike prosp., 10a, 423822
Naberezhnye Chelny
ksenia.shubenkova@gmail.com

Vadim Mavrin
Kazan Federal University
Syuyumbike prosp., 10a,
423822 Naberezhnye Chelny
vadim_mmite@rambler.ru

ABSTRACT

The development of the Internet, sustainable communication channels, cloud technologies and digital platforms, as well as information "explosion" of data, have ensured the transition from the local automation of enterprises to the open information systems and global industrial networks, which overstep segregate enterprise and interact with each other. Such systems and networks provide a transition from industrial automation into a new, fourth, stage of industrialization. The concept of "Industry 4.0" was formulated in 2011 by the President of the World Economic Forum in Davos, Klaus Schwab. Breakthrough developments in such areas as artificial intelligence, nanotechnology and others lead not only to the creation of new market segments, but also to a fundamental change in existing business models. Combination of the growth of Internet's dissemination, mobile devices, development of data analysis methods, "Internet of Things" and machine learning are changing the expectations and requests of consumers. Digitalization helps to focus on the customer, so mass production of a new type allows the industrial manufacturing of an individual product. At the same time, the problem of organizing the interaction between the designer and the technologist becomes urgent. One of such tasks is checking design documentation for manufacturability. The article describes one of the variants of solving this problem by developing an intelligent information system to verify manufacturability of design documentation of automotive company. The conceptual scheme of this system, interaction of its modules and its operational algorithms are presented in the paper. The adequacy of the proposed solution is checked by conducting a multifactor simulation. Verification and validation of the system is based on the case study.

Key Words: Industry 4.0, Internet of Things, Automotive Industry, Intelligent Information System

Imputation Strategies for Dealing with Missing Values in Multi-item Scales

Oya Kalaycıoğlu
Bolu Abant İzzet Baysal University
Department of Econometrics
oyakalaycioglu@ibu.edu.tr

ABSTRACT

In multi-item scales, missing data may arise in one or more item scores. Various methods exist to handle missing item scores, however the performance of these methods are not thoroughly investigated in terms of their effects on the validity and reliability of the scales. The aim of this study is to evaluate various methods for handling missing items in multi-item scales via simulations based on the real data collected from a questionnaire consists of 39 Likert-type items and four subscales.

The methods that have been compared include complete case analysis (CC), single imputation methods (SIM), multiple imputation to total scale scores (scale-MI) and multiple imputation to the item scores (item-MI). Item-MI, which uses all of the items in the questionnaire as predictors in the imputation model, may suffer from over fitting problems when the number of incomplete items is large. Therefore two different techniques were used to reduce the number of predictors in the imputation model. First, the predictors in the imputation model were selected with forward selection approach and second, rather than using item scores, scale scores were used as predictors when imputing an item. With simulation studies, mean of the sub-scale scores were compared using the bias and coverage as the performance parameters for different missing data handling techniques under various settings of missing data mechanisms (MCAR, MAR and MNAR). Additionally, commonly used measures to ensure validity and reliability of the multi-item scales were assessed.

Multiple imputation of each scale item separately outperformed in terms of bias and coverage of the mean sub-scale scores. When the validity and reliability measures are concerned item-MI is also preferable in either MAR or MNAR scenarios, however if the number of items in a multi-item instrument is large it is recommended to use an appropriate variable selection technique for the imputation model. All methods were most sensitive when missing data were MNAR. The methods did not show substantial differences if the missingness proportion is less than 10%. In general, single imputation methods resulted in large bias and therefore it is not advised to use any form of single imputation method to handle missing item scores.

Key Words: *Multiple Imputation, Missing Data, Multi-Item Scales, Simulation*

Inventory Control Policy Decisions in the Presence of Servers Subject to Random Failure

Serife Özkar

Balıkesir University, Department
of International Trade and
Logistics
serife.ozkar@balikesir.edu.tr

N. Onur Bakır

Altınbaş University,
Department of Industrial
Engineering
onur.bakir@altinbas.edu.tr

ABSTRACT

Control policies are extensively studied within queueing theory and inventory theory. In practice, it is not always optimal to resume service immediately upon the arrival of a customer particularly if the server has a significant switching or setup cost. In this case, one can argue that the per unit time operational cost well exceeds the customer waiting cost. Thus, it is better to delay service until a predetermined number N of customers are accumulated in the waiting line. The decision on N in the presence of the trade-off between operational cost and customer waiting becomes more complex when the system server is assumed to be subject to failure.

We study an inventory model with service time, and a server subject to failure. Customers arrive to a single server according to a Poisson process. Service times of customers are exponentially distributed. The inventory is governed by an (s, S) policy. Whenever the server becomes idle, it is switched off and the server is activated as soon as the number of customers in the waiting line hits N . If the inventory level drops to s , then an order for replenishment is placed and lead time is assumed to be zero. The server can fail during active service and a repair action is taken immediately upon failure. Two alternative types of repair are available: minor or major, where major repair makes the server as good as new, and minor repair merely brings the server to operational condition. We obtain the steady-state probabilities of the system and then related performance measures. We perform some numerical examples by using a cost function structured to obtain the optimal values of the control variables s , S , and N .

Key Words: *Inventory system with service time, N-policy, Server breakdown*

Inversion of Electrical Resistivity Data: An Application from Gümüşhane University Football Field

Mahmut Sarı
Gümüşhane University,
Department of Geophysics
msari@gumushane.edu.tr

Serkan Öztürk
Gümüşhane University,
Department of Geophysics
serkanozturk@gumushane.edu.tr

ABSTRACT

In this study, it was attempted to determine the problematic area, where subsidence occurred in Gümüşhane University Football Field, by using the electrical resistivity method. The measurement was made by using a multichannel electrical resistivity device along 1 line with the total length of 100m. 41 electrodes in total were used, and spacing between each electrode was selected as 2.5 m. Individuals working in the field of environment and engineering use the measurement results generally obtained by the electrical resistivity method for mapping the conductivity characteristics of the underground. Two different arrays, being the Wenner and Schlumberger arrays, were used in the electrical resistivity method. The Wenner array was preferred because it can differentiate horizontal layers under a common point. The apparent resistivity measurement calculation for the Wenner array is given with the equation of $\rho_a = (2\pi a) V/I$. In the equation, V is the volt, and I is the current. The Schlumberger array was preferred since it provides high vertical resolution. The equation of the apparent resistivity measurement for the Schlumberger array is given with the $\rho_a = [\pi(L^2 - l^2)/2l] V/I$ formula. In the equation, L denotes the distance between the current electrodes, and l denotes the distance between the potential electrodes. Inversion operations were applied to the measurements taken, and 2-dimensional electrical resistivity sections were obtained. Upon examining the sections, it was observed that the resistivity values changed between 23.9-523 Ω on average. However, the resistivity values were found to change between 826-1305 Ω in (0-4 m) approximate areas close to the surface in the problematic area. It was concluded that the problematic area with subsidence resulted from the rocks exhibiting high resistivity values.

Key Words: *Gümüşhane, Inversion, Electrical resistivity method, Appearance resistance*

Investigation of The Attitudes of Academicians About The Individual Pension System: The Case of Eskisehir Osmangazi University

Fatih Cemrek

Associate Professor,
Eskişehir Osmangazi University
Faculty of Science and Letters,
Dept. Of Statistics
Meşelik Campus F1 Blok 26480
Odunpazarı/Eskişehir
fcemrek@ogu.edu.tr

ABSTRACT

Individual Pension System (IPS) in Turkey came into operation with the law dated October 7, 2001. First pension schemes were approved and put into practice on October 27, 2003. With IPS, the government's providing contribution in the ratio of 25 % of the contribution rate of the employee's contribution increased the participation rate to the system after January 1, 2013. Afterwards, a new period obligating all the employers under the age of 45 to participate started on January 1, 2017.

The system allows the individuals to save money while they are still working, and with these savings they can have extra income during their period of retirement. In Individual Pension System, contributions paid by the contributors are canalized to investment via private pension funds, and thus it is aimed to provide a retirement income which will be sufficient during the retirement period.

Thanks to this system, individuals prosper during their retirement, and long term pension funds provide significant finance for the country economy. According to Pension Monitoring Center, by the date of July 20, 2018, the contributor fund amount is ₺ 71626,1 million; the government contribution fund is ₺ 9805,5 milllion; the number of the contributors is 6977575; the contribution amount is ₺ 566548,3 million, and the amount canalized to investment is ₺ 55786,1 million.

In this study, it was aimed to investigate the attitudes of academicians about the Individual Pension System. A survey having three sections was designed for this purpose. In the first part, there are demographic questions, and the second part includes questions about saving and investment. In the third part, the questions are 5 point-Likert type scale questions to reveal the academicians attitudes on the Individual Pension System. The survey was conducted on randomly chosen 250 academicians working at ESOGU in the 2016-2017 educational year. The questions asking about the academicians attitudes on the Individual Pension System were applied factor analysis. As the result of the factor analysis, the questions were grouped in three factors and total variance expression ratio was determined as 64,4 %.

Key Words: *Individual Pension System Individual, Eskisehir Osmangazi University, Academician, Factor Analysis*

Investigating the Impact of Polysomy 17 in Breast Cancer Patients With Amplification through Meta-Analysis

Ezel Ozge Tas
Uludag University, Institute of
Health Science, Department of
Biostatistics, Turkey
E-mail: ezelozgetas@gmail.com

Sebnem Sag
Uludag University, Faculty of
Medicine, Department of
Medical Genetics, Turkey
E-mail: ozemri77@yahoo.com

Ilker Ercan
Uludag University, Faculty
of Medicine, Department of
Biostatistics, Turkey
E-mail: iercan@msn.com

ABSTRACT

Since studies regarding the effect of P17 in breast cancer cases in some clinical findings are few in number and are in small sample sizes, during the meta-analysis the effects of P17 in patients with amplification on lymph node involvement and grade were examined. Prior to meta-analysis, publication bias of the studies was examined. In cases of publication bias, the trim and fill method was applied. Heterogeneity of the studies was evaluated. In cases where heterogeneity was determined in the publications the DerSimonian-Laird method was carried out using the random effects model, while when there was homogeneity in the publications, the Mantel Haenszel method was applied using the fixed effects model. As a result of this study, when meta-analysis results are evaluated generally, it was observed that in lymph node involvement, polysomy 17 was a risk factor in patients with amplification. With regard to grade, polysomy 17 was a risk factor when grade increased in patients with amplification.

Key Words: *Meta Analysis, Polysomy 17, Breast Cancer, HER2, Amplification*

Investigating the Use of Mobile Banking Application by the Technology Acceptance Model (TAM)

İlkay Altındağ

Department of Banking, Faculty
of Applied Sciences, Necmettin
Erbakan University, 42060, Konya,
Turkey
ialtindag@konya.edu.tr

Ebru Özer Topaloğlu

Department Of Management
Information Systems, Faculty of
Applied Sciences, Necmettin
Erbakan University, 42060,
Konya, Turkey
eozer@konya.edu.tr

ABSTRACT

The technology acceptance model (TAM) is an information systems theory that models how users are using and adopting a technology. This theory has two main purposes. The first is to estimate the acceptability of a tool and the second is to define the modifications that must be brought to the system in order to make it acceptable to the user. The aim of this paper is to examine the relationships between the factors determining the behavior of bank customers who use the mobile banking application by using the Technology Acceptance Model (TAM). Mobile banking is a service provided by a bank or other financial institution that enables its customers to perform their financial transactions without needing to go to the bank office using a mobile device such as a smartphone or tablet. Unlike classic Internet banking, software is used for mobile banking devices developed by financial institutions. It is called usually Mobil Banking Application. With this application, almost all banking transactions can be done very quickly. Due to the advantages it provides, there has been a significant increase in the number of customers using mobile banking in recent years. This research was tested using data collected from 687 questionnaires which was applied to bank customers in Konya. The collected data were analyzed by structural equation modeling.

Key Words: *Mobile Banking Service, Structural Equation Model, Technology Acceptance Model*

Investigation of European Football Clubs Efficiencies with Bi-Objective Multi Criteria Data Envelopment Analysis (BiO-MCDEA)

Talip Arsu
Aksaray University
Faculty of Economics and Administrative
Sciences
Department of Business Administration
arsutalip@gmail.com

ABSTRACT

Football, the most popular sporting event in the world, has turned into an industry line besides as a sporting event especially in Europe. The revenues and brand values of football clubs have become competing with many industries and brands. To be successful, they need to increase their revenue and in order to increase their revenue, soccer clubs, which clubs should be successful, must be both sporty and financially effective organizations. Therefore, the aim of the study is to examine the last three years' efficiency values of 10 clubs in Europe which are at the top of both sportive and financial sense. For this purpose, an efficiency analysis was performed with the Bi-Objective Multi Criteria Data Envelopment Analysis (BiO-MCDEA) model, which is a multi-criteria data envelopment analysis (MCDEA) based goal programming, using inputs and outputs considering sports success and financial success. The BiO-MCDEA method is a two-objective weighting method developed by Ghasemi et al (2014). The reason for choosing MCDEA methods instead of classical data envelope analysis methods is to avoid the low discriminatory disadvantage of classical data envelope analysis. The reason for choosing the BiO-MCDEA model is to avoid subjective errors due to the necessity of selecting from 3 different efficiency values resulting from the MCDEA method.

The average number of fans participating in matches played by the clubs on their own stadium, the number of fans following the social media account of the clubs and the total market value of the players that the clubs own were used as the input and UEFA club score and club total revenues at that time were used as output.

10 football clubs used in the study are: Manchester United, Real Madrid, FC Barcelona, Bayern Munich, Manchester City, Arsenal, Paris Saint-Germain, Chelsea, Liverpool, Juventus. Effective clubs are as follows: Manchester United, Paris Saint-Germain and Chelsea in 2017, Paris Saint-Germain and Liverpool in 2016, Arsenal and Paris Saint-Germain in 2015. Only Paris Saint-Germain club is found effective in the last 3 years. The common feature of effective clubs in 2015 and 2016 is that they earn higher revenue than their market value that year. Manchester United, Manchester City and Liverpool in 2015 and Manchester United in 2016, although they earned more than their market value, were not considered as effective clubs for UEFA club scores are low. In 2017, UEFA club score determined the effective clubs, due to there wasn't a club which has higher revenue than market value.

Considering the findings from the study, it is not enough that only a UEFA club score, revenue or market value is high for a club can be effective. An evaluation can be made for all European clubs by increasing the number of clubs and the number of input-output assessed in future studies. This assessment will yield beneficial results for the clubs, rating agencies and all stakeholders.

Key Words: *Football, Multi Criteria Data Envelopment Analysis (MCDEA), Bi-Objective Multi Criteria Data Envelopment Analysis (BiO-MCDEA)*

Investigation of The Effect of The Phase Space Parameters on Prediction Performance in Chaotic Time Series

Ayşe İşi

Gazi University, Ankara, Turkey

isiayse@gmail.com

ABSTRACT

The prediction of chaotic time series is done by using

$$X_t = [x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(d-1)\tau}]$$

time-delayed vectors, where x_t is an observed time series, τ is the time delay and d is the embedding dimension. Prediction values are obtained by assuming that there is a functional relationship between current state of the system X_t and future value x_{t+T} as $x_{t+T} = f(X_t)$.

Several methods have been proposed in the literature to determine the time delay (τ) and the embedding dimension (d), which are phase space parameters (Fraser and Swinney, 1986; Kennel et al, 1992). In this study, phase space parameters are determined differently from these methods and their effect on estimation performance is investigated.

According to Zhang and Man (1998), when $\tau > 16,8$ the system is chaotic. According to Takens' (1980) embedding theorem, the embedding dimension (d) is defined as $d_a \leq d \leq 2d_a + 1$, where d_a is the fractal dimension of the attractor. For this reason, τ was chosen between 17 and 34; d was chosen between 3 and 6. Prediction performance of the Nearest Neighbor Method predictions using different τ and d values were compared using RMSE criterion.

As a result of the experimental study using the FTSE 100 index values, it has been observed that different τ values do not change d values. Better estimation performance is obtained when phase space parameters are determined by the methods proposed in this study.

Keywords: Chaotic Time Series, Chaotic Prediction, Time Delay, Embedding Dimension, Phase Space Parameters

Likelihood Function for Zero-inflated Models with Fuzzy Counts

Godrick Oketch
Department of Statistics,
Yildiz Technical University
godrickmaradona@gmail.com

Filiz Karaman
Department of Statistics,
Yildiz Technical University
filiz_karaman@hotmail.com

ABSTRACT

Statistics literature suggests zero inflated models in situations when excess of zeros exist in counts. These models have been widely applied to model counts that are precisely reported. However in situations when some counts other than zeros in a dataset are reported with uncertainty, the idea of introducing fuzzy counts may be helpful to model such counts. This paper identifies counts in a dataset that are reported as fuzzy then proposes a likelihood function that not only includes zero inflation of the counts but also alpha cuts of fuzzy numbers to help estimate coefficients of a model. We eventually illustrate the model using data from the National Longitudinal Study of Adolescent Health to model some independent variables on number of sex partners among teens. The results show that the model gives better predictions when imprecisely reported data points are considered fuzzy. The model becomes even better when the impreciseness of the fuzzy numbers is increased by changing the alpha cuts.

Key Words: *Likelihood Function, Fuzzy Regression, Zero-Inflated Count Models*

Liu Pretest and Stein-Type Shrinkage Estimation in Linear Regression Models

Bahadır Yüzbaşı
Inonu University,
Malatya, Turkey
byzb@hotmail.com

S. Ejaz Ahmed
Brock University,
St. Catharines, Canada
sahmed5@brocku.ca

Yasin Asar
Necmettin Erbakan
University, Konya, Turkey
yasar@konya.edu.tr
yasinasar@hotmail.com

ABSTRACT

In a multiple linear regression model, it is usually assumed that the explanatory variables are independent of each other. However, the multicollinearity problem arises when the explanatory variables are dependent. In this case, some biased estimations, such as shrinkage estimation, principal components estimation (PCE), ridge estimation, partial least squares (PLS) estimation Liu estimator and Liu-type estimator were proposed to improve the least square estimation (LSE). In this study, we present the preliminary test, Stein-type and positive part Liu estimators in the linear models when the parameter vector β is partitioned into two parts, namely, the main effects β_1 and the nuisance effects β_2 such that $\beta = (\beta_1, \beta_2)$. We consider the case that a priori known or suspected set of the explanatory variables do not contribute to predict the response so that a sub-model may be enough for this purpose. Thus, the main interest is to estimate β_1 when β_2 is close to zero. Therefore, we investigate the performance of the suggested estimators via asymptotically and both conducted a Monte Carlo simulation study and applied a real data example to evaluate the relative efficiency of the suggested estimators, where we demonstrate the superiority of the proposed estimators.

Key Words: *Liu Estimation, Pretest and Shrinkage Estimation, Multicollinearity*

Market Segmentation With Clustering Analysis –SSO Sample

Songül Bulut
State Supply Office
Ankara, Turkey
songul.bulut@dmo.gov.tr

ABSTRACT

In this essay, the importance and benefits of market segmentation to be done by clustering analysis from data mining methods. With clustering analysis, customers can categorize their segments and find out the products they prefer, and marketing strategies can be defined in this direction. Marketing is a process which the customers' needs and demands are well understood and a number of strategies are generated and presented to institution.

In this study, State Supply Office's customers are classified using order number, sales, product range, and sales method parameters. In practice, hierarchical and non-hierarchical clustering analysis methods are used. To create market segmentation by dividing our customers into subgroups by clustering analysis and so which product offer is available for which customer group, find out how to implement a marketing strategy, protect our existing customers, reach out to potential customers and to provide customer satisfaction and to provide effective reliable sales service to the public.

At the end of this study, the marketing strategy will be determined according to the segments of the customers and it is aimed to determine what kind of campaign should be done in which segment and customer needs and requirements will be better determined.

Key Words: *Cluster Analysis, Market Segmentation, Data Mining, State Supply Office (SSO)*

Markovian Queueing System with Ordered Entry

H. Okan Isguder
Dokuz Eylul University, Faculty of
Science, Department of Statistics,
Izmir, Turkey
okan.isguder@deu.edu.tr

Neslihan Avcu
Dokuz Eylul University,
Engineering Faculty, Electrical
and Electronics Engineering
Department, Izmir, Turkey
neslihan.avcu@deu.edu.tr

ABSTRACT

In this study, Markovian queueing system with heterogeneous servers and ordered entry is analyzed. Inter-arrival times are assumed to be independent and have exponential distributions. There are c servers and the k^{th} server has an exponential distribution with parameter $m_k (k=1,2,\dots,c)$. Any arrival is served by the server whose index number is the lowest one among the idle servers with probability 1. This service discipline is known as the ordered entry and it was firstly introduced by Palm to analyze a system with homogenous servers [1]. Additionally, the multi-server queueing system with heterogeneous servers was primarily considered by Gumbel in 1960 [2]. In his study, the waiting line was assumed to be infinite and also, arrival customer was served in one of the idle servers with equal probability. The difference of the study from these previous works [1-2] is not only the chosen service discipline but also the system capacity. In this study, the considered system has an ordered entry service discipline with finite capacity. Since the system capacity is K , the arrivals wait in the queue when the all servers are busy. Therefore, an arrival leaves without having any service when system capacity is achieved.

The mathematical model of defined system is obtained and analyzed using the matrix geometric solution. The steady-state probabilities and the distribution function of the waiting times in the queue are obtained. The performance measures such as mean waiting time, mean number in the queue etc. are calculated.

The dimensions of the matrices in the model increase with the huge number of servers and the capacity of the system as in real examples. Due to the computational complexities of the calculations, the exact solutions can be computed numerically for a limited capacity. The results with the closed-form solution and the matrix geometric solution calculated in MATLAB for different capacities are compared.

Key Words: *Heterogeneous, Ordered Entry, Matrix Geometric Solution, Closed-Form Solution*

M-Competitions: ATA Method and Others

Güçkan Yapar
Dokuz Eylül University,
Department of Statistics
guckan.yapar@deu.edu.tr

ABSTRACT

Forecasting plays a critical role for almost all of science and the “M” competitions organized by Spyros Makridakis have had an enormous influence on the field of forecasting. They focused attention on what models produced good forecasts, rather than on the mathematical properties of those models. In this presentation, a new forecasting approach, ATA method, will be introduced and the forecasting accuracy of this new forecasting method that is alternative to two major forecasting approaches: exponential smoothing (ES) and ARIMA, will be evaluated using M3 and M4 competition data set. Finally, I will present ideas on how the forecasting performance of the method can be improved so that desirable results can be obtained for the most recent of the series M4- competition.

Key Words: *Exponential Smoothing, Forecasting, Initial Value, Forecasting Competition, Smoothing Parameter*

Measurement of the Region's Functions on the Example of the Analysis of the Level of Innovation

Marek Szajt

Czestochowa University of Technology
marek.szajt@wz.pcz.pl

Agata Mesjasz-Lech

Czestochowa University of Technology
agata.mesjasz-lech@wz.pcz.pl

ABSTRACT

In the literature we meet with assigning particular regions to a given function. Regions are defined as industrial, agricultural, tourism as well as innovative. In the initial analyzes, the role of the region was determined by employment or the impact of production of a given sector on a given region or its importance in a wider administrative sense. Currently, other - more complex measures are also used. The region's specialization was treated as a determinant of its development with the use of competitive advantage in a given area. The aim of the article is to verify the possibility of adapting the measures used in measuring the tourist function to measure the innovative function.

Most of the research on innovations uses tools in the field of descriptive statistics, mathematical statistics, numerical taxonomy or econometrics. This article focuses on the use of spatial methods. As a tool, the quotient and location index, the index of specialization, and the exogenous sector diversification index were used. The results, clearly less pronounced than their counterparts for other industries, allow to indicate the proposed measures as useful in the analysis of the importance of innovation for the development of the region.

Key Words: *Spatial Statistics, Innovations, Regional Analysis*

Measuring Product Sustainability: A Risk-Focused Approach

Mürüvvet Deniz Sezer
Dokuz Eylül University
Industrial Engineering Department
mdenizsezer@gmail.com

Hasan Selim
Dokuz Eylül University
Industrial Engineering Department
hasan.selim@deu.edu.tr

ABSTRACT

Sustainability is the holistic approach that aims making future-oriented programming with economic, social and environmental dimensions. Sustainable product development is one of the most significant issues facing in today's companies due to rapid depletion of raw material resources, increase in the social and environmental impact of products, legislative regulations and awareness with regard to social and environmental issues. Within this framework, measuring product sustainability plays essential role in gaining and protecting competitive advantage. However, it is necessary to consider all risks associated with sustainability in order to measure the product sustainability effectively. In this regard, the objective of this study is to develop a risk-driven approach for measuring product sustainability. Previous studies in the related field are restricted due to the lack of dynamic and multidimensional measurement. However, in the real-world settings, because of decision-making environment constantly changing, it requires a dynamic decision-making approach. This study proposes an approach for measuring product sustainability that consider all risks and risk interactions. To deal with the dynamic and multidimensional features of product sustainability, the proposed approach employs system dynamics modelling. To confirm the applicability of the proposed approach a real-world application is presented.

Key Words: *Product Sustainability, System Dynamics, Risk Management*

Measuring the Financial Performances of Companies Listed on BIST-Services Index with Data Envelopment Analysis

Murat Gül

Department of Statistics, Giresun
murat.gul@giresun.edu.tr

Olcay Karakuş

Department of Statistics, Giresun
olcayykarakus@gmail.com

ABSTRACT

The concepts of efficiency and effectiveness in daily life which are two important indicators of performance are used interchangeably, although in reality the meanings they carry are different from each other. The concept of efficiency refers to the proportional relationship between output and inputs. The concept of activity, its strategic goals and defined objectives of the organization as a result of the activities they perform in order to achieve these objectives and goals that determines the degree of achievement of a performance size. Several analysis methods have been developed for the measurement of the activity. Data envelopment analysis was used in the method of measuring the research activity.

Within the content of this study; financial performance evaluation of services companies listed on BIST-Services Index is carried out based on Data Envelopment Analysis (DEA). For that purpose, in order to expose the financial performances of the mentioned companies, inputs-outputs are defined by using financial statements and annual report for resolution units. According to the findings resulted by the analysis, it is revealed that the inadequacy on which accounting branches of the companies that are not found as effective. Thus, the efficiency of the mentioned companies has been examined and measured

Key Words: *Services Sector, Performance Evaluation, Data Envelopment Analysis*

Mix EWMA-CUSUM Control Chart In the Presence Of Measurement Error

Muhammad Noor-ul-amin
COMSATS Institute of Information
Technology, Lahore
nooramin.stats@gmail.com

Afshan Riaz
COMSATS Institute of
Information Technology, Lahore
afshanriazpst001@yahoo.com

ABSTRACT

In this study, we considered the mix EWMA-CUSUM control chart to examine the effect of measurement error. The model used is the one involving linear covariates. We estimate the Average run lengths (ARLs) by using Monte-Carlo simulation method. It is observed on the basis of ARLs performance of the mix EWMA-CUSUM control chart is seriously effected in the presence of measurement error on the basis of ARL.

Key Words: *EWMA, CUSUM, Monte-Carlo*

Modeling the Financial Attitudes and Behaviors of Y Generation via Structural Equation Modeling

Murat Akşit
Big Cat Research, İstanbul, Turkey
murat@bigcatresearch.com

Sinan Saraçlı
Department of Statistics, Faculty of Science
Literature, Afyonkarahisar, Afyon Kocatepe
University, Turkey
ssaracli@aku.edu.tr

ABSTRACT

Members of this generation composed of individuals born between 1980-2000, is a generation that has to work for living. In their lives, technology is the symbol of many things. They are narcissi, individual and entrepreneur. This generation, which does not like to work, loves entertainment and winning, acts aggressive to authority, consumes fast and has dissatisfied manner. Their most important characteristics are freedom and technology. Basic specifications of Y generation are their ability to use technology and loneliness. One other characteristics of this generation is self-confidence (Çelik and Arslan Gürcüoğlu, 2016).

SEM is a statistical technique for testing and estimating causal relationships using a combination of statistical data and qualitative causal assumptions. It is used in social, behavioral and educational sciences, namely, psychology, biology, economy, marketing and medicine. SEM is a comprehensive statistical method used in testing hypotheses about causal relationships among observed and unobserved (latent) variables and has proved useful in solving problems in formulating theoretical constructions (Schumacker & Lomax, 2004; Reisinger & Turner, 1999; Raykov & Marcoulides, 2006; Yılmaz, 2004).

In this study to model the financial attitudes and behaviors of Y generation via structural equation modeling, a Likert scale type questionnaire, ranging from 1 'strongly disagree' to 5 'strongly agree' is applied to 372 individuals via Computer Assisted Personal Interviewing (CAPI) on Pollyfy application. To analyze the data, SPSS and LISREL software are used and the results are given in related tables and figures.

Key Words: *Financial Health, Financial Behavior, Structural Equation Modeling.*

Modelling of Wind Speed Data for Konya

| | | |
|---|--|---|
| İsmail Kınacı Selcuk University, Konya, Turkey ikinaci@selcuk.edu.tr | Emine Karakaya Selcuk University, Konya, Turkey kkarakaya@selcuk.edu.tr | Coşkun Kuş Selcuk University, Konya, Turkey coskun@selcuk.edu.tr |
|---|--|---|

ABSTRACT

Recently, importance of wind energy is increasing in the world. In this study, the hourly wind speed data has been modelled by some statistical models such as Exponential-Poisson, APT-Weibull, ALT-Exponential and Mixed Weibull models. Finally, we estimate the potential wind energy production of Doğanhisar, Altınekin and Güneysınır which are three districts of Konya for per wind turbine by using considered models.

Key Words: *Inference, Renewable Energy, Statistical Modelling, Wind Speed*

Modeling Nonlinear Relation by Using Running Interval Smoother, LOWESS and Different Quantile Estimators

Burak Dilber
Dokuz Eylül University, İzmir
burakdilber91@gmail.com

A. Fırat Özdemir
Dokuz Eylül University, İzmir
firat.ozdemir@deu.edu.tr

Gözde Navruz
Dokuz Eylül University, İzmir
gnavruz@gmail.com

ABSTRACT

When dealing with curvature, the problem of estimating and plotting a regression line is an important issue in applied statistics. The weighted averages of the Y values with the weights of a function of how close the vector of predictor values is to the point of interest are known as smoothers or nonparametric regression estimators. In this study, the problem of estimating and plotting a regression line when the goal is to determine the conditional quantile of some random variable Y given X is considered. For this purpose, the running interval smoother and Cleveland's LOWESS are applied consecutively in literature. The running interval smoother is used since it enables using any robust estimator and it is followed by Cleveland's LOWESS for further smoothing. This approach which consists of two smoothing methods, is applied with three quantile estimators: Harrell Davis quantile estimator, a newly proposed NO quantile estimator and the default quantile estimator of R programming language. The performance of these estimators, is investigated under both theoretical distributions and real data sets in terms of mean squared error and mean maximum absolute error performance criteria.

Key Words: *Running Interval Smoother, LOWESS, Harrell-Davis Estimator, NO Estimator*

Modifying Clements' Method with Alternative Skewness and Kurtosis Estimators

Burcu Aytacıoğlu

Ege University, Faculty of Science,
Department of Statistics, İzmir, Turkey.
burcu.aytacoglu@ege.edu.tr

Cemil Giray Genç

Ege University, Faculty of Science,
Department of Statistics, İzmir, Turkey.
cgiraygenc@gmail.com

ABSTRACT

In many industries process capability studies are conducted in order to determine the capability of the process to produce acceptable products and it is one of the important activities of statistical process control. It has an essential role in continuous improvement of quality and productivity. In order to express the capability of a process, process capability indexes are frequently used. The most widely used traditional capability indexes are C_p , C_{pl} , C_{pu} and C_{pk} . These standard indexes are usually computed under the assumption that the process data follow normal distribution. In many situations, however, normality assumption may be violated and the use of these traditional capability indexes may cause misleading interpretation about the capability of the process. For example, semiconductor processes, some chemical processes and coating processes often follow skewed distributions. In literature, several methods are proposed for determining process capability indexes under non-normality. One of the most widely discussed methods is Clements' method which is proposed in 1989.

Clements' method uses the Pearson family of curves for calculating capability indexes for any shape of distribution. It requires the estimation of the mean, standard deviation, skewness and kurtosis. Moreover, this method makes use of the classical estimators of skewness and kurtosis. In this study, we discussed the use of more robust estimators of skewness and kurtosis in the calculation of process capability index by Clements' method. For this purpose, capability indexes with the use of these robust estimators are simulated and the mean square errors (MSEs) of them are reported in conjunction with the MSEs of the capability index that uses classical estimators of skewness and kurtosis. Comparison is done through simulating Weibull data with several different shape parameter values. The results revealed that the use of robust estimators of skewness and kurtosis in the calculation of process capability index by Clements' method outperforms the classical approach unless the skewness is not considerably large.

Key Words: *Process Capability Index, Clements' Method, Skewed Population, Robust Estimator*

Monitoring of Blood Parameters of Dogs and Cats Using Control Chart Techniques

Liaquat Ahmad

Department of Statistics and
Computer Science, University of
Veterinary and Animal Sciences,
Lahore 54000, Pakistan
liaquatahmad@uvas.edu.pk

Sabqa Khan

Department of Statistics and
Computer Science, University of
Veterinary and Animal Sciences,
Lahore 54000, Pakistan
sabqakhan7@gmail.com

Muhammad Aslam

Department of Statistics,
Faculty of Science, King
Abdulaziz University,
Jeddah 21551, Saudi Arabia
aslam_ravian@hotmail.com

ABSTRACT

Monitoring of medical outcomes demands for the early and quick detection of the unusual change in the interested parameter. In this research paper the application of Statistical Process Control (SPC) using the control charts have been presented for the veterinary based data. The traditional Shewhart control chart is not effective for monitoring the small or moderate changes in the targeted quality characteristic. In this paper the quick detection technique of Exponentially Weighted Moving Average (EWMA) has been used for the efficient monitoring of the data collected from January to April, 2016 consisting of eight parameters of blood test in dogs and cats. According to the results the EWMA chart shows the out-of-control observation at the 5th sample while the X-bar chart shows the out-of-control observation at the 24th sample of white blood cell parameter of dogs. It has been observed and shown that the application of EWMA technique on the blood parameters is an effective method for monitoring. The diagrams of X-bar chart and EWMA based X-bar chart have been constructed for all parameters.

Key Words: *Statistical Process Control, X-bar Chart, EWMA Chart, Blood Parameters*

More Efficient Estimator for Unit Root and Cointegration: RALS Approach

Berhan Çoban
berhancoban@windowlive.com

Esin Firuzan
Dokuz Eylül Üniversitesi,
Fen Fakültesi, İstatistik Bölümü
esin.firuzan@deu.edu.tr

ABSTRACT

Ordinary Least Squares (OLS) estimation technique is a major tool for estimating the parameter of the linear regression model. The accuracy and validity of regression model depend some basic error assumptions: independency, linearity, normality and homoscedasticity.

Several noticeable features of econometric/economic time series are well known with a fat-tailed distribution. This destructive behavior leads to aperiodic volatility of error term and non-normality in series. In these circumstances OLS estimator is still unbiasedness and consistency, but inefficient.

Generally, cointegration, unit root and causality analyses are performed by least square estimation can also be affected by the violation of homoscedasticity and normality of error term.

In this study, we propose new estimator for unit root and cointegration analysis, which is more efficient than classical OLS estimation. Residual Augmented Least Squares (RALS) approach, which utilizes information on the skewness and excess kurtosis, is more efficient than classical OLS estimation. The RALS estimator, which is closely related to the GMM estimator, is one of a widespread variety of alternative robust estimation methods, which have more efficiency for non-normal data.

RALS method first introduced by Im (1996), then extended by Im & Schmidt (2008) to the functional form of estimator under non-normality of error term. Im, Lee & Tieslau (2014) proposed unit root tests based on RALS (2&3) and RALS (t5). Meng et al. (2014) improved the RALS methodology connected with LM test, hereby they gained more powerful and robust test results with non-normal errors and some arrangements of nonlinearity. From cointegration viewpoint, Taylor & Peel (1998) has suggested RALS cointegration test, which is robust to excess skewness and kurtosis of the error term. Lee et al. (2015) developed present cointegration tests Engle-Granger, Modified Engle Granger (Modified-EG), Autoregressive Distributed Lag test (ADL) and Error Correction Model (ECM) owing to RALS approach. Modified tests indicated that resulting estimators can be more efficient and the tests may become more powerful under the non-normal error term.

In this study, we investigate asymptotic properties of cointegration test based on RALS methodology. Lastly, to obtain more powerful or/and efficient tests are shown through by Residual Augmented Least Squares (RALS).

Key Words: *Residual Augmented Least Squares (RALS), Non-normality, Unit Root, Cointegration*

Multicomponent Stress-strength Reliability Estimation for A Model Based on Bivariate Generalized Exponential Distribution

Mustafa Nadar

Istanbul Technical University,
Turkey
nadar@itu.edu.tr

ABSTRACT

A multicomponent stress-strength model, usually called an s -out-of- k : G system, is a system which has k s – independent and identically distributed strengths components and a common stress. The system functions as long as s ($1 \leq s \leq k$) or more components can withstand the stress. This model, which has both industrial and military applications, was first introduced by Bhattacharyya and Johnson (1974). In this paper, we consider a system which have k identical strength components and each component is constructed by two pair of dependent elements. These elements $(X_{11}, X_{12}); (Y_{11}, Y_{12}), \dots, (X_{k1}, X_{k2}); (Y_{k1}, Y_{k2})$ follows from bivariate generalized exponential with $BVGE(\alpha_1, \alpha_2, \alpha_3, \lambda)$ (see Kundu and Gupta(2006)) and each element is exposed to a common random stress T which follows generalized exponential distribution with $GE(\lambda, \theta)$. The system is regarded as operating only if at least s out of k ($1 \leq s \leq k$) strength variables exceeds the random stress. Otherwise the system will fail to function as intended. As an example for the proposed model, consider a multi axle trailer. It may have either six- or four-axle lines, with each axle line consisting of multiple wheels. They are used for transporting massive objects such as large bridge sections, oil refining equipment and other objects that are too big or heavy for trucks. The multicomponent reliability of the system is given by $R_{s,k} = P(\text{at least } s \text{ out of the } (U_1, \dots, U_k) \text{ exceeds } T)$ where $U_i = \min(Z_{i1}, Z_{i2})$ and $Z_{i1} = \max(X_{i1}, X_{i2}), Z_{i2} = \max(Y_{i1}, Y_{i2}), i = 1, \dots, k$. We estimate $R_{s,k}$ by using frequentist and Bayesian approach. The Bayes estimate of $R_{s,k}$ have been developed by using Lindley's approximation and the Markov Chain Monte Carlo methods due to the lack of explicit forms. The reliability estimators compared by using the estimated risks through Monte Carlo simulations.

Key Words: *Bivariate Generalized Exponential Distribution, Stress-strength Model, Multicomponent Reliability*

Multiple Regression Analysis Based on ANN Using with The Different Type of Error Function

Asiye Zühal Küçükmustafa

The University of Avrasya,
Higher School of Applied Sciences,
The Department of International
Business and Trade
zuhalkucukmustafa@avrasya.edu.tr

Mehmet Arif Demirci

The University of Muş
Alparslan, Higher School of
Health, The Department of
Healthcare Management
ma.demirci@alparslan.edu.tr

Vedide Rezan Uslu

The University of Ondokuz
Mayıs, Faculty of Art &
Science, Department of
Statistics
rezzanu@omu.edu.tr

ABSTRACT

Regression analysis is one of the most powerful statistical techniques as long as assumptions of the model given are not violated by the data. When the errors exhibit correlation, heteroscedasticity, non-normality, or the data have outliers or there is a multicollinearity problem among independent variables, the interpretations on the estimates of the regression coefficients, the confidence intervals or the fitted values can cause misleading results. Residuals analysis can help the researcher to understand whether the assumptions are violated or not. When the case is occurred the researcher has to get rid of the problem. Many remedies can be tried. One of the remedies purposed in the literature is to apply robust regression since it does not require the restrictive assumptions. In this study as an alternative robust regression technique we propose to use Artificial Neural Network (ANN) with differently defined fitness function. By diversifying the fitness function of ANN we aim to compare to the performances of the algorithms for the test data. For this purpose we have used three differently defined error function based on Least Squares, Least Absolute Deviation and Least Median Squares in ANN algorithm. The proposed methods have applied to the real data set, which are problematic in the sense of assumptions and the obtained results have been compared with respect to some types of accuracy measures.

Key Words: *Artificial Neural Network, Robustness, Least Median Square and Least Absolute Deviation*

Network Meta-Analysis for Decision Making in Medicine

Didem Derici Yıldırım
Mersin University Department
of
Biostatistics and
Medical Informatics
didemderici@hotmail.com

Damla Hazal Saral
Mersin University
Department
of Biostatistics and
Medical Informatics
dmlhz15@gmail.com

Bahar Taşdelen
Mersin University Department of
Biostatistics and Medical
Informatics
bahartasdelen@gmail.com

ABSTRACT

Pairwise meta-analysis compares the relative efficiency of two interventions or drugs based on direct evidence obtained from randomized controlled trials. For many diseases, there are multiple treatment alternatives. This situation makes difficult to investigate treatment efficiency indirectly. Decision making is too complex to determine just pairwise meta-analysis. A novel method has been developed for such kind of difficulties.

Network Meta-Analysis (NMA) compares the direct and indirect effectiveness of multiple treatments in similar patient populations. NMA has a complex and Bayesian structure contrary to frequentist approach of traditional meta-analysis. Besides this, NMA allows the visualization of more evidence and estimating the effect size of indirect comparisons among all treatments. Fixed and random effects models are used for comparison but more comprehensive than pairwise meta-analysis.

For systematic review of network meta-analysis, two databases should be searched (PUBMED, Cochrane Library ex.) for appropriate evidence as minimum. Thus, we aim to introduce the key features and assumptions of network meta-analysis, methods and statistical software used in NMA. We illustrate its implementation in a clinical example dataset and interpret the results.

Key Words: *Network Meta-Analysis, Multiple Treatments, Decision Making*

New (r, Q) Inventory Control Models in the Presence of Misplaced items

G.Yazgi Tütüncü
İzmir University of Economics
Department of Mathematics,
İzmir, Turkey
yazgi.tutuncu@ieu.edu.tr

Ceki Franko
İzmir American Collegiate
İzmir, Turkey
cekfranko@yahoo.com

Linda Zhang
IESEG School of
Management, Paris,
France
l.zhang@ieseg.fr

ABSTRACT

In the literature most available (r, Q) inventory control models assume that the actual inventory records are the same as these in the computer systems. In fact, misplaced items, i.e., items that are unavailable on the correct shelves or storage areas though they are in a computer system, are inevitable in any retail store in practice. The presence of the misplaced items results in the discrepancy between physical and recorded inventory levels. With the above unrealistic assumption, the available (r, Q) models may, thus, distort the inventory decision making in practice.

In this study, we develop new (r, Q) inventory control models considering the misplaced items to provide the retail stores with realistic decision-making support. To cover the practical situations as comprehensive as possible, we consider two cases in developing the new models: backorders and lost-sales. In both cases, we consider stochastic demand. We also propose solution algorithms for model solving. Numerical examples are conducted to demonstrate the applicability and potential of the new (r, Q) models and solution algorithms in helping retail stores make realistic inventory decisions. Based on the new models and the results of numerical examples, we further obtain managerial implications.

Key Words: (r, Q) inventory control, Misplaced items, Backorders, Lost-sales, Stochastic demand

Nonlinear Neural Network for Portfolio Optimization based on ARIMA

Ilgım Yaman
Affiliation and address
ilgimyaman@gmail.com

ABSTRACT

Standart Portfolio selection problem which is proposed by Hanry Markowitz in 1952 is very popular financial problem in optimization world. In the Portfolio optimization problem main goal is minimizing the risk, while maximizing the expected return of portfolio. Also, portfolio optimization problem is a quadratic optimization problem. Hence many heuristic methods were used to solve portfolio optimization method such as particle swarm optimization, ant colony optimization etc. Indeed, stock markets demands are not satisfied with these methods in financial world. A new neural network which is represented in 2014 by Yan based on solving primal and dual problems simultaneously. In this study, the nonlinear neural network is adopted to Markowitz's mean-variance portfolio optimization. Afterwards, ARIMA model is used to estimate expected returns of assets. Istanbul Stock Exchange-30 (ISE-30) data within 10.06.2015-14.05.2017 is taken from <https://tr.investing.com/>. Data comprise 480 days' expected ratio of 30 assets. In order to select optimum portfolio. Results are compared with calculated expected return based on ARIMA model

Key Words: *Statistics, Regression, Principal Components*

On the Efficiency of Proportion Estimators Under Modified Ranked Set Sampling Methods

Aylin Göçoğlu

The Graduate School of Natural and Applied Sciences, Department of Statistics, Dokuz Eylül University, İzmir, TURKEY
a.gocoglu@gmail.com

Neslihan Demirel

Faculty of Science, Department of Statistics, Dokuz Eylül University, İzmir, TURKEY.
neslihan.ortabas@deu.edu.tr

ABSTRACT

Ranked set sampling (RSS) is becoming a popular sampling method in recent years. The RSS method is an efficient alternative to simple random sampling (SRS) when the variable of interest is difficult, expensive or time consuming to measure but is easy to rank. The aim of this study is to estimate the population proportion and associated variance estimators for a binary variable ranked via a concomitant variable under modified ranked set sampling methods which are Median RSS, Extreme RSS, Pair RSS, Percentile RSS and LRSS. The Monte Carlo simulation study is performed via R Project with 10,000 repetitions. The performance of the estimators is constructed in terms of bias, mean squared error and relative efficiency for different set and cycle sizes for Normal and Log-Normal distributions. The results indicate that the proportion estimator under modified RSS methods performs better than the proportion estimator under SRS. Moreover, the study is supported with a real data example.

Key Words: *Ranked Set Sampling, Proportion Estimator, Relative Efficiency*

On the General Formula of Renewal Function of Erlang Distribution

Ozlem Sevinc (Ardic)

Central Bank of the Republic of
Turkey,
Anafartalar Mah. Istiklal Cad. No:10,
06050, Ulus, Altindag, Ankara
ozlem.sevinc@tcmb.gov.tr

Tahir Khaniyev

TOBB University of Economics and
Technology, Department of
Industrial Engineering, Sogutozu
Cad. No:43, 06650, Ankara
tahirkhanyev@etu.edu.tr

ABSTRACT

A great number of problems arising in the theories of stock control, mathematical insurance, queuing theory, reliability, stochastic finance and so on can be expressed with the aid of renewal processes, renewal-reward processes, random walk processes and etc. Laplace transformation is widely used for deriving probabilistic characteristics of the renewal processes. Especially, renewal function which is the expected value of number of renewals on a given time has great value on renewal processes. Although the renewal function is so important for renewal theory, it is so hard to compute the exact expression of renewal function because of n -fold convolutions. There are various studies about computing renewal function with asymptotic and approximation methods. In this study, with the help of Laplace transformation method, the exact formulas of renewal function ($U_n(x)$) generated by n^{th} degree Erlang distributed random variables with α parameter are derived.

Key Words: *Renewal Function, Erlang Distribution, Laplace Transformation*

On the Renormalized Entropy-based Complexity Analysis by BIST100 Data

Adnan Karaibrahimoglu
Süleyman Demirel University,
Faculty of Medicine, Dept. of
Biostatistics and Medical
Informatics
adnankaraibrahim@gmail.com

Özgür Afşar
Ege University, Faculty of
Science, Dept. of
Mathematical Physics
ozgur.afsar@ege.edu.tr

Şenay Çelikörs
Süleyman Demirel
University, Dept. of
Biostatistics and Medical
Informatics
senaycelikors@gmail.com

ABSTRACT

The BIST100 index is used as the main index for Borsa İstanbul Equity Market. It consists of 100 stocks selected among the stocks of companies traded on the BIST Stars and BIST Main markets. The aim of this study is to analyze stock exchange data of the opening session prices in BIST100 Index by renormalized entropy-based complexity analysis. It is well known that standard entropy-based complexity measures such as Shannon and relative (Kullback-Leibler) entropies possess always positive values and second law of thermodynamics also says that entropy of a closed system has to increase. Since a stock market is not a closed system, affected by many environment conditions, it can't be evaluated within usual concepts. Our first aim is to compare two cases by equating their mean energies so that the cases obey the ordinary closed system formalism. Moreover, it is fact that productive stocks need to increase their organizational order. Second aim is to check this condition for the stocks in this study. Therefore, one can clearly explain why a system decreases its renormalized entropy, that implies increasing order, in the context of self-organization. Third aim is to build a predictive modeling for stocks whether a stock will bring in or lost for its share dealers.

The data was obtained from the legal website of Istanbul Stock Exchange (BIST) by an official application after permission of the authority. The data consists of the opening session prices of 100 stocks between the dates of Jan. 1st 2014 and Dec. 30th 2016, covering 3 years period. However, some stocks were not active during 3 years period; therefore, these stocks were omitted from the data. At the end, there were 91 stocks, 666 daily prices and the information about the positivity of the yield categorized as "positive and negative". Calculation of renormalized entropy and the other entropies (namely, Shannon and Kullback-Leibler) requires the estimation of distributions. In order to achieve this, we use the autoregressive spectral distributions in our calculations.

We analyzed only first four stocks due to time constraint and, as it is expected, renormalized entropy has given us positive results for lost stocks (indicating the decreasing order) and negative results for profitable stocks (increasing organizational order).

Key Words: *Shannon entropy, Renormalized entropy, Autoregressive model, Stock Exchange index*

Optimal Bandwidth Selection for a Kernel Density with a Location-Scale Property

| | | |
|--------------------------|--------------------------|--------------------------|
| Celal Aydın | <u>Necla Gündüz</u> | Jale Balibeyoğlu |
| Gazi University | Gazi University | Gazi University |
| Department Of Statistics | Department Of Statistics | Department Of Statistics |
| Ankara-TURKEY | Ankara-TURKEY | Ankara-TURKEY |
| aydinc@gazi.edu.tr | ngunduz@gazi.edu.tr | ngunduz@gazi.edu.tr |

ABSTRACT

In kernel density estimation, the most important step is to select a proper bandwidth. In this study, classical kernel smoothing and adjusted kernel smoothing is considered. Adjusted kernel smoothing is applied by using kernel density from a member of location-scale family. It is believed that using the kernel function from location-scale family provide some kind of control on the bias of kernel smoothing. By using the minimizing mean squared error as a criteria for kernel smoothing, the optimal bandwidth exist as a function of density functionals, therefore, their estimates are needed.

A comparison of performance of the two kernel density estimation approach is made by simulation study for some different distributions. Additionally, these two approaches are applied to a real data and the results are summarized.

Key Words: *Kernel estimation, Location-scale family, Bias, Density functional estimate, Optimal bandwidth.*

Optimal Determination of the Parameters of Some Biased Estimators Using Genetic Algorithm

Erkut Tekeli

Department of Computer Technologies, Kozan
Vocational School, Çukurova University,
Kozan / Adana, Turkey
etekeli@cu.edu.tr

Selahattin Kaçıranlar

Department of Statistics, Faculty of Science
and Letters, Çukurova University, Adana,
Turkey
skacir@cu.edu.tr

ABSTRACT

Linear regression model and ordinary least squares (OLS) method are widely used in many fields of natural and social science. In the presence of the multicollinearity, OLS estimator is unstable and often gives misleading information. Ridge, Liu and two-parameter estimators are some alternative to the OLS method when multicollinearity is present in multiple linear regression model. The selection of the parameters is an important issue to give good performance of the estimators. In this paper, a new algorithm of estimating k value for ridge estimator, d value for Liu estimator and k and d values for two-parameter estimator were introduced by using genetic algorithm (GA). This methods are based on minimizing the statistic measures; mean square error (MSE), mean absolute error (MAE), mean absolute prediction error (MAPE) using GA and managing the condition number and variance inflation factors values to be less than or equal to 10. A numerical example is presented by using different estimators of the biasing parameters to illustrate the results. We showed that the new method permit to reduce the multicollinearity problem and it minimizes the MSE, MAE and MAPE. In addition to the numerical example, an extensive Monte Carlo experiment is conducted.

Key Words: *Liu Estimator, Ridge Estimator, Two Parameter Estimator, Genetic Algorithm, Parameter Optimization*

Optimising the Biasing Parameters of Liu Type Estimator Based on Ridge Regression Using with Genetic and Differential Evolution Algorithm

Mehmet Arif Demirci

The University of Muş Alparslan,
Higher School of Health,
The Department of Healthcare
Management
ma.demirci@alparslan.edu.tr

Vedide Rezan Uslu

The University of Ondokuz Mayıs
Faculty of Art & Science
Department of Statistics
rezzanu@omu.edu.tr

ABSTRACT

When there is severe multicollinearity problem, which refers that predictors are correlated with other predictors in the model, it causes the variances of the regression parameter estimates increase. When this case is occurred the estimates can be very sensitive to minor changes in the model. In the literature there are some alternative techniques to least squares. Most of them provide biased estimates of the regression coefficients. In order to achieve consistency from sample to sample these generally aim to get much smaller variances for the parameter estimates while they produce a bias for the parameter estimate. One of them is known Liu Type estimator in linear regression model. It has got two parameters k and d known as biasing parameters. In this study we propose a new approach that finds the biasing parameters of Liu Type estimator based on Ridge Regression estimator with heuristic algorithms. In addition to the novelty, the proposed approach defines the objective function based on Mean Prediction Error, providing to keep Variance Inflation Factor and Condition Number around the level of literately accepted. The proposed approach has applied to the real data which has got severe multicollinearity problem and the results have been compared according to some accuracy measures.

Key Words: *Liu Type Estimator, Ride Estimator, Genetic Algorithm, Differential Evolution Algorithm*

Outlier Detection and Variable Selection for Count Data

Fatma Sevinç KURNAZ
Yildiz Technical University
Faculty of Arts and Sciences
Department of Statistics
fskurnaz@yildiz.edu.tr

ABSTRACT

Maximum likelihood (ML) estimator is a very common method in Poisson regression for count data. However, it is well known that quite often the real data sets contain a significant level of outliers and ML estimator can give misleading information in such a case. The purpose of this study is to introduce a new method which is able to do both outlier detection and variable selection at the same time for poisson regression. The proposed method is illustrated by an extensive simulation study as well as a real data example.

Key Words: *Poisson Regression, Outliers, Lasso Type, Robustness*

Parameter Estimation for Uniform-Geometric Distribution Based on Censored Sample

Mehtap Koca Yılmaz Yunus Akdoğan Kadir Karakaya
Selçuk University Selçuk University Selçuk University
mektaps@hotmail.com yakdogan@selcuk.edu.tr kkarakaya@selcuk.edu.tr

ABSTRACT

In this study, we obtained estimators for the unknown parameters of Uniform-Geometric distribution (UGD) based on censoring sample using estimation methods which are maximum likelihood estimation(MLE), method of moments(MM) and method of proportions(MP). Performance of the estimators are compared using Monte Carlo simulation. A real data example is also provided.

Key Words: *Censored Sample, Estimation, Monte Carlo Simulation, Uniform-Geometric Distribution*

Parametric Joint Modelling for Longitudinal and Survival Data

Elif Dil Duru Karasoy
Hacettepe University Hacettepe University
elifdil@hacettepe.edu.tr Duru@hacettepe.edu.tr

ÖZET

Longitudinal data are obtained by measuring repeatedly the properties of the same units over time. The purpose of this data set analysis is to change the dependent variable over time and to identify the factors that cause this change and parametric regression, nonparametric regression and semi parametric regression methods are used. Survival data includes the time from the beginning of a given event to the occurrence of a certain event and in the analysis of this data set semi parametric regression and parametric regression methods are utilized.

Longitudinal and survival data are analyzed in the context of study but they are usually analyzed separately. However, combined with survival time of the units and the longitudinal observation related to this survival time, a joint model should be established to obtain unbiased and effective results. Joint model is obtained by connecting longitudinal sub-model and survival sub-model with joint random effects. In general, joint model is created using linear mixed model for longitudinal sub-model and Cox regression model for survival sub-model. Nevertheless, in cases where the survival time has a specific parametric distribution, joint model should be created using parametric survival sub-model.

In this study, a joint model of longitudinal and survival model with different parametric distribution is performed. The model is used to test the efficacy and reliability of two alternative medicines to Zidovudin treatment over 467 individuals with Aids in Australia before July 1, 1991 and the results are interpreted.

Key Words: *Parametric Survival Analysis, Longitudinal Data, Survival Data and Joint Model*

Pattern Recognition in Control Charts with Neural Networks

Alper Vahaplar
Dokuz Eylül University,
Department of Computer Science,
İzmir / Türkiye
alper.vahaplar@deu.edu.tr

Senem Şahan Vahaplar
Dokuz Eylül University,
Department of Statistics,
İzmir / Türkiye
senem.sahan@deu.edu.tr

ABSTRACT

Quality control charts are among popular and efficient tools of statistical process control. They contain valuable information about the current situation of an on-going process. The idea behind them is monitoring the process and analyzing the variability. Unnatural patterns in control charts show that the process is affected by assignable causes and a process is an out-of-control process when assignable causes are present. So for this, patterns on control charts should be evaluated carefully. Pattern recognition aims to discover nonrandom or systematic patterns and their reasons.

Neural networks are parallel computing mechanisms that derive meaning from complicated data. They can be used to detect specific patterns. Neural network based pattern recognition has been studied by several researchers. This study describes the possible patterns that can be observed in a control chart and offers a neural network analysis for detecting patterns in the processes by analyzing the control chart measurements.

Key Words: *Statistical Process Control, Control Charts, Pattern Recognition, Neural Networks*

Penalized Logistic Regression for Classification: A Case Study for non-OECD Countries

Ezgi Nazman
Gazi University, Faculty of
Science, Department of
Statistics, 06500, Ankara
ezgicabuk@gazi.edu.tr

Semra Erbaş
Gazi University, Faculty of
Science, Department of
Statistics, 06500, Ankara
serbas@gazi.edu.tr

Hülya Olmuş
Gazi University, Faculty of
Science, Department of
Statistics, 06500 Ankara
hulya@gazi.edu.tr

ABSTRACT

Data Envelopment Analysis (DEA) has been used to measure efficiencies of the Decision Making Units (DMUs) containing multiple inputs and outputs. Therefore, DEA has been combined with multivariate statistical methods. Thanks to DEA, DMUs can be classified into two different main groups according to their efficiency scores. These two groups have been a guide to evaluate DEA classification results using binary Logistic Regression (LR). Not only binary LR tends to estimate parameters as infinite but also increase in bias of estimated model parameters when number of efficient DMUs is low. In literature, Firth's penalized LR has been suggested for rare event situation instead of binary LR because it reduces first-order bias in maximum likelihood estimations of parameters while allowing computation of parameter estimations. In this study, efficiency situations of non-OECD countries were determined by using DEA. Then, bias in predicted probability of binary LR and Firth's penalized LR according to DEA classification results were evaluated with Root-Mean-Squared Error (RMSE).

Key Words: *Data Envelopment Analysis, Firth's Penalized Logistic Regression, Binary Logistic Regression, Rare Event, Non-OECD Countries*

Penalized L_2E Estimator for Gaussian Mixture Model

Fikriye Kabakci

Faculty of Arts & Sciences
Department of Mathematics
Recep Tayyip Erdogan
University
fikriye.kabakci@erdogan.edu.tr

Yüksel Öner

Faculty of Arts & Sciences
Department of Statistics
Samsun Ondokuz Mayıs
University
yoner@omu.edu.tr

Umashanger Thayasivam

Department of Mathematics
Rowan University
USA
thayasivam@rowan.edu

ABSTRACT

This work was intended to give short review on parameter estimation of Gaussian mixture model (GMM) with L_2E . In the literature consistent estimators for determining the number of components in GMM proposed by penalized likelihood method by Huang, et al.(2013). Besides, Lee, J. (2010) proposed penalized L_2E estimator to estimate mixture complexity for non-normal mixtures. The goal of our study is to define penalty approach to L_2E for model parameter estimation.

We wish to modify L_2E by applying a penalty for the weights of GMM as well as the covariance structure to improve the robustness of variance elements. Theoretical approach is verified by some numerical experiments.

Key Words: *Finite Mixture Model, L_2E , Mpd, Robustness, SCAD Penalty*

Performance Comparisons of Some Classification Algorithms Used in Machine Learning

Bünyamin Sarıbacak
University of Ondokuz Mayıs,
Faculty of Education,
Department of Computer
Education and Instructional
Technology,
Samsun, Turkey
bunyamin@omu.edu.tr

Erol Terzi
University of Ondokuz Mayıs,
Faculty of Science and Art,
Department of Statistic, Samsun,
Turkey
eroltr@omu.edu.tr

ABSTRACT

The classification is the distribution of the data between the various classes defined on the data set. Classification algorithms learn this form of distribution from a given set of training and then try to classify it correctly when it comes to test data for which the class is not specified. The values that specify these classes on the dataset are called labels and are used to determine the class of data during both training and testing. In this study, mushroom records from the Audubon Society Field Guide of North American Mushrooms (1981) were used. The fungi are divided into two classes as poisonous and edible. It has also been found that there is no simple rule to determine the edibility of a mushroom. Mushrooms were tried to be classified using machine learning classification algorithms such as ZeroR, OneR, Naive Bayesian, Decision Tree, K Nearest Neighbors, Artificial Neural Network and Support Vector Machine algorithms. Weka 3.8 software is preferred for Machine Learning application. In the resultant comparison tables, high classification performance was achieved with very small differences among some methods. In particular, it has been observed that the success rates of Decision Tree, K Nearest Neighbors and Support Vector Machine algorithms are very high.

Key Words: *Machine Learning, Weka, Classification, Classification Algorithms*

Practicality of Some Variations of Ranked Set Sampling

Monjed H. Samuh
Palestine Polytechnic University
Hebron, Palestine
monjedsamuh@ppu.edu

M. Hafidz Omar
King Fahd University of
Petroleum & Minerals
Kingdom of Saudi Arabia
omarmh@kfupm.edu.sa

M. Pear Hossain
Bangabandhu Sheikh Mujibur
Rahman Science & Technology
University
Gopalganj, Bangladesh
pearhossain35@bsmrstu.edu.bd

ABSTRACT

In this paper, we discuss the practicality of some variations of ranked set sampling (RSS) such as median RSS, double median RSS, and double RSS. The Hellinger distance is used as a measure of practicality. Although double median RSS is the most efficient approach among the RSS variations considered, it is shown in this paper that it is the least practical.

Key Words: *Efficiency, Hellinger Distance, Median, Practicality, Ranked Set Sampling*

Precursory Seismicity Rate Changes associated with the earthquakes in the East Anatolian Region of Turkey: An Application of Region-Time-Length (RTL) Algorithm

Serkan Öztürk
Gümüşhane University,
Department of Geophysics
serkanozturk@gumushane.edu.tr

Santi Pailoplee
Chulalongkorn University,
Department of Geology
Pailoplee.S@gmail.com

ABSTRACT

There are a number of statistical methods for the analyzing, appraising and mapping the space-time analyses of precursory seismicity rate changes on the earthquake occurrences. These models suggest that specific regional and temporal patterns including the phenomenon of precursory quiescence (decrease) and activation (increase) against the background seismicity may be related with the seismic and tectonic process. Many researchers use different approaches and parameters such as β -value, Z -value, Relative Intensity (RI), Pattern Informatics (PI) and Region-Time-Length (RTL). However, RTL method is a weighted coefficient statistical approach that use the i) location (R), ii) occurrence time (T), and iii) magnitude related with rupture length (L) of earthquake when earthquake rate changes before large earthquakes are being investigated. The basic assumption of this approach is that each prior earthquake has some influence on the main event under investigation, and positive values of RTL represent seismic activation whereas negative values point out seismic quiescence. Thus, this algorithm is an effective tool to detect the precursory seismic quiescence. Considering the potential applications of RTL algorithm in the earthquake forecasting studies, we aimed to detect the possible areas of the next earthquakes occurrences in the East Anatolian Region of Turkey by using this algorithm. We used KOERI (Kandilli Observatory and Earthquake Research Institute) catalog including 44,359 shallow crustal earthquakes with duration magnitude $M_d \geq 1.0$ from 1970 until 2018 is utilized. In order to check out the relation between RTL score and the forthcoming earthquake occurrences, we selected seven earthquakes with $M_d \geq 6.0$ as case studies and we analyzed retrospectively these events temporarily and spatially with RTL algorithm. After a number of iterative tests, the most suitable characteristic parameters for the East Anatolian Region of Turkey, i.e., radius (R_{max}) and time (T_{max}), were considered as 110 km and 4.5 year, respectively. This means that the clearest RTL anomalies in the temporal investigation are obtained with these values. Based on detailed temporal investigation, all the case studies show clear quiescent anomalies. However, based on spatial investigation, some case studies do not conform spatially to the mentioned anomalies. It is well known that the East Anatolian part of Turkey is one of the most seismically and tectonically active regions. Thus, since many large earthquakes occurred in the past and recent years, more detailed investigations for seismic hazard assessment have become more important.

Key Words: East Anatolia, RTL Algorithm, Precursory Quiescence, Seismicity Rate Changes

Preparation MRI imageable novel HEC-PVP-Magnetic Embolic Particles with anticancer drug release for Transcatheter Arterial Embolization

Maide Gökçe Bekaroğlu

Istanbul Technical University,
Faculty of Science and Letters,
Dept. of Physics
gbekaroglu@live.com

Fuat Nurili

Department of Radiology,
Memorial Sloan Kettering
Cancer Center
fuatnurili@gmail.com

Sevim İşçi

Istanbul Technical University
Faculty of Science and Letters
Dept. of Physics
iscisev@itu.edu.tr

ABSTRACT

Embolization with drug-loaded particles is a relatively novel method and have attracted increasing interest in the treatment of many diseases. These drug-loaded particles cut off the nutrition and oxygen supply necessary for tumor survival by blocking the blood flow toward tumors and releasing chemotherapeutic agents in a sustained manner thereby achieving double-effect with devascularization and local chemotherapy for progressive shrinkage of the tumor. Micron sized embolic materials allow for more precise control of the level of occlusion and transcatheter arterial chemoembolization (TACE) using drug-loaded microparticles showed less systemic toxicity and drug-related side-effects.

Stabilization of colloidal particles would require understanding of polymer behaviour in solution as well as at the colloid/solution interface for a complete quantitative description. Polymers can act as flocculant as well as dispersant for the colloidal particles depending on the properties and the concentration of the polymer. At the low concentration ranges, the adherence of the particles to polymer, resulting to a fractional coverage, causes to big flocs and networks in the medium called interparticle bridging or bridging flocculation. Further addition of polymer to the dispersion increases the polymer concentration beyond to an optimum concentration of flocculation and also increases the osmotic pressure in the medium. In this case, adhering to the particle surface cause the stabilization and enveloped surfaces resist to a further approach of particles to each other.

MRI imageable novel HEC-PVP-Magnetic Embolic particles intended for TACE and Magnetic hyperthermia were prepared from Iron Oxides linked with PVP using bridging flocculation process. Additional HEC were incorporated (coated) onto the particles to load a drug. All synthesized particles calibrated in the 40-500 microns range and easily injectable through a microcatheter. An anticancer drug, doxorubicin, was successfully incorporated into these particles for local release and thus chemotherapy-induced cell death could be achieved. The particles were determined to exhibit no toxicity against hFOB cells. In addition, a preliminary study on the in vivo embolization with these particles performed in a rabbit renal embolization model, demonstrated that these particles were compatible with, capable of occluding the arteries, and were imageable with MRI.

Key Words: *Colloidal particles, flocculation, stabilization, drug delivery,*

Pricing Of Long-Term Care Insurance For Turkey

Çiğdem Lazoğlu
Hacettepe Üniversitesi
cigdemkobal@hacettepe.edu.tr

Murat Büyükyazıcı
Hacettepe Üniversitesi
muratby@hacettepe.edu.tr

ABSTRACT

An increase in the incidence and duration of chronic illnesses have occurred with the extension of the expected life span. Chronic diseases lead to deterioration of the physical condition of the patients. These disorders cause individuals to be unable to meet their daily needs and the quality of life to deteriorate. The patient needs other a person or people to meet their daily needs. Life expectancy in Turkey is increasing. Expected increase in life expectancy will increase the demand for care services. Increasing demand will put pressure on the state budget. As a solution to this problem long-term care insurance models should be developed for older people.

Long-term care insurance is usually modeled with simple Markov, which is a single state for dependency. But the degree of disease for elderly individuals vary from day to day. The degree of disease affects the duration in dependency. Therefore, it is inadequate to explain in a single situation the duration in state of dependency. In this study, the long-term care insurance model, which is in the literature, has been adapted to Turkey. The state of dependency is divided into four according to the degree of dependency and the transition probabilities are assumed to be semi-Markov process. Transition possibilities in the semi-Markov process are not only dependent on the current state, but also on the duration of the present state.

For the duration of dependency, parameter values existing in the literature are used. The life span of healthy individuals for [40,60] age range was estimated by Monte Carlo simulation using the general mortality rate in Turkey and the incidence rate of France becoming in dependency from a healthy state. The study is based on the assumption that the probability of death does not change over time. Premiums and reserves for each age are calculated by cash flow method.

Key Words: *Long-Term Care Insurance, Monte Carlo Simulation Method, Semi-Markov Process, Cox Proportional Hazard Model, Frailty Model*

Probability Density Function Estimation Using Scaled Inverse Chi-Squared Kernel

Elif Erçelik
Istanbul Technical University,
Turkey
ercelikel@itu.edu.tr

Mustafa Nadar
Istanbul Technical University,
Turkey
nadar@itu.edu.tr

ABSTRACT

Standard kernel estimators were developed primarily for the densities with unbounded support. However, symmetric kernel estimator is not an appropriate choice for fitting densities with compact or semi-infinite support as it causes boundary bias due to the allocation of weight outside the support when smoothing applied near the boundary. Several methods have been proposed to fix this problem. One possible approach proposed by Chen, which we also follow, to solve the boundary bias problem is to use nonparametric estimators based on kernels of asymmetric distributions with nonnegative support. In this work, we propose a new kernel estimator for estimation of the probability density functions which is based on scaled inverse chi-squared probability density function used as a kernel. Similar to the other kernel estimators, proposed one is free of boundary bias, nonnegative with natural varying shape and achieves the optimal rate of convergence $O\left(n^{-\frac{4}{5}}\right)$ for the mean integrated squared error. Finite sample comparisons with other kernel estimators are reported for different distributions and sample sizes via Monte Carlo simulations.

Key Words: *Boundary Bias, Density Estimation, Scaled Inverse Chi-squared Kernel*

Regression Mixture Model with Two Components for Count Data: An Application to Traffic Accidents

Hande Konşuk Ünlü

Ayten Yiğiter

Hilal Özcebe

Burcu Küçük Biçer

Hacettepe University,
Institute of Public Health

Hacettepe University,
Faculty of Science,
Department of Statistics

Hacettepe University,
Faculty of Medicine,
Department of Public
Health

Yukse İhtisas
University, Faculty of
Medicine, Department
of Public Health

hkonsuk@hacettepe.edu.tr

yigiter@hacettepe.edu.tr

hilalozcebe@gmail.com

drburcubicer@gmail.com

ABSTRACT

In the modeling of count data, Poisson distribution is frequently used. In the presence of explanatory variable information, Poisson regression model is applied. However, the feature of equality of expected value and variance of the Poisson distribution could not be always met in real life data. For example, the variance may be greater than the expected value, which is called as over-dispersion. For this case, the Negative Binomial distribution which is the mixed distribution of Poisson and Gamma is used. Another situation that can be encountered is the zero inflated data structure. In this situation, zero-inflated Poisson model or zero-inflated Negative Binomial model could be used. Sometimes, however, the data may be a heterogeneous structure consisting of a mixture of multiple populations, rather than a single population. Mixed models are used to analyze such heterogeneous data.

In this study, the number of injury cases in traffic accidents resulting in death or injured was modeled by using a new mixed distribution with two components. In addition, under the presence of explanatory variable the regression mixture model with two components (Poisson and Zero-truncated Poisson) is proposed and parameter estimations of the model are obtained by using the EM algorithm.

Key Words: *Poisson, Zero Truncated Poisson, Poisson Regression Mixture Distribution, EM Algorithm*

Reliability Analysis of a Multi-state System with Identical Units Having Two Dependent Components

Funda Iscioglu

Ege University, Faculty of
Science, Department of Statistics,
3550, Izmir, Turkey
fundaiscioglu@gmail.com

Aysegul Erem

Cyprus International University,
Faculty of Arts and Sciences,
Lefkosa, Mersin 10, Turkey
aysegulerem@gmail.com

ABSTRACT

In most real life applications, a system and its components can have a range of performance levels from perfect functioning to complete failure. Multi-state modelling and the reliability analysis is then used to model and evaluate the performance of such systems. We deal with the performance evaluation of a multi-state system having n identical units, each consists of two components. The components of each unit are assumed to be s -dependent. However the units work s -independently with each other. We let the system and each component of each unit have performance levels "0, 1 and 2" where "0" indicates "failure", "1" indicates "partially working" and "2" indicates "completely working" states. The degradation of the components follows Markov Process and also FGM distribution is used to model the s -dependence of the components. We evaluate the reliability analysis of a multi-state k -out-of- n : G system under the assumptions given and also evaluate some dynamic performance measures for the system such as the mean residual and mean past lifetime functions of that system based on the use of bivariate order statistics. The performance of the system is especially examined depending on the s -dependence parameter change, the degradation rates' change and also for different number of units for the system. The results are supported with some numerical examples and graphical representations.

Key Words: *Multi-state System, Order Statistics, Reliability Analysis*

Restricted Bridge Estimator

Hüseyin Güler
Çukurova University,
Faculty of Economics
Administrative Sciences,
Department of Econometrics
hguler@cu.edu.tr

Ebru Özgür Güler
Çukurova University,
Faculty of Economics
Administrative Sciences,
Department of Econometrics
ozgurebru@cu.edu.tr

ABSTRACT

Parameters of a linear regression model might be estimated with Ordinary Least Squares (OLS) if the assumptions of the model are met. In addition to that, if there are linear restrictions on model parameters, it is possible to use Restricted Least Squares (RLS). However, both OLS and RLS suffers from multicollinearity if the predictors of the model are highly correlated or if the number of predictors is large.

A solution to multicollinearity is to use biased estimators like Ridge Regression (RR). RR shrunk the model parameters to zero with the help of a biasing parameter set by the researcher. However, for sparse regression models where a part of the model parameters are actually zero, RR estimates of model parameters doesn't have to be zero. This means that RR is not suitable for model selection and parameter estimation for a sparse model.

In contrast to RR, Least Absolute Shrinkage and Selection Operator (LASSO) performs well in sparse models, that is, LASSO can be used for model selection and estimation simultaneously. A generalized form of LASSO is called Bridge estimator and it includes LASSO, OLS, and RR for different penalty functions. It is shown in the literature that Bridge estimator outperforms LASSO, OLS, and RR estimators under different settings.

A Restricted LASSO (RL) estimator is proposed in the literature for linear regression model with linear restrictions. In this paper, we define a new estimator called Restricted Bridge (RB) that satisfies the linear restrictions of the model. We compare RB estimator with RL, Bridge, LASSO, and OLS estimators according to correct model selection percentage and mean squared error criteria with Monte Carlo simulations. We also implement aforementioned estimators to prostate dataset that is widely used in the literature for an application.

Key Words: *Restricted Bridge Estimator, Restricted LASSO, Sparse Regression, Model Selection, Monte Carlo Simulation*

Results of a Project of Marriages Between Disabled People

Trabzon İl Sağlık Müdürlüğü
Gülbahar Hatun Mah. İnönü Cad.
Ahmet Can BALI Sok. No:15
Ortahisar / TRABZON
trabzon.ebys@saglik.gov.tr

Ceyhun MAZLUM
Gülbahar Hatun Mah. İnönü Cad.
Ahmet Can BALI Sok. No:15
Ortahisar / TRABZON
cyhnmzlm@gmail.com

Selma YILDIZ
Gülbahar Hatun Mah. İnönü
Cad. Ahmet Can BALI Sok.
No:15 Ortahisar / TRABZON
ktuselma@gmail.com

ABSTRACT

Turkey's disabled community comprises about % 12,29 of the population. Although, this rate is considerably high for our country it is hard to claim that the necessary precautions. In this sense, it could not be claimed that enough awareness toward disabilities has not been raised. Generally, speaking about the research findings taking into consideration the information which must be known about disability it is phenomenon as well as it is reason we summarised the ways of disability avoidance and prevention in medical fields as follows;

This study's aims are;

- 1) To obtain statistically interpretable data about disabled people
- 2) To do a research about what could be done to prevent transmission of genetic related illnesses
- 3) To study the necessity of Health Counselling Centres for the Disabled
- 4) To find out the most important issues that disabled people want to be studied

This study is done with 208 married disabled people in Trabzon Province and its cities. A questionnaire is used to reach the sample which targeted individuals with genetic and hereditary illnesses.

Of the 208 people who participated in this research 167 were male (%80) and 41 were female (%20). The participant of this study have got 379 children and among this children 25(%6,59) of them are disabled. The number of miscarriage was 26.

Of the 208 people who participated in our questionnaire 103 (%49,52) were born with congenital disabilities and 99 (%47,60) participants said that they become disabled later in their life. 6 (%2,88) participants did not answer this question. Due to some unanswered question in the questionnaire we could not find out whether these participants were genetic or hereditary disabled or whether they were born with disabilities after birth or not.

Of the 208 people who participated in this survey, 207 responded to the question about their type of disabilities.

| | |
|--|---|
| Sight impaired : 79 people (%38) | Mentally –emotionally disabled: 3 people (%1) |
| Hearing impaired: 8 people (% 4) | Mentally disabled: 8 people (%4) |
| Speech and hearing impaired: 21 people (%10) | Chronic illness disabled: 21 people (%10) |
| Orthopedically disabled: 66 people (%32) | Chronic illness – sight impaired: 1 person (%1) |

Most of the participants (%75,24) in this research stated that they are hopeful about their future. %65,86 of the participants said that they need to be informed about health issues. While, %36,14 of the participants believe that their children will be disabled, %29,70 of the participants stated that they are informed about their children's changes of being born with disabilities. On the other hand, %6.43 of the participants said that they have knowledge about Genetic Counselling Centre and %1,98 of them said that they were directed to the Centres.

As a result, since genetic transition has an important role in formation of disabilities, an establishment of disabled health counselling centre where disabled people can be informed about Genetic Counselling Centres before they apply to Medical Genetic Clinic Centres. Therefore, it will be possible to predetermine situations which cause disabilities and allow children to be born healthy.

Key Words: *Statistics, Descriptive Statistics, Cross Table*

Risk Evaluation for Support: Predictions for Elder-Life in Community Tool

Meltem Tuna
Program Manager / Lead Analyst
Ottawa Hospital Research Institute Ottawa, Ontario, Canada /
Institute for Clinical Evaluative Sciences, Ottawa and Toronto, Ontario,
Canada
The Ottawa Hospital, Civic Campus
1053 Carling Avenue, Box 684
Administrative Services Building - 1st Floor, Room 1014
Ottawa, ON K1Y 4E9
mtuna@ohri.ca

ABSTRACT

The Resident Assessment Instrument Home Care Assessment System (RAIHC) is a, person-centered assessment system that informs and guides comprehensive care and service planning in community-based settings. It focuses on the person's functioning and quality of life by assessing needs, strengths, and preferences, and facilitates referrals when appropriate.

Risk Evaluation for Support: Predictions for Elder-Life in Community Tool (RESPECT) was developed and validated using RAIHC data to predict the risk of death in 6 months. The cohort includes over 490 thousand Ontarians who received at least one home care assessment with over 1.3 million assessments between 2007 and 2014. The algorithm was developed on the first five years of the data and validated on the last year. Discrimination and calibration were satisfactory between the derivation and validation cohort, with a C-statistics of 0.75 and calibration plot intercept 0.015, slope 0.873.

Thousands of predictive health algorithms have been developed over the past 20 years, however less than 2% have been assessed outside their original setting and even fewer have been applied and evaluated in practice. To facilitate the dissemination and exchange of predictive algorithms in health our team has developed streamlined, semi-automated and standardized workflows and libraries.

The implementation of predictive analytics in real-world settings such as within electronic medical records or decision aid applications requires well described algorithms that are easy to replicate and implement, especially as access to big health data increases and algorithms become increasingly complex.

Key Words: *Home Care Assessment, Predictive Algorithms, Dissemination and Implementation of Algorithms*

Robust Estimators in Linear Regression: A Simulation Study and an Application on Healty Care

Neslihan Gökmen

Istanbul Technical University
ngokmen@itu.edu.tr

Aydın Erar

Department of Statistics, MSGSÜ
aydin.erar@msgsu.edu.tr

ABSTRACT

The least squares (OLS) estimation method is used to obtain unbiased and least variance estimators for the parameters in the linear regression. The OLS method has poor performance in the presence of outliers and in this case the choice of robust regression methods is recommended. In order to obtain robust estimators against outliers, robust regression methods have high breakdown point. In this study, a simulation study and an application were carried out to compare the robust regression methods with OLS using "rlm" and "lm" functions in R program. In addition, in the simulation study, different objective functions were selected and compared with the OLS. The sum of mean square error (total MSE) and coefficient of determination (R^2) was used to evaluate the model performance. In order to demonstrate the success of robust methods in detecting outliers, 98 male patients with type 2 diabetes mellitus between the ages of 18 and 80 years participated in a study and modeled with parameters that could affect the vitamin D level. The outliers were detected by robust methods and performance of the estimated OLS model was improved without outliers.

Key Words: *Robust Regression, MM Regression, M-Estimators*

Robust Estimation of Response Surface

Talha Arslan

Van Yuzuncu Yil University,
Department of Econometrics,
65080 Van, Turkey
mstalhaarslan@yyu.edu.tr

Sukru Acitas

Eskisehir Technique University,
Department of Statistics, 26470
Eskisehir, Turkey
sacitas@anadolu.edu.tr

Birdal Senoglu

Ankara University,
Department of Statistics, 06100
Ankara, Turkey
senoglu@science.ankara.edu.tr

ABSTRACT

Response surface is a well-known and widely used methodology incorporating mathematical and statistical optimization techniques in developing new products or existing processes [1]. In estimating the response surface, functional relationship between the controllable variable x and the response variable y is utilized via first or second order model. Traditionally, model parameters are estimated using the least squares (LS) method. However, LS estimators are not optimal under non-normality and in the presence of the outlier(s) existing in the data set. This leads to the inefficient estimates of the response surface. The motivation of this study comes from tire tread compound data [2,3]. We reanalyzed the mentioned data set and realize that error terms are distributed as Jones and Faddy's [4] skew t (JFST). We use Tiku's [5] modified maximum likelihood (MML) methodology to obtain the estimators of the model parameters; see also Acitas [6] in the context of multiple linear regression. Different than ML estimators, MML estimators have closed forms besides being easy to compute. Furthermore, they are robust to the outlier(s). At the end of the analysis, it is seen that the mean square error (MSE) value belonging to the model constructed using the proposed methodology is much smaller than the MSE value of the corresponding model based on normal theory. In other words, using MML estimators in constructing the response surface provides better accuracy.

Key Words: *Response Surface Methodology, Jones and Faddy's skew t , Modified Likelihood*

Robust Estimators in Generalized Linear Models with Binary Responses in the case of Influential Points

Burcu Mestav

Canakkale Onsekiz Mart
University, Faculty of Arts and
Science, Department of Statistics,
17020, Terzioğlu Campus,
Canakkale
burcumestav@comu.edu.tr

Kumru Urgancı Tekin

Canakkale Onsekiz Mart
University, Faculty of Arts and
Science, Department of Statistics,
17020, Terzioğlu Campus,
Canakkale
kumruurganci@comu.edu.tr

Neslihan Iyıt

Selçuk University, Faculty of
Science, Department of
Statistics, 42031, Alaeddin
Keykubat Campus, Konya
niyit@selcuk.edu.tr

ABSTRACT

Generalized linear models (GLMs) are described as assembly a class of regression models for discrete, categorical and continuous response variables by Nelder and Wedderburn (1972). GLMs are widely used in many fields and parameter estimations in GLMs are made by maximum likelihood (ML) method. But it is known that ML estimator for these models is very sensitive to influential points. Observations can potentially be influential in different ways. If an observation has a response value that is very different from the predicted value based on the model, then that observation is called an outlier. On the other hand, if an observation has a particularly unusual combination of predictor values, then that observation is said to have high leverage. It is also possible for an observation to be both an outlier and have high leverage. To overcome these situations, robust estimators for GLMs have been invented. Many robust estimators have been developed to deal with such problems and to make the best predictions for the achievement of consistent results. The aim of this study is to make more consistent and unbiased parameter estimates with creating data sets containing influential points in different percentage amounts and in different sample sizes. In accordance with this aim, simulation studies are done for the chosen estimators (weighted maximum likelihood estimator (WMLE), conditionally unbiased bounded-influence estimator (CUBIF), consistent misclassification estimator, Mallows type estimator, Huber type estimator, Bianco-Yohai estimator (BY), weighted Bianco-Yohai estimator (WBY)) are compared appropriately for GLMs with binary response. As a result of this study, it is seen that some of the proposed robust estimators give better performance than others. Simulation studies and data analysis have been implemented in R-Project (<http://www.r-project.org/>) package programme.

Key Words: *Generalized Linear Models, Binary Response, Outliers, High Leverage Points, Robust Estimators*

Robust Estimators in Portfolio Optimization

Baki Unal

Hacettepe University
Department of Economics
bakiunal@gmail.com

Cagdas Hakan Aladag

Hacettepe University
Department of Statistics
chaladag@gmail.com

ABSTRACT

A major step in the direction of the quantitative management of portfolios was made by Harry Markowitz in 1952. The ideas introduced by Markowitz build the foundations of what is now popularly referred to as mean-variance analysis, mean-variance optimization and Modern Portfolio Theory (MPT). Mean-variance analysis provides a framework to construct and select portfolios, based on the expected performance of the investments and the risk appetite of the investor. It is the process of weighing risk against expected return. Markowitz argued that risk can be measured by the variance of returns. He suggested an optimization problem which minimizes the portfolio risk (variance) for a given level of portfolio expected return. Main inputs of this optimization problem are expected returns and variance-covariance matrix of the assets. Exact values of these inputs are not known so they must be estimated using statistical techniques. Mean-variance optimization framework is very sensitive to the accuracy of inputs and outliers in the data can lead to suboptimal results. Robust statistical techniques offer estimators that are less sensitive to outliers and other sampling errors. In this study, performances of classical and robust estimators of the optimization inputs will be compared for Turkish financial data by using back-testing technique. For robust estimators, class of M- and MM estimators, estimators based on robust scaling, Stahel–Donoho estimator and Orthogonalized Gnanadesikan–Kettenring estimator are used.

Key Words: *Mean-Variance Optimization, Robust Statistics, Portfolio Optimization, Back-testing*

Searching the Differences Through the Tails of Distributions

Gözde Navruz

Dokuz Eylül University, İzmir
gnavruz@gmail.com

A. Fırat Özdemir

Dokuz Eylül University, İzmir
firat.ozdemir@deu.edu.tr

ABSTRACT

One of the main objectives in applied statistics is to determine whether two independent groups differ, and if so, to understand how they do. Student's t test is the conventional method for doing this, however it depends on some restrictive assumptions. So, attention might be focused on some more robust approaches. There are robust alternatives for comparing two independent groups, but it is often of interest to analyze the differences which occur in the tails of distributions, namely the quantiles. When comparing two independent groups by a classical method, the test may state that there is no significant difference between populations. Conversely, comparing the groups through quantiles may indicate there are significant differences in the tails of distributions. From this point of view, two independent groups are compared by using a newly proposed quantile estimator in conjunction with a percentile bootstrap method. Results are investigated for both theoretical distributions and real data sets.

Key Words: *Quantile Estimators, Two Independent Groups, NO Estimator, Percentile Bootstrap*

Shiny App to Real Time Tweet Sentiment Classification Using Deep Learning

Andrés Felipe Rodríguez Pérez

Student of Statistics

Universidad Santo Tomás

andres.rodriquezp@usantotomas.edu.co

Wilmer Pineda Rios

PhD(c) in Statistics

Universidad Santo Tomás

wilmerpineda@usantotomas.edu.co

ABSTRACT

Digital transformation age and growth of social networks in recent years has changed the way that people get communicated. Text data generated by millions users on Twitter, about certain topics, contains unstructured information useful to identify the polarity or feeling that is being generated around a specific subject, brand or person. The process to achieve an accurate classification of the tweet polarity implies the use of Data Mining techniques and implementation of Deep Learning Models. In this talk, we describe the complete process of unstructured data processing extracted from Twitter about a specific topic, the conversion into useful entries of a classification model, the selection and calibration of model that achieves the best success rates through different Deep Learning Techniques such as Support Vector Machines (SVM), Neural Networks. Additionally, we show the visualization of this information derived from the textual data through a R Studio Shiny App focused on decision making.

Key Words: *Deep Learning, Social Networks, Twitter, Sentiment Analysis*

Simulated Annealing Based Simulation Optimization Method for Resource Constrained Project Scheduling Problem with Multiple Modes

Nurhayat Sahinkaya

Dokuz Eylül University
The Graduate School of Natural
and Applied Sciences
nurhayatsahinkaya@gmail.com

Gokalp Yildiz

Dokuz Eylül University
Dept. of Industrial Engineering,
35390, Buca-İZMİR/TURKEY
gokalp.yildiz@deu.edu.tr

ABSTRACT

This study aims at developing Simulated Annealing (SA) Based Simulation Optimization Method for Resource Constrained Project Scheduling Problem (RCPSP) with Multiple Execution Modes. In the literature, RCPSPs are generally solved by using heuristics, metaheuristics and mathematical programming methods. In real life, the duration of activities and the availability of resources throughout the project may be stochastic. Therefore, some of the inputs may be represented as stochastic variables. In order to manage these variabilities more easily, a generic simulation model is developed by using ARENA Simulation Software and it is integrated with the simulated annealing based simulation optimization method which is coded in MATLAB. In the proposed method, firstly, SA algorithm finds a candidate solution which consists of activity priority list and execution mode number of each activity, and this candidate solution is passed to the simulation model as input. Then, secondly, the simulation model is executed for estimating the makespan of this candidate solution and it passes the estimated makespan to the SA algorithm as fitness value. This process continues until termination criterion is met. The computational studies carried out show that this method can be used for solving RCPSPs and it easily handles all types of stochastic variations.

Key Words: *Resource Constrained Project Scheduling, Simulation Optimization, Simulated Annealing*

Site Effect Classification Estimated from H/V Spectral Ratio of Single Station Microtremor Data in Kürtün-Özkürtün Districts of Gümüşhane, Turkey

Serkan Öztürk

Gümüşhane University, Department
of Geophysics
serkanozturk@gumushane.edu.tr

Yasemin Beker

Gümüşhane University,
Department of Geophysics
bekaryasemin@gmail.com

Levent Pehlivan

Gümüşhane University, Department of
Geophysics
levent_pehlivan@gumushane.edu.tr

Mahmut Sarı

Gümüşhane University,
Department of Geophysics
msari@gumushane.edu.tr

ABSTRACT

Geophysical and seismological methods have been used to define site effects since measurement of geotechnical variables using borehole samples is expensive, invasive, time-consuming. However, a single station microtremor horizontal to vertical (H/V) spectral ratio technique has gained popularity to evaluate the dominant frequency or period of soil sites and it is very useful tool for engineers to measure the intensity of earthquake ground motion and the capacity of buildings to resist the earthquakes. This technique requires observations of ground motion owing to seismic ambient noise at a site and a relatively simple processing. It also requires data acquisition only in a short time and includes minimal computational time and resources. Microtremor data is preferred in the recognition of the soil layers, prediction of shear-wave velocity of the ground. Moreover, observations and analysis of microtremor data is a very effective and low-cost technique for earthquake hazard microzonation. The aim of this study is to make a site effect classification using the ratio of the horizontal to vertical components of Fourier amplitude spectra, described as H/V spectral ratio of Nakamura method in Kürtün-Özkürtün districts of Gümüşhane city, Turkey, based on the single station microtremor data measurements. For this purpose, predominant frequency, H/V spectral ratio and predominant period were determined and site effect classifications were made for 24 measurement points according to predominant periods. The locations of the observation points and the distance between these points were chosen as close to the settlement areas, taking into account the size and layout of the survey area. Recording time is determined by considering the noise content and it generally changes between 10 and 30 minutes. H/V spectral ratio changes between 1.071 and 3.942, the predominant frequency varies from 2.589 to 8.886 Hz, and the predominant period are calculated between 0.106 and 0.386. These results show that predominant periods vary in relation to the soil formation and two transient zones can be proposed as Z_1 and Z_2 between different geologic structures in Kürtün-Özkürtün region of Gümüşhane including of hard rock composed of gravel, sand and other soils mainly consisting of tertiary or older layers, hard sandy clay, sandy gravel, loam or sandy alluvial deposits whose depths are 5m or greater. Also, we observed a promising correlations between the predominant periods and surface geology. Thus, single station microtremor data can be used for seismic microzonation and disaster mitigation, even when detailed soil profile data is not available.

Key Words: *Microtremor, Kürtün-Özkürtün, Nakamura Method, H/V Ratio, Predominant Period*

Acknowledgements: This study is supported by Gümüşhane University Scientific Research Project (GUBAP, Turkey) with project no 17.F5117.02.01.

Spatial Econometric Modeling of Social Capital Effect on Internal Migration

Esra Satıcı

General Directorate of Turkish
Highways, Ankara
esra.satıcı@gmail.com

Halis Kiral

Social Sciences University of
Ankara
halis.kiral@asbu.edu.tr

Bariş Alpaslan

Social Sciences University of
Ankara
baris.alpaslan@asbu.edu.tr

ABSTRACT

Social capital has long been a question of great interest in a wide range of fields. Despite its widely varying definitions in the literature, according to Putnam (1995,p.664), social capital can be, for instance, defined as follows: “features of social life-networks, norms and trust that enable participants to act together more effectively to pursue shared objectives”. When the regional development levels cannot be explained solely by physical capital, the term social capital has started to be considered and it has then become an important concept in the economics discipline.

Given that there are several different ways in the literature to measure social capital, in our study we first determined indicators that could be used for its measurement, and then derived the social capital index at the province level in Turkey with the help of the factor analysis. Internal migration in Turkey is then modeled using the spatial econometric model in which the social capital index is considered to be an auxiliary variable, thereby allowing us to analyze not only spatial effects of social capital but its relation with internal migration.

Key Words: *Social Capital, Migration, Spatial, Factor Analysis*

Spatio Temporal Bayesian Beta Model to Estimate the Forced Displacement Ratio by Municipalities in Colombia 2002-2016

Natalia Niño Martínez
Student of Statistics
Universidad Santo Tomás
natalianino@usantotomas.edu.co

Wilmer Pineda-Ríos
PhD(c) in Statistics
Universidad Santo Tomás
wilmerpineda@usantotomas.edu.co

Dagoberto Bermúdez Rubio

PhD(c) in Statistics
Universidad Santo Tomás
dagobertobermudez@usantotomas.edu.co

Andrés Cruz Pérez

Master in Statistics
Universidad Panamericana
eacruzp@unipanamericana.edu.co

ABSTRACT

Forced displacement in Colombia has been since the nineties, the most visible consequence of the internal armed conflict in the country. Between 1985 and 2017, according to the single registry of victims, around six million people have had to travel to protect their lives and the lives of their families. Although it cannot be assumed that forced displacement is constant, it has been permanent in the last 25 years. For the specific case of this investigation, counting data expressed in proportions are studied, that is, it has a Beta distribution, making use of the generalized linear models approach with temporal space variation and the Bayesian approach. The estimation of the models will provide us with important results in the identification of the incidence of the armed conflict, presence of illicit crops and violations of human rights in the proportion of forced displacement by municipalities in Colombia, to make significant contributions to the strategies envisaged in the integral reparation stages contemplated in the country's public policy.

Key Words: *Forced Displacement, Vulnerability, Spatial Correlation, Internal Armed Conflict, Armed Actions, Threat*

Spatio-Temporal Multivariate Imputation of Missing Air Quality Data

Sebnem Er
Department of Statistical Sciences
University of Cape Town
Sebnem.Er@uct.ac.za

ABSTRACT

The air we breathe is increasingly getting polluted by fossil-fuel emissions of carbon dioxide, particles found in the air caused by increased transport and industrial development. Exposure to air pollution has a well-recognized adverse effect on the health of individuals living in the country. Several research studies show that air pollution can have several short and long-term effects such as causing lung and cardiovascular problems amongst people, if the levels are too high. This consequently has an adverse effect on the economy from a public health point of view. As a result, air pollution has become an increasingly concerning global problem. For this reason, there is a need for accurately analysing air quality data and determining the levels of pollution.

There have been several studies investigating the indoor air pollution effects on human health in South Africa however, very little has been done with regards to the effects of outdoor air pollution. This is mainly due to the fact that malfunctions and communication errors cause usually large amounts of missing data. As the presence of missing data can have a detrimental effect on the analysis of the data set, it is impossible to accurately predict and model the data set and measure the impacts of air quality on public health risk factors. Therefore, the missing values need to be replaced with plausible imputed data values. However, air quality data, measured hourly by monitoring stations within the large geographical areas around South Africa, is a spatio-temporal data. This property makes the data unsuitable for most of the common imputation methods frequently used, as the method needs to account for the spatial and temporal effects as well as the multivariate dimension of the data such as the effects of wind speed and rain environmental factors. Therefore, it is important to find suitable methods to impute missing values to determine the pollution levels in South Africa and eventually analyse the effects on human health.

The aim of the research is to determine which of the imputation methods will be most appropriate to use for the data (air quality) to have minimal error and bias when the data is modelled.

Key Words: *Missing Value Imputation, Spatio-temporal Data, Air Quality Data*

Stacking Ensemble Approach for the Classification of Survival Status: HCC patients

Metin YANGIN

Mimar Sinan University

metin.yangin@msgsu.edu.tr

Bilge BAŞER

Mimar Sinan University

bilge.baser@msgsu.edu.tr

Ayça ÇAKMAK PEHLİVANLI

Mimar Sinan University

ayca.pehlivanli@msgsu.edu.tr

ABSTRACT

Liver cancer is the one of the most frequently diagnosed cancer and **Hepatocellular Carcinoma (HCC)** is the third leading cause of cancer deaths with over 500,000 worldwide per year. Clinicians smear the liver instead of treatment; however, it takes too much time for developing medicine to fight the illness. Therefore, some in silico diagnostic tools such as machine learning algorithms have been used for classifying survival situation by using clinical data. In this study, the clinical HCC data was used which was collected from a hospital in Portugal.

Since the data has small size with missing values and imbalanced structure, it is necessary to implement an oversampling method to build larger and more balanced data set. In this study, the oversampling approach **SMOTE** is applied for this purpose. In addition to this, several data imputation methods such as **Replace Missing Values** and **Multiple Imputation by Chained Equations** which depends on **Classification and Regression Trees** are used for missing data in the preprocessing phase. Finally, **Stacking Ensemble Approach** is applied to classify the survival situation as alive or death.

Stacking is an ensemble approach that combines several base level classifiers into the multiple classifiers. The process is generated by different learning algorithms on a single dataset. In the next, a meta level classifier is chosen from the combination of base level classifiers.

In the study, base level classifiers are run individually for classifying survival status, then the results of **classification accuracy** of base level classifiers which are under **70%** to be chosen for assembling stacking ensemble approach. As a result of examining all the combinations base level classifiers obtained, it was seen that stacking ensemble methods improve the performance of accuracy between 3% and 8% in classical approaches which are used in HCC survival status.

Key Words: *SMOTE, machine learning, Hepatocellular Carcinoma, Classification, Stacking ensemble approach.*

Statistical Investigation of the Effect of Burnout and Job Satisfaction on Academicians

Ebru ÖZGÜR GÜLER
Department of Econometrics,
Faculty of Economics and
Administrative Sciences,
Çukurova University, Adana
ozgurebru@cu.edu.tr

Dilek VEYSİKARANI
Department of Econometrics,
Faculty of Economics and
Administrative Sciences,
Munzur University, Tunceli
dilekveysikarani@munzur.edu.tr

ABSTRACT

Individuals who work directly with people or face-to-face occupational groups appear to have higher burnout levels than individuals who work in other occupational groups. Tensions arise in the working life, the role conflicts, and the inability of employees to reach their expectations results in individuals become unhappy, anxious and have a negative feeling about the institution they are working with. In this situation, the feeling of burnout seen in the individuals prevents the employees from working efficiently.

In this context, the main purpose of this paper is to examine the effect of the concept of burnout on job satisfaction, which is thought to be among the most important issues affecting working life. For this purpose, a survey is applied on academicians who are one of the professional groups whose exhaustion is frequently seen. Turkish versions of Maslach Burnout scale and Hackman and Oldham Job Satisfaction scale are used in the survey. The reliability of scales are verified with Cronbach's alpha. In addition to that, factor analysis is used to check validity and to determine sub-dimensions of burnout and job satisfaction scales. Using t and ANOVA tests, the effect of demographics on burnout and job satisfaction in the academicians are investigated. Finally, regression analysis is conducted to investigate the effect of burnout on job satisfaction. By doing so, basic factors leading to burnout are determined and suggestions are made for those belonging to this profession group in the light of findings obtained.

Key Words: *Burnout, Job Satisfaction, Factor Analysis*

Statistical Methods in Optimization of Food Materials

Alev Yüksel AYDAR
Manisa Celal Bayar
University,
Department of Food
Engineering
alevyuksel.aydar@cbu.edu.tr

ABSTRACT

Statistical methods play a vital role in science and in various fields of industry. An experiment is an application of treatments applied to experimental units and is then part of a scientific method based on the measurement of one or more responses. The process and the operation of the system should be well observed. For this reason, in order to obtain a final result, an experimenter should plan and designs experiments and analyzes the results. Response Surface Methodology (RSM) is one of the most widely used experimental designs for optimization. This is a useful method because it allows researchers to evaluate the effects of multiple factors and their interactions in one or more response variables. Recent studies widely aim high yield and quality in extraction of food materials and determining optimum conditions for this extraction process. Box-Behnken design (BBD), central composite design (CCD), cube style-central composite design (SCCD), face central composite design (FCCD) and face central design (FCD) which are main designs applied in food science has been reviewed

Key Words: *Statistical Methods, Experimental Designs, Food, Optimization*

Statistical Power in Meta-Analysis Using Correlation Coefficient as Effect Size

| | |
|--------------------------|-------------------------|
| Burçin Öner | Bülent Çelik |
| Gazi University | Gazi University |
| Science Faculty | Science Faculty |
| Statistic Department | Statistic Department |
| burcinoner6858@gmail.com | bulent06celik@gmail.com |

ABSTRACT

Meta-analysis is a method of combining the results of a large number of independent studies about a specific subject and analyzing the obtained findings by using statistical techniques. In meta-analysis, it is used a numerical index as an estimate of effect size for describing the results of each study using and then these estimates across studies are combined to obtain a summary result.

Moreover, it should be known that calculations of the power of statistical tests are important in planning research studies and in interpreting situations in which a result has not proven to be statistically significant. Although statistical power is often considered in the design of primary research studies, it is rarely considered in meta-analysis. Statistical power is important in a meta-analysis study, although few studies have examined the performance of simulated power in meta-analysis.

In this paper, calculations of statistical power for statistical tests that are used according to the kind of model in meta-analysis using correlation coefficient as effect size are conducted. It is investigated whether there is any different or not between the simulation power and analytical power. Besides calculations of power according to fixed effects model and random effects model are conducted for the tests of heterogeneity and component of variance. Thus, recommendations are provided for researchers who are interested in power of meta-analysis.

Key Words: *Meta-analysis, Correlation, Simulation Power, Analytical Power, Heterogeneity*

Statistical Shape Analysis of Hand and Wrist in Children

| | | | | |
|--|--|---|---|---|
| Ural Koc Mengücek Gazi Training Research Hospital Department t of Radiology Erzincan Turkey dr_uralkoc@ hotmail.com | Ilker Ercan Uludag University, Faculty of Medicine, Department of Biostatistics, Bursa Turkey iercan@ msn.com | Semih Bolu Adıyaman Training Research Hospital Department of Pediatric Endocrinology, Adıyaman Turkey drsmhbl@ hotmail.com | <u>Ayşegül Yabacı</u> Bezmialem Vakıf University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, İstanbul Turkey Uludag University, Institute of Health Sciences, Department of Biostatistics, Bursa Turkey ayabaci@ bezmialem.edu.tr | Onur Taydaş Mengücek Gazi Training Research Hospital Depart ment of Radiology Erzincan Turkey taydasonur@ gmail.com |
|--|--|---|---|---|

ABSTRACT

Geometric morphometrics is a landmark-based method that was developed to analyze form, and thus morphological changes. Landmark-based measurement, based on size and shape information, has a long history in anthropology and other fields. Landmark location analysis has been used in forensics, computer assisted neurosurgery, anthropological studies, and MRI-based morphological analyses of the brain. This methodology offers a new and promising direction for morphological analysis by permitting the multivariate and integrated study of morphological configurations, instead of linear measurements. It reveals shape differences between different human populations.

In our study, 20 anatomical landmarks were identified on hand and wrist. Landmark coordinate data were collected from two-dimensional hand and wrist radiography of 452 children. For landmark reliability, the intra-rater reliability coefficient was computed via two-facet crossed based on the generalizability theory.

Cluster analysis was used to classify the hand and wrist. The Procrustes distance was used to compute similarity between individuals. The complete linkage method was used for merging clusters. The shape deformations were evaluated using Thin Plate Spline (TPS) analysis. To obtain the overall measurements of shape variability, the root mean square of Kendall's Riemannian distance to the mean shape was calculated.

In our study, which investigates the similarities between the hands in general and landmark-based terms, the subjects could be split into three groups according to their similarity in terms of hand shapes. Three clusters were determined at the dissimilarity level of 0.231. The first comprises 36.50% (165/452) and the second cluster comprises 9.73% (44/452) and the third cluster comprises 53.76% (243/452) of the total subjects. The general shape variability of the hands were found in cluster-I, 0.076; cluster-II, 0.077 and cluster-III, 0.082. Three clusters were evaluated according to bone age and gender.

Key Words: *Statistical Shape Analysis, Hand and Wrist Shape, Morphometry*

Stock Certificate Selection for Optimal Investment by Using Path Analysis and Multi Criteria Decision Making Methods

Özlem GÖCEN

Hacettepe University,
Department of Statistics
06800 BEYTEPE ANKARA TURKEY
ozlem.gocen@hacettepe.edu.tr

İrem KARABEBE

Hacettepe University,
Department of Statistics,
06800 BEYTEPE ANKARA TURKEY
irem.karabebe@hacettepe.edu.tr

ABSTRACT

Decision-making is choosing the most appropriate one from various activities in order to achieve an aim, taking into account the possibilities and circumstances available. Good decision is based on logical, quantitative approach, taking into consideration alternative options, using all available data, and deciding on the appeal. In this study, the optimal stocks for investment were investigated by using Multi Criteria Decision Making Methods (MCDM) from stocks of date 29.12.2017 which were traded in ISE(BİST).

The most frequently variables used in stock certificate selection are the Commodity Channel Index (CCI), Relative Strength Index (RSI), Stochastic Index (STOCH), Momentum Index, Price / Earnings Rate, Market Value / Variables are explained in the first part of the study, taking into consideration ten variables: Book Value (PD / DD), Chande Momentum Oscillator (CMO), Moving Average Convergence Divergence (MACD), Stochastic RSI and William% R. In addition to the direct relationship between the dependent variables and the independent variables, the RSI, STOCH, MACD, and William%R were selected as the independent variable for the above-mentioned ten variables, as well as the path analysis used instead of the Regression and Correlation Analysis. In the second part of the study, the TOPSIS, the Analytical Hierarchy Process (AHP), the Gray Relationship Analysis (GRA) and the ELECTRE I methods, which will enable the decision maker to obtain the decision result as quickly and as easily as possible, Path Analysis.

In the first part of the prepared study, the general information about the variables is given. In the second part, general information about Path Analysis and then application of path analysis is done. In the third part, general information about MCDM methods and an application on stocks traded in BİST are given. The analyzes were done using R and SPSS packet programs and MS Excel.

Key Words: *Stock Certificates , Multi-Criteria Decision Making, Path Analysis*

Study and modeling of the extreme rainfalls (case: Khemis Miliana, Algeria)

GUERMAH Toufik RASSOUL Abdelaziz
Université Hassiba Ben Bouali, Chlef, Algeria.
LRDSI Laboratory, Blida University, Algeria
National High School of Hydraulics,
Blida, Algeria
GEE Laboratory
tewfik.guermah@gmail.com
a.rassoul@ensh.dz

ABSTRACT

The main topic of this work is the statistical analysis of extreme values (EVA) with applications to hydrology, more specifically, to rainfalls. Statistical inference of rainfall is very important as we consider the risk of damage to agriculture, ecology, infrastructure systems and also risk of drought. The main aim of this study is to find out the most adequate fitting distributions of rainfalls taken in Khemis-Miliana station (Algeria) during the period 1975-2006. The methods of Block Maxima (BM) and the Peak Over Threshold (POT) are adopted to fit the data, after testing of course stationarity of time series in hand. Concerning estimation of parameters, we use: Maximum Likelihood Estimation (MLE), Probability Weighted Moments (PWM) and Profile of Maximum Likelihood (PMLE) for both models (GEV and GP). With these models, we derive estimates of T-years return levels for different periods T.

Key Words: GEV, GPD, POT, Block Maxima, extreme values, extremes quantile, return period, return level.

Temperature interpolation using spatio-temporal Kriging for incomplete data

Daniela López Sarmiento
MSc(c) in Applied Statistics
Universidad Santo Tomás

E-mail:
daniela.lopez@usantotomas.edu.co

Wilmer Pineda-Ríos
PhD(c) in Statistics
Universidad Santo Tomás

E-mail:
wilmerpineda@usantotomas.edu.co

ABSTRACT

The analysis of spatial and temporal correlations is useful to know the character and causes of the variability, but it is also important to predict in points of neighboring observations, or even in the same place in other moments of time. In this work, we present a method of prediction for incomplete spatio-temporal data. Various spatio-temporal covariance models have been implemented. As an application, we implement spatio-temporal interpolations using these models. The main variable is the daily mean temperature across England in 2017.

Key Words: *Interpolation, Spatio-temporal Models, Ordinary Kriging*

The Effect of Biodiesel Market and Policies in European Union and Turkey on Economic Growth

Fatih ÇEMREK

Associate Professor,
Eskişehir Osmangazi University
Faculty of Science and Letters,
Dept. Of Statistics
Meşelik Campus F1 Blok 26480
Odunpazarı/Eskişehir
fcmrek@ogu.edu.tr

Hüseyin Naci BAYRAÇ

Assistant Professor,
Eskişehir Osmangazi University
Faculty of Economics and
Administrative Sciences,
Dept. Of Economics,
Meşelik Campus F1 Blok 26480
nbayrac@ogu.edu.tr

ABSTRACT

Along with the fluctuations in fossil fuel prices in global economy, the gradual decrease in such fuel reserves has been the major factor to search for alternative energy resources. The production of biodiesel which is one of the modern bio fuels has been soaring up. Biodiesel, also named Biogasoline, Diesel-Bi, Green Diesel, is a product used as fuel which is produced through appropriate chemical / biochemical processes of oils derived from vegetable such as canola (rape), sunflower, soy, cotton, jathropa, palm and safflower or animal fats.

Biodiesel produced from vegetable oils and animal fats is different from petroleum based diesel in chemical structure; however, it is equivalent to petrol in function. It is possible to use biodiesel as pure or after blending it with petroleum based diesel in any kind of diesel engine, without any modifications on the engine or with some small changes.

Biodiesel has taken an important place in energy and agriculture policies as this kind of fuel not only contributes to different issues such as providing security of energy supply, maintaining agricultural development and struggling with global warming, but it also increases the vegetable oil production and enables recycling of vegetable oil wastes. Most countries have been paying attention to such implementations which are also named renewable energy plants agriculture. Various incentive and support policies are applied for oily seeds used in the production of biodiesel, and a variety of mixture ratios are determined and obligated to be utilized.

The aim of this study is to examine the structure of biodiesel markets in EU and Turkey and the relationship between fuel production and economic growth with panel data analysis, and to evaluate the effects of applied policies.

Key Words: *Biodiesel, European Union, Turkey, Panel Data Analysis*

* This work was supported by Eskişehir Osmangazi University Scientific Research Projects (Project No: 2017-1843

The effect of Methods Used for Missing Data Imputation on Classification Success

Hüseyin Budak
Tanı Pazarlama ve İletişim
Hizmetleri A.Ş., İstanbul, Turkey
huseyin.budak@hotmail.com.tr

Enis Gümüştas
Mimar Sinan FA Universtiy,
Statistics Dept., Bomonti,
İstanbul, Turkey
20172107001@ogr.msgsu.edu.tr

ABSTRACT

The missing data problem is frequently encountered in many fields, such as field research, clinical study results, business data. Missing data can cause various problems such as increasing errors in statistical estimates, making false predictions, loss of information due to missing data rate and inadequacy of work. Many methods have been developed based on deletion and imputation approaches as a basis for solving missing data problems. The methods that can be used for different data sets also vary depending on factors such as sample size, missing data amount and missing data mechanism.

Data sets that are used to solve real-life problems are often not fed from a single source. In data sets created by combining from many different sources, there is a loss of data due to mistakes in the process of data collection, merging or the like. Before starting statistical analysis, it is necessary to analyze the missing data and perform the necessary data cleaning operations and thus improve the quality of the data.

In this study, it was aimed to compare the methods used in the imputation of missing data over the success of classification. In addition to descriptive and predictive methods based on basic statistics for comparison purposes, machine learning methods applied in the imputation of missing data were also used. The problem of predicting credit default risk is preferred for the application. It is aimed to predict the loan repayment status of bank customers with the data obtained from a bank and with the missing observations in different variables. Before the application, Little's MCAR test was applied in order to detect the missing data mechanism in the data set. In application, different data sets were created by filling in the missing values in the data set related to each imputation method. Afterwards, classification success rates were obtained by using these data sets and LightGBM package which is a gradient boosting algorithm and the related imputation methods were compared using these rates.

Key Words: *Missing Value, Imputation Methods, Machine Learning Algorithms, Classification*

The Effects of Ranking Error Models on Mean Estimators Based on Ranked Set Sampling

Sami AKDENİZ

Department of Statistics;
Dokuz Eylül University;
İzmir, Turkey
samiakdeniz19@gmail.com

Tuğba Özkal YILDIZ

Department of Statistics;
Dokuz Eylül University;
İzmir, Turkey
tugba.ozkal@deu.edu.tr

Yusuf Can SEVİL

Department of Statistics;
Dokuz Eylül University;
İzmir, Turkey
yusuf.sevil@ogr.deu.edu.tr

ABSTRACT

Ranked set sampling (RSS) is a cost efficient sampling technique when actually measuring sampling units is difficult but ranking them is relatively easy. RSS was firstly used by McIntyre (1952) to estimate pasture yield of a cultivated area in Australia. Detailed mathematical development of RSS was presented independently by Takahasi and Wakimoto (1968). Dell and Clutter (1972) showed that sample mean based on RSS is an unbiased estimator of the population mean. In this study, we examine mean estimators that were proposed by McIntyre (1952), Samawi et al. (1996) and Muttalak (2003) for RSS, extreme RSS (ERSS) and percentile RSS (PRSS). In each of the sampling methods, two ranking models which are visual judgement ranking and concomitant variable ranking are used. Also, additive error model for visual judgement ranking (Dell and Clutter, 1972) and for concomitant variable ranking (Stokes, 1977) is studied. In simulation study, mean squared errors (MSE) of the mean estimators based on RSS, ERSS and PRSS are computed and compared with MSE of mean estimator based on simple random sampling.

Key Words: *Ranked Set Sampling, Ranking Error Models, Concomitant Variable Ranking, Visual Judgement Ranking, Mean Squared Error*

The Impact of Logistics on the Economic Series in Turkey

Berhan ÇOBAN
berhancoban@windowslive.com

Şerife ÖZKAR
Balıkesir University
International Trade and Logistics
serife.ozkar@balikesir.edu.tr

ABSTRACT

Logistics is one of the tools that play an important role in the change and improvement of economic component such as manufacture sector. The development of the logistics industry can be support business flow, technology flow and will promote the development of the transportation, commerce, finance, information, tourism and other industries. Therefore, the attributes, the logistics have been a basis or support industry in many central cities in countries and it has becoming the economic growth or support of regions. At the same time, logistics industry supplies important macro contributions to economy by creating employment, and composing national income and foreign investment flow.

In this study, we investigate how the logistics variables affect to the economic growth in Turkey. The transportation infrastructure, turnover index of transportation and storage (TITS) etc. data are used for logistics industry and for economic growth data of GDP, import and export for foreign trade etc.

The presence of cointegration relationship between the variables indicates that these variables move together in the long term. Also, if there is cointegration relationship between the variables investigated short-and long-term causal relationship. In this study, conventional and recently developed tests unit root and cointegration tests are performed to the existence of cointegration between logistic industry and economic variables.

The study reveals strong evidence that there is long term relationship between logistic sector and macroeconomic series of Turkey. It is possible to state that the logistics sector development in Turkey is essential determinants of economic structure. Thus, policy implementations of authorities, which encourage investments in this area, will influence positively economic growth.

Key Words: *Cointegration, Unit Root, Logistics Industry, Economic Growth*

The Iterative Method for Multidimensional Objects Ranking

Małgorzata Markowska

Wrocław University of Economics, Poland
malgorzata.markowska@ue.wroc.pl

Danuta Strahl

Academia WSB, Poland
dstrahl@wsb.edu.pl

ABSTRACT

The first method for linear ordering of multidimensional objects, proposed by Hellwig (1968) was based on the distance to an artificial object defined by “the best” observed values, called a development pattern. The smaller the distance, the higher the rank assigned to the object. The anti-pattern method calculates distances to “the worst” objects. The presented paper discusses a new iterative procedure extending the aforementioned methods. Objects are ranked one at a time, and then, the just ranked object is eliminated from the analysed set. New formulas are proposed for the calculation of composite index. This approach make enables to avoid unintentional weighting caused by strong skewness in diagnostic variables. In the paper, the iterative procedure is compared to the classical one. New procedure is illustrated by the industrial Relative Index of Enterprise Innovation Activity in Polish provinces, based on 39 variables from the official Survey conducted by Central Statistical Office in Poland.

Key Words: *Linear Ordering, Development Pattern, Multivariate Analysis*

The Lindley- Frechet Distribution and Properties with Lifetime Data Application

Selen Çakmakyapan

İstanbul Medeniyet Üniversitesi
İstatistik Bölümü

E-mail: selen.cakmakyapan@medeniyet.edu.tr

Gamze Ozel

Haecettepe Üniversitesi
İstatistik Bölümü

E-mail: gamzeozl@hacettepe.edu.tr

ABSTRACT

The Frechet distribution is an absolutely continuous model which has wide applicability in extreme value theory. In this paper, the Lindley- Frechet (LF) distribution which is an extension of the Lindley distribution is introduced and its properties are studied. The LF distribution represents a flexible model for the lifetime data. Some properties of the proposed distribution are obtained. The maximum likelihood method is used to estimate the model parameters and an application on the lifetime data set is presented for the illustration of the usefulness of the distribution. We provide a simulation study to illustrate the performance of the maximum likelihood estimates. The flexibility of the proposed distribution is illustrated by fitting a real data set. We use some goodness-of-fit statistics to verify the adequacy of the proposed model. We prove empirically that it is appropriate for lifetime applications.

Key Words: *Probability Distribution, Lindley Distribution, Frechet Distribution*

The Portfolio Optimization based on Genetic Algorithm Approach

Vedide Rezan Uslu
The University of Ondokuz Mayıs
Faculty of Art & Science
Department of Statistics
rezzanu@omu.edu.tr

Azize Zehra Çelenli Başaran
The University of Ondokuz Mayıs
Faculty of Art & Science
Department of Statistics
azizezehracelenli@gmail.com

ABSTRACT

All of the financial assets owned by an investor constitute the portfolio. Since an investor wishes to have maximum return and minimum risk then he/she has to decide for the diversity of the portfolio. This decision is referred as the portfolio optimization problem.

In this study as a solution approach for this optimization problem Genetic Algorithm, which is one of the heuristic algorithms, has been used. The proposed approach has been applied BIST-30 Index for the period January 2017-December 2017. This time series data set is actually the closing price of the financial assets contained from BIST-30 and the data has been analyzed by two approaches, which are Genetic Algorithm and Quadratic Programming for achieving the optimum portfolio. The results are compared to reveal that the proposed approach has provided the better result. As a novelty, we have defined Sharpe Performance Ratio, which determines the portfolio performance, as the objective function of Genetic Algorithm. The portfolio has been evaluated with respect to Sharp Performance Ratio. The portfolio which has got the maximum ratio is the optimum result. We have finally achieved the expected return and risk of the optimum portfolio as well as the weights of the financial assets contained in the optimum portfolio. We have observed that the results from Genetic Algorithm provide the high return and the high performance ratio.

Key Words: *Genetic Algorithm, Portfolio optimization, Sharpe Performance Ratio*

The Proposed F Approach for Estimation of Missing Landmark

Fatma Ezgi Can
Izmir Katip Celebi University,
Faculty of Medicine, Department
of Biostatistics and Medical
Informatics, Izmir Turkey
Uludag University, Institute of
Health Sciences, Department of
Biostatistics, Bursa Turkey
canezgi@uludag.edu.tr

Ilker Ercan
Uludag University, Faculty of
Medicine, Department of
Biostatistics, Bursa Turkey
iercan@msn.com

ABSTRACT

Statistical shape analysis is a geometric analysis of statistics measured from sets of shapes that determines the features of similar shapes or of different groups comprising similar shapes. The most important input to the use of geometric information in statistical shape analysis is landmarks. Landmark is the point corresponding to the position of a particular feature on the object of interest in two- or three-dimensional space. Missing data in landmarks in shape analysis studies leads to loss of information. We propose an approach based on F-statistic to estimate missing landmarks if there is a missing landmark in two-dimensional (2D) space in this study.

In the proposed approach after the reference landmarks are determined and the Bookstein coordinates are transformed, iterations have been made to estimate the coordinates of the missing landmark using the Euclidean distance and circle equation. Landmark estimates were made with the proposed F approach, as a result of the iterations made. The proposed approach is compared with the methods used in the estimation of missing data such as EM algorithm, multiple regression imputation and principal component analysis.

The data are generated based on multivariate normal distribution from isotropic and non-isotropic models and different simulation scenarios are considered. Landmark counts were taken as 3, 6, 9 and sample sizes were taken as 30, 50, 100.

Different scenarios and sample sizes were considered according to simulation study results, it is seen that the proposed approach gives better results than other missing data estimation methods.

Key Words: *Missing Data, Statistical Shape Analysis, Landmark*

The Use of Joinpoint Regression Analysis in Core Healthcare Indicators in Turkey, 2003-2016.

Pakize Yiğit

Istanbul Medipol University
Kavacık Mah. Beykoz,
Istanbul
pyigit@medipol.edu.tr

Seda Kumru

Istanbul Medipol University
Kavacık Mah. Beykoz,
Istanbul
skumru@medipol.edu.tr

ABSTRACT

In Turkey, Health Transformation Program, which has been introduced since 2003, aimed to improve the quality and efficiency of health system. This paper aims to identify statistically significant change in the trend of the basic healthcare indicators the Turkish healthcare system between 2003-2016 years with Joinpoint Regression Analysis. Joinpoint Regression Analysis is one of the statistical methods used to identify the best-fitting points if there is a statistically significant change in the trend. For the evaluation of health indicators were chosen the indicators including infant mortality rate, crude mortality rate, health workers per 10 000 population, health expenditure per capita, public share of current expenditure on health, hospital beds per 10 000 population, life expectancy at birth. These indicators were stated as core indicators for health sector by WHO. Infant mortality rates from 2003 to 2010 decreased by 12,8% (95% confidence interval (CI)=-15,6 to -10) per year, after 2010 there was not statistically significant decreased. Maternal mortality rates from 2003 to 2007 decreased by 23,6% (95% CI=-26,1 to -21) per year from 2007 to 2010 there was not significant decreased and from 2010 to 2016 decreased by 1,9% (95% CI=-3,6 to -0,1). Furthermore, other health indicators have improvement as a result of health transition program.

Key Words: *Joinpoint Regression, Healthcare Indicators, Trend*

Time Series Based Decision Support System to Reduce Unexpected Airline Diverts

| | | | | |
|--------------------------|---------------------|---------------------|-------------------------|---------------------|
| <u>Hazal Berve Doğan</u> | Derya Gözen | Başak Gever | Umut Demirezen | Tahir Khaniyev |
| STM Defense | STM Defense | STM Defense | STM Defense | TOBB University of |
| Technologies and | Technologies and | Technologies and | Technologies and | Economics and |
| Trade Inc., | Trade Inc., | Trade Inc., | Trade Inc., | Technology, |
| Department of | Department of | Department of | Department of | Department of |
| Technology, 06510, | Financial Affairs , | Artificial | Artificial Intelligence | Industrial |
| Ankara, Turkey | 06510, | Intelligence | Engineering and | Engineering, 06560, |
| hberve.dogan@ | Ankara, Turkey | Engineering and | Data Fusion, 06510, | Ankara, Turkey |
| stm.com.tr | derya.gozen@ | Data Fusion, 06510, | Ankara, Turkey | tahirkhanyev@ |
| | stm.com.tr | Ankara, Turkey | udemirezen@ | etu.edu.tr |
| | | basak.gever@ | gmail.com | |
| | | stm.com.tr | | |

ABSTRACT

In this study, a decision support system is designed in order to minimize the number of flights that are diverted unexpectedly. The aim is to reduce the expenses that arise when the aircraft is not able to land on the targeted airport due to the unfavorable weather conditions, such as rescheduling the timetable, overuse of aircraft fuel than planned, passengers' accommodation and ticket reissue. In order to reduce such temporal and financial losses caused by diverted flights, decision to take off or not is made before departure, while the decision to land or not is made during flight, after a brief analysis based on weather data of target airport. For the aircraft to land on target airport as scheduled, it is crucial that the weather forecasts for visibility range, ceiling and wind speed are within the limits of the safe flight requirements. Considering the significance of this decision regarding by finance, there is a need for a decision support system that is capable of boosting the process through optimal decision-making by forecasting airport weather conditions. In the first part of the study, weather is forecast using regression and time series analysis, of which methods can be detailed as auto regressive (AR), moving average (MA), auto regressive integrated moving average (ARIMA) and vector auto regressive (VAR). Although such forecast methods are relatively effective in achieving the desired result, neural network and fuzzy logic techniques are expected to present more accurate forecast with their complicated and advanced algorithm structure. Thus; in the second part of the study these methods will be used for more logical results.

Key Words: *Divert, Weather Forecast, Regression, Time Series, Decision Support System*

Time Series Forecasting of Intermittent Demand by Using ATA Method and Computational Intelligence

Tuğçe Ekiz Yılmaz
Dokuz Eylül University
0 (532) 624 72 35
ekiztugce@gmail.com

Güçkan Yapar
Dokuz Eylül University
0 (532) 436 66 57
guckan.yapar@deu.edu.tr

ABSTRACT

Intermittent demand forecasting is substantially important for firms and commercial activities. Recently, many researchers have focused on the methods which are used to forecast for intermittent demand and developed varied forecasting techniques. The outstanding method among these techniques is Croston method that is developed when inspired by exponential smoothing methods. Primarily, exponential smoothing (ES) method which is universally used more than 50 years is still one of the most prevalent forecasting techniques. Moreover, this model's superiority and success is proved by famous M-competition. However, its supremacy and widespread use in many areas, ES models have some drawbacks such as initial value that negatively affect the accuracy of forecasts. On the other hand, Croston method is widely used in forecasting of intermittent demand and inventory (stock) control. Since these demands usually include zero values, using the groundbreaking method developed by Croston becomes inevitable in this data. Nevertheless, there are some proved shortcomings such as biasness in this method. In this study, our aim is to compare the consequences of Croston and ATA method which is a new method for an alternative of exponential smoothing and the other goal is that ATA method's properties are adjusted to Croston method. Hence, ATA method will be used to intermittent demand forecasting and also proposed this method will be compared statistically to all methods.

Key Words: *Croston Method, Exponential Smoothing, Intermittent Demand, ATA Method*

Turkish Energy Sector Analysis: A Bootstrap Data Envelopment Analysis Approach

Esra Betül KINACI
Gazi University, Faculty of
Science, Department of
Statistics
esrakinaci@gazi.edu.tr

H. Hasan ÖRKÇÜ
Gazi University, Faculty of
Science, Department of
Statistics
hhorkcu@gazi.edu.tr

Hasan BAL
Gazi University, Faculty of
Science, Department of
Statistics
hasanbal@gazi.edu.tr

ABSTRACT

Data Envelopment Analysis (DEA) is a linear programming based method and is used to evaluate the relative efficiency of Decision-Making Units (DMU) responsible for output or outputs using similar inputs. DEA is one of the most preferred methods since it does not want a functional relation between input and outputs. However, estimation of the efficiency scores obtained with DEA is a bias. This is one of the disadvantages of DEA. Especially in small-sized observations, this bias will get bigger values. To solve this problem, Lönthgren and Tambour (1997) and Simar and Wilson (1997, 2000) combined the DEA method with a bootstrap method and proposed a new approach. In this way, the efficiency scores can be estimated. Later on, the efficiency scores were purged from of bias value and unbiased efficiency scores were estimated. In this study, the energy performance of Turkey in terms of consumption of energy sources was evaluated with the DEA. This analysis was applied to several types of energy sources and Turkey's energy sector analysis was evaluated for 2015 and 2016. DMUs are composed of BRICS countries (Brazil, Russia, India, China and South Korea) and Turkey. In the study, the experimental distribution of the efficiency scores obtained by the DEA was obtained and the bias of the efficiency scores were estimated through this experimental distribution by using different algorithm types of bootstrap method. Confidence intervals for unbiased estimates are calculated. The obtained results are given in charts and tables.

Key Words: *Data Envelopment Analysis, Energy Efficiency, Bootstrap*

Understanding Drought with Copula Functions: Case Study for Konya Province

O. Ozan Evkaya
Atılım University, Ankara
ozan.evkaya@atilim.edu.tr

Ceylan Yozgatlıgil
METU, Ankara
ceylan@metu.edu.tr

A. Sevtap Selcuk-Kestel
METU, Ankara
skestel@metu.edu.tr

ABSTRACT

Drought is one of the most drastic and complex natural phenomenon in the world, a major reason for undesired agricultural, economic and environmental damages. For this reason, the objective identification of drought and its characteristics deserve more detailed models. Most of the studies in the literature related to drought monitoring and analysis are based on special widespread quantitative drought indices. On the other hand, the generated drought indices have been derived using only one drought related parameter. However, the recent studies illustrated that the drought is a multi-scalar natural hazard having serious effects clearly. In this respect, the main objective of this paper is twofold. Firstly, the construction of a new copula based bivariate drought index has been considered with precipitation and potential evapotranspiration time series data. For the index generation, two weather variables are jointly considered via copula approach to establish a novel drought indicator. Additionally, most widely used drought characteristics, namely drought duration, severity and peak intensity are elaborated for the historical drought events in terms of available drought indices. The performance of the copula based drought index is examined by comparing the existed univariate drought indices like SPI and SPEI over different time scales. For this case study, monthly based weather data between 1950-2010 for weather stations in Konya province is implemented.

Key Words: *Drought, Drought Indices, Copulas, Drought Characteristics*

Use of Heijunka Technique to Reduce Production Variability by Leveling Excessive Demand Increase/Decrease

Dr. Atakan GERGER
PFW Aerospace GmbH
atakangerger@gmail.com

Asso.Prof.Dr.Ali Rıza FİRUZAN
Dokuz Eylül University
aliriza.firuzan@deu.edu.tr

Dr. Ümit KUVVETLİ
ESHOT
umit.kuvvetli@gmail.com

ABSTRACT

Companies often manufacture their products in high-quantity lots to cover the customer demand and increase production efficiency. Manufacturing the same product in large quantities may reduce setup times; however, it may cause prolonged lead times, inflated stocks, probability of exorbitant defects, excessive idle time or overtime as well. As a result, this will likely lead to profit loss for the company and stress for employees due to widely varying production plans.

In order to avoid these circumstances and secure their existence among competition, companies require new approaches. A significant one of these approaches is Heijunka, first used at Toyota as a lean manufacturing technique. Heijunka, a cornerstone of the Toyota manufacturing system, is a form of cyclic scheduling used to cover different customer demands within the same period of time and reduce production variability by leveling excessive demand increase/decrease. It does not only eliminate the waste (muda), but also the overburdening (muri) and the unevenness (mura) simultaneously. In this manner, variations in the production due to unpredictability of customer demands and planning deficiencies are minimized and statistical techniques are intensively used to ensure these results.

Key Words: *Industrial Statistics, Heijunka, Production Levelling, Lean Management, Optimization*

Using Multi-Response Taguchi Method (MRSN) for the Achievement of Product Optimization with Six Sigma Methodology in the Automotive Industry

Dr. Atakan GERGER
PFW Aerospace GmbH
atakangerger@gmail.com

Assoc. Prof. Dr. Ali Rıza FİRUZAN
Dokuz Eylül University
aliriza.firuzan@deu.edu.tr

ABSTRACT

This Six Sigma project was conducted to improve the manufacturing process and achieve product optimization, eliminating cause of a customer complaint received by the company. The outcome of this study will provide a DMAIC method for the adoption of Six Sigma in the production context, which will be used as a basis for further development of a practical Six Sigma project evaluation framework.

This implementation study suggests a conceptual Six Sigma/DOE hybrid framework with an aim to integrate Taguchi methods and Six Sigma for process improvement in a complex industrial environment. The studies employing the Taguchi experiment design usually focus on a single variable and neglect the effects of the adjustments on remaining quality characteristics, in this study a multi-response experiment design was preferred and all of the quality characteristics are taken into account.

As the result of implementing Six Sigma methodology, the sigma level increased from 2.21 (baseline) to 4.80.

Key Words: *Industrial statistics application, multi-response Taguchi design of experiment, Six Sigma, measurement system analyze, DMAIC,*

What sample sizes is needed in Validation Studies?

Arzu BAYGÜL

Beykent University

E-mail: arzubaygul@beykent.edu.tr

Neslihan GÖKMEN

Istanbul Technical University

E-mail: ngokmen@itu.edu.tr

ABSTRACT

Scales are used in health sciences increasingly. Reliability and validity are essential for their use. An important question related to their use in health sciences, is what sample sizes provide meaningful estimates of reliability and validity. Unfortunately there are not any precise sample size calculation formula about this situation. Many of publications show that five to ten sample for each item (question) is enough for validity and reliability studies.

The purpose of our study is to find the answer of this question: "What extent does sample size influence the stability of reliability and validity estimates?"

In our study, we generate sub-samples from a database which contains the answers of all questions of Asthma Self Management Questionnaire on 200 samples. We generated randomly 50 sub-samples $n=100$, 50 randomly selected sub-samples $n=50$. At all sub-samples we calculated Cronbach alpha coefficient and applied Kaiser-Meyer Olkin Test for sampling adequacy.

As a results of the study reliability estimates were stable in magnitude and interpretation in all sub-samples, whereas Kaiser Meyer Olkin test differ in lower sample sizes. These findings should be considered provisional and more work is needed to determine if these estimates are generalisable, consistent, and useful.

Key Words: *Reliability and Validation, Factor Analysis, Kaiser-Meyer Olkin Test, Sample Size, Cronbach Alpha*

Workforce Forecasting

Buse Dedeakayoğulları

Datamining Consultant at OBASE,
İstanbul, Turkey

Buse.Dedeakayogullari@obase.com

ABSTRACT

Fashion, food, cosmetics and technology companies with multiple stores in the retail sector, serve more than 7 days a week and 12 hours a day. For these companies that serve their customers in various store locations with different characteristics, customer-staff satisfaction and customers turning in to potential buyers is directly related to continuous customer service. Therefore, it is very important that the staff is assigned the right place at the right time. Workforce management becomes challenging as companies have both full-time and part-time employees, customer frequency and characteristics fluctuate and location-based workplace characteristics and product range vary. Currently, arrangements are made by store managers based on personal interpretation. This leads to incomplete or redundant staff planning. The goal of this study is to estimate the number of employees assigned to stores at certain times of the day, for companies that have multiple store locations and a certain workforce, considering customer frequency.

This study examines hourly recorded data between the dates of January 2014 and February 2015 for 390 stores of a technology retailer. There can be found 30 variables in the raw data set. The problem has numerous independent variables such as store size, sales cost, number of special days and the number of visitors. In order to estimate the number of employees assigned per store, multiple linear regression methods have been tested and models created based on the correlation results of the independent variables. In conclusion, for the workforce planning estimation, the model that gave the best results in the process was tested and the results were compared to the actual number of employees assigned per store per hour.

Key Words: *Workforce Planning, Regression, Forecasting*

