4th INTERNATIONAL RESEARCHERS, STATISTICIANS AND YOUNG STATISTICIANS CONGRESS

28-30 April 2018
Ilica Hotel Spa & Wellness Thermal Resort Çeşme, İzmir / TURKEY

BOOK OF ABSTRACTS
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>THANK YOU FROM CONGRESS CHAIR</td>
<td>22</td>
</tr>
<tr>
<td>COMMITTEES</td>
<td>23-24</td>
</tr>
<tr>
<td>SCIENTIFIC PROGRAM</td>
<td>25-30</td>
</tr>
<tr>
<td>KEYNOTE SPEAKERS</td>
<td>31-34</td>
</tr>
<tr>
<td>ORAL PRESENTATIONS</td>
<td>37-200</td>
</tr>
<tr>
<td>POSTER PRESENTATIONS</td>
<td>201-217</td>
</tr>
<tr>
<td>INDEX</td>
<td>218-221</td>
</tr>
</tbody>
</table>

*S1- A COMPARATIVE STUDY ON MODELLING CLAIM FREQUENCY IN THE PRESENCE OF ZERO-INFLATION*

Övgücan KARADAĞ ERDEMİR¹, Özge KARADAĞ

*S3- PREDICTION OF CLAIM PROBABILITY USING TREE-BASED METHODS*

Aslıhan SENTÜRK ACAR¹

*S4- COMPARISON BETWEEN MULTINOMIAL LOGISTIC REGRESSION AND ARTIFICIAL NEURAL NETWORKS: AN APPLICATION*

Deniz ALPTEKİR¹, Bülent ALPTEKİR², Çağdaş Hakan ALADAĞ¹

*S5- AN ALTERNATIVE ROBIT REGRESSION MODEL TO PROBIT REGRESSION MODEL ON BAYESIAN APPROACH*

Emre KOCABALKAN¹, Burcu MESTAV¹

*S6- THE ROLE OF BIG DATA IN INDUSTRY 4.0*

Hüseyin GÜNLÜ¹, Leyla DEMİR¹

*S7- RECENT DEVELOPMENTS IN QUALITY MONITORING OF MASS SPECTROMETRY-BASED PROTEOMIC EXPERIMENTS*

Eralp DOGU¹, Sara TAHERF, Olga VITEK²
S8- A GENERAL OVERVIEW FOR ASSESSING BIOLOGICAL TIME SHIFTS IN TIME SERIES GENE EXPRESSION DATA
Semiha ÖZGÜL1, Aslı SUNER1

S9- IS THE OUTLIER DETECTION APPROPRIATE FOR PROTEIN-PROTEIN INTERACTION DATA?
Ezgi AYYILDIZ1, Vilda PURUTÇUOĞLU1

S10- COMPARISON OF WAVELET SUBBANDS FOR EPILEPTIC SEIZURE DETECTION USING EEG SIGNALS
Hakan ÖZTÜRK1, Mevlüt TÜRE1, İmran KURT ÖMÜRLÜ1, Fulden CANTAŞ1

S11- BAYESIAN META-ANALYSIS OF PREVALENCE: ALZHEIMER'S DISEASE IN EUROPE
Esin AVCI1

S12- PERFORMANCE EVALUATION OF SUPERVISED MACHINE LEARNING ALGORITHMS FOR PREDICTING DIABETES MELLITUS
Yüksel ÖZKAN1, Banu SARAR YÜREKLİ2, Aslı SUNER1

S13- COMPARISON OF PIECEWISE AND POLYNOMIAL REGRESSION ANALYSES IN SIMULATION AND REAL DATA SETS
Buğra VAROL1, İmran KURT ÖMÜRLÜ1, Mevlüt TÜRE1

S14- DETERMINATION OF RISK FACTORS FOR ICU MORTALITY WITH SINGLE AND SERIAL BIOMARKER VALUES
Naime MERİC KONAR1, Eda KARAİSMALIOĞLU2, Ahmet Ergun KARAAGAOĞLU3

S15- COMPARING THE PERFORMANCE OF DIFFERENT STATISTICAL SOFTWARE PACKAGES FOR POWER ANALYSIS IN CLINICAL TRIALS
Gülden HAKVERDİ1, Yüksel ÖZKAN1, Aslı SUNER1

S16- ROBUSTNESS OF ANALYSIS OF COVARIANCES (ANCOVA) UNDER OF DISTRIBUTIONS ASSUMPTIONS AND VARIANCE HOMOGENEITY
Mustafa Agah TEKİNDAL1, Özlem Güllü KAYMAZ2, Can ATEŞ3, Beyza DOĞANAY ERDOĞAN4
S17- EVALUATION OF SELECTED SOME ENVIRONMENTAL FACTORS AFFECTING THE MATHEMATICS SUCCESS OF STUDENTS WITH MULTIPLE CORRESPONDENCE ANALYSIS

Mehmet Tahir HUYUT¹, Siddik KESKİN²

S18- LONGITUDINAL MODELING OF GENETIC INFLUENCES BY USING LATENT CLASS GROWTH ANALYSIS.

Özge KARADAĞ¹, Serpil AKTAŞ ALTUNAY ²

S19- EXAMINATION OF INSULIN RESISTANCE IN RHEUMATIC PATIENTS WITH LOGISTIC REGRESSION ANALYSIS

Murat GÜL¹, İlker İLHANLI², Yunus DURMAZ³

S20- RISK ADJUSTED HOSPITAL MORTALITY PREDICTION MODEL: A CASE STUDY IN A TURKISH TRAINING AND RESEARCH HOSPITAL

Fatma GÜNṬÜRKÜN¹, Özgül VUPA ÇİLENGİROĞLU¹

S21- META-ANALYSIS OF MEDIAN AND RANGE REPORTED STUDIES: DIFFERENCE BETWEEN GENDER IN TERMS OF INTERNET ADDICTION SCALE SCORES IN TURKEY

Esin AVCI¹

S22- DIFFERENTIAL NETWORK ANALYSIS ON MOUSE BRAIN PARTS USING PLSR

Aylin ALIN¹, Ayça ÖLMEZ², Gökhan KARAKÜLAH¹, Ash SUNER⁴

S23- COMPARISON OF NEURAL NETWORKS AND MULTIVARIATE NONLINEAR REGRESSION: CASE STUDY FOR ANIMAL SCIENCE

Ash AKILLI¹, Ozkan GORGULU²

S24- USE OF ARTIFICIAL INTELLIGENCE METHODS IN 305-DAY MILK YIELD ESTIMATION

Ozkan GORGULU¹, Ash AKILLI²

S25- TRANSMUTED DISTRIBUTIONS AND AN APPLICATION TO FINANCE

Nuri ÇELİK¹, Demet SEZER²
S26- RETROSPECTIVE RECONSTRUCTION OF ECONOMIC TIME SERIES

Margarita ROHR

S27- RIDIT AND EXPONENTIAL WEIGHTS FOR MULTIRATER AGREEMENT STUDIES IN LATENT CLASS MODELS

Serpil AKTAS ALTUNAY, Ayfer Ezgi YILMAZ

S28- THE LIFE SATISFACTION OF TURKEY: A COMPARISON OF MULTIPLE CORRESPONDENCE ANALYSIS AND JOINT CORRESPONDENCE ANALYSIS

Elif ÇOKER

S29- POWER ANALYSIS OF MANTEL-HAENSZEL CHI-SQUARE STATISTIC BY USING SEVERAL SCORES FOR ORDERED CONTINGENCY TABLES

Hülya OLMUŞ, Ezgi NAZMAN, Semra ERBAŞ

S30- DYNAMIC MODELING OF MEASURES OF CREDIT QUALITY

Laura VANA

S31- A COMPARATIVE EVALUATION OF THE FACTORS AFFECTING THE ELECTIVE COURSE SELECTION

Alican BOZYİĞİT, Can ATILGAN

S32- USING PARTICLE SWARM OPTIMIZATION FOR ESTIMATING THE PARAMETERS OF KAPPA DISTRIBUTION VIA MAXIMUM LIKELIHOOD METHOD: AN APPLICATION TO HYDROLOGY

Sukru ACITAS, Cagdas Hakan ALADAG, Birdal ŞENOĞLU

S33- TEXT MINING AND APPLICATIONS WITH TWITTER DATA IN R

Feten Başak DEMİR, Ayten YİĞİTER

S34- EVALUATION OF PERFORMANCES OF UNSUPERVISED ANOMALY DETECTION ALGORITHMS

Beyza KIZILKAYA, Engin YILDIZTEPE
S35- MODELING CURRENT SITUATION OF FORESTRY VILLAGES IN MURAT RIVER BASIN BY LOGISTIC REGRESSION AND TWO-STEP CLUSTER ALGORITHMS

Şenay ÇELİKÖRS1, Adnan KARABRAHİMOĞLU1, Alaattin YÜKSEL2, Alperen MERAL3

S36- NEWS ANALYZE USING TEXT MINING METHODS

Kaan TOPRAK1, Ugur SEVİK1, Tolga BERBER1

S37- A HYBRID CLASSIFICATION METHOD ON THE PREDICTION OF FOOTBALL MATCHES RESULT

Ismail Hakki KINALIOĞLU1, Coşkun KUŞ2, Ismail KINACI2

S38- ON THE PERFORMANCE OF CLUSTER ENSEMBLES FOR DOCUMENT CLUSTERING

Aytuğ ONAN1

S39- SENTIMENT ANALYSIS ON USER COMMENTS IN VIDEO SHARING NETWORKS

İsmail Hakki KINALIOĞLU1, Adem KARATAŞ2

S40- DISEASE DIAGNOSIS with SUPPORT VECTOR MACHINES: AN APPLICATION for DIAGNOSIS of DIABETES

Melda KOKOÇ1, Güvenç ARSLAN2

S41- A COMPARISON OF PAGE AND JONCKHEERE TESTS FOR REPEATED MEASURES IN RANDOMIZED BLOCK DESIGNS

Hatice Tul Kübra AKDUR1, Deniz OZONUR1, Hulya BAYRAK1

S42- LINEAR CONTRASTS IN ONE-WAY ANOVA UNDER TYPE II CENSORED SAMPLES WHEN THE DISTRIBUTION OF THE ERROR TERMS IS NON-IDENTICAL JONES AND FADDY’S SKEW T

Talha ARSLAN1, Sukru ACITAS2, Birdal ŞENOĞLU3

S43- OPTIMIZATION OF CHANNEL COVER SEAL PRODUCTION PARAMETERS USING RESPONSE SURFACE METHODOLOGY

Ali İhsan BOYACI1, Kasım BAYNAL1
S44- OPTIMIZATION OF MULTIPLE RESPONSES USING PRINCIPAL COMPONENT ANALYSIS AND RESPONSE SURFACE METHOD FOR AN EXPERIMENTAL DESIGN

Duygu KILIC1, Deniz ÖZÜN2, Hatice Tul Kübra AKDUR2, Hülya BAYRAK2

S45- MULTI-OBJECTIVE PARAMETER OPTIMIZATION OF CUSCORE CONTROL CHART FOR ARMA (1,1) COMPUTER STOCK DATA

Özgür Cem ISIK1

S46- A CRITERION TO EVALUATE THE POWER OF THE TEST

Mustafa CAVUS1, Berna YAZICI1, Ahmet SEZER1

S47- CLASSIFICATION OF AIR QUALITY MONITORING STATIONS VIA STATISTICAL METHODS: A CASE STUDY FOR KOCAELİ

Murat ÇOLAK1, Gülsen AYDIN KESKİN1, Şenay ÇETİN DOĞRUPARMAK2, Hatice ESEN1

S48- AN OVERVIEW OF E-WASTE ESTIMATION METHODS AND THEIR APPLICABILITY FOR THE CASE OF TURKEY

Zeynep OZSUT BOGAR1, Ozan CAPRAZ2, Leyla ÖZGÜR POLAT1, Askiner GUNGOR1

S49- STATISTICAL ASSESSMENT OF AIR POLLUTANT CONCENTRATION IN ALIKAHYA REGION OF KOCAELİ PROVINCE

Gülşen AYDIN KESKİN1, Murat ÇOLAK1, Şenay ÇETİN DOĞRUPARMAK2, Hatice ESEN1

S50- FORECASTING VALUE-AT-RISK WITH GENERALIZED ASYMMETRIC STUDENT-T DISTRIBUTION: APPLICATION TO S&P-500 INDEX

Emrah ALTUN1, Hüseyin TATLIDİL2

S51- A COMPARISON BETWEEN ARMA-GARCH AND KALMAN-FILTERED ARMA-GARCH IN PORTFOLIO ASSET SELECTION: CASABLANCA STOCK EXCHANGE

Abdelhamid Hamidi ALAOUI1, Fadwa El KHALIL1

S52- ON BOOTSTRAP PREDICTION INTERVALS FOR GARCH MODELS

Beste H. BEYAZTAS1
S53- INVESTIGATING DETERMINANTS OF FOREIGN DIRECT INVESTMENT WITH BAYESIAN MODEL AVERAGING

Erhan ÇENE¹, Filiz KARAMAN¹

S54- DO THE COUNTRIES’ MONETARY POLICIES HAVE SPATIAL IMPACT?

Cengiz ARIKAN¹, Yeliz YALÇIN¹

S55- ARBITRAGE STRATEGY IN ELECTRICITY MARKETS- A CASE STUDY

Betül Zehra KARAGÜL¹

S56- COMPARISON OF SIMPLE METHODS USED TO DETERMINE THE SOLAR RADIATION ON CLEAR SKY ON THE BASIS OF FUZZY LOGIC

Veysel ÇOBAN¹, Sezi ÇEVİK ONAR¹

S57- HESITANT FUZZY ENGINEERING ECONOMIC ANALYSIS OF SOLAR POWER PLANTS

Veysel ÇOBAN¹, Sezi ÇEVİK ONAR¹

S58- OUTLIER DETECTION METHODS FOR TIME SERIES DATASETS

Ekin Can ERKUŞ¹, Vilda PURUTÇUOĞLU²

S59- MAXIMUM ENTROPY DISTRIBUTIONS BASED ON FRACTIONAL CALCULUS

Mehmet Niyazi ÇANKAYA¹

S60- DYNAMIC LOT-SIZING DECISIONS UNDER FUZZY DEMAND AND THE COMPARISON WITH HEURISTIC METHODS

Metin ÖNER¹, Yağmur USTA²

S61- A LEAN PRODUCTION STUDY IN AN AUTOMOTIVE COMPANY

Hatice ESEN¹, Murat ÇOLAK¹, Gülsen AYDIN KESKİN¹
<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>S62</td>
<td>COMPARISON OF THE ITERATIVE ALGORITHMS FOR PARAMETER ESTIMATION IN NONLINEAR REGRESSION</td>
<td>Gamze MUSLUOGLU¹, Aysen DENER AKKAYA¹</td>
</tr>
<tr>
<td>S63</td>
<td>EXAMINATION ON THE PERFORMANCE OF SOME BIASING PARAMETERS FOR SHILLER’S ESTIMATOR IN DISTRIBUTED LAG MODEL</td>
<td>Nimet ÖZBAY¹, Selma TOKER¹</td>
</tr>
<tr>
<td>S67</td>
<td>INVESTIGATION OF SAMPLING BIAS IN RARE EVENT LOGISTIC REGRESSION</td>
<td>Olay ALPAY¹, Emel ÇANKAYA¹</td>
</tr>
<tr>
<td>S68</td>
<td>MODEL SELECTION IN JOINT LOCATION SCALE AND SKEWNESS MODELS OF SKEW-T DISTRIBUTION USING INFORMATION COMPLEXITY</td>
<td>Yesim GUNEY¹, Olay ARSLAN¹, Hamparsum BOZDOGAN²</td>
</tr>
<tr>
<td>S69</td>
<td>ASSESSING THE PERFORMANCE OF MODIFIED RIDGE ESTIMATOR IN SIMULTANEOUS SYSTEMS</td>
<td>Selma TOKER¹, Nimet ÖZBAY¹</td>
</tr>
<tr>
<td>S70</td>
<td>PARAMETER ESTIMATION TECHNIQUE USING PRIOR INFORMATION OF PARAMETERS IN NONLINEAR MULTICOLLINEAR DATA</td>
<td>Ali ERKOÇ¹, M. Aydın ERAR¹</td>
</tr>
<tr>
<td>S71</td>
<td>AN INTERCITY BUS SCHEDULING PROBLEM WITH CENTRAL CITY LOCATION</td>
<td>Hüseyin GÜDEN¹, Barış Keçeci², Ertan Yakıcı³, Mumtaz Karatas³</td>
</tr>
<tr>
<td>S72</td>
<td>SOLVING NONLINEAR PROGRAMMING PROBLEMS WITH PARTICLE SWARM OPTIMIZATION</td>
<td>Bülent ALPTEKİN¹, Çağdaş Hakan ALADAĞ²</td>
</tr>
<tr>
<td>S73</td>
<td>A COMBINED CUSTOMER AND SUPPLIER LOGISTICS PROBLEM FOR A COMPANY: A MATHEMATICAL MODEL</td>
<td>Mujgan SAĞIR ÖZDEMİR</td>
</tr>
<tr>
<td>S74</td>
<td>OPTIMIZATION OF MULTI-ITEM VENDOR MACHINE OPERATION</td>
<td>Ertan YAKICI¹, Hakan MERDANOĞLU², Osman Tufan DOĞAN², Serhan DURAN³</td>
</tr>
</tbody>
</table>

4th International Researchers, Statisticians And Young Statisticians Congress / 28-30 April 2018
S75- PARTICLE SWARM OPTIMIZATION FOR MULTI-OBJECTIVE OPTIMIZATION PROBLEMS
Bülent ALPTEKİN¹, Deniz ALPTEKİN², Çağdaş Hakan ALADAĞ²

S76- MULTI-OBJECTIVE MATHEMATICAL MODELS FOR SENSOR NETWORK COVERAGE PROBLEMS
Mumtaz KARATAS¹, Tuğcen HATİPOĞLU²

S77- BAGGED PI-SIGMA ARTIFICIAL NEURAL NETWORKS WITH ARTIFICIAL BEE COLONY ALGORITHM FOR FORECASTING
Erol EĞİOĞLU¹, Ufuk YOLCU², Eren BAŞ¹, Ali Zafer DALAR¹

S78- COMPARISON OF CLASSIFICATION ACHIEVEMENTS OF LOGISTIC REGRESSION AND DECISION TREES ALGORITHMS ON TIMSS 2015 DATA
Enes FILIZ¹, Tugay KARADAĞ², Oykum Esra ASKİN¹, Ersoy OZ¹

S80- RANDOM FOREST AND BOOSTING TREE ALGORITHMS IN THE PREDICTION AND DIMENSIONALITY REDUCTION MODELING OF BIOMASS ENERGY SYSTEM
Halil AKBAŞ¹, Gültekin ÖZDEMİR²

S81- LARGE SCALE PAİRWISE BİPARTİTE RANKİNG USİNG FAST GRADİENT DESCENT
Engin TAS¹

S82- ATTITUDES TOWARDS IMMIGRANTS IN GERMANY
Malika IDIZOVA¹

S83- DETERMINATION OF CUSTOMER PREFERENCES USING CONJOINT ANALYSIS: AN APPLICATION FOR CAR SELECTION
İrem SARBAT¹, Seren ÖZMEHMET TAŞAN¹

S84- MULTICRITERIA DECISION MAKING METHODS FOR ADMISSION TO AN AIR TRAFFIC CONTROL DEPARTMENT
Müjgan SAĞİR ÖZDEMİR¹, Mustafa ÖZDEMİR²
S85- INVESTIGATING CORPORATE INTEREST IN SOCIAL INNOVATION
Sercan MADANLAR¹, Burcu FELEKOĞLU²

S86- TOWARDS ON DEVELOP A NEW SOLUTION APPROACH TO MULTIPLE CRITERIA DECISION PROBLEMS
Ayşe Övgü KINAY¹, Barış Tekin TEZEL¹

S87- COMPARISON OF TYPE II REGRESSION AND STRUCTURAL EQUATION MODELING AS MEASUREMENT ERROR MODELS
Sinan SARACLI¹

S88- COMPARISON of LINK FUNCTIONS in BINARY RESPONSE MODEL by USING FIRTH’S PENALIZED MAXIMUM LIKELIHOOD ESTIMATION METHOD
Ezgi NAZMAN¹, Semra ERBAŞ², Hülya OLMUŞ²

S89- INDEPENDENT COMPONENT ANALYSIS: AN ALTERNATIVE APPROACH FOR MULTICOLLINEARITY PROBLEM
Nurbanu BURSA¹, Hüseyin TATLIDİL¹

S90- ON ESTIMATION OF KENDALL’S DISTRIBUTION FUNCTION USING BERNSTEIN POLYNOMIAL APPROACH FOR SOME ARCHIMEDEAN COPULA FUNCTIONS
Selim Orhun SUSAM¹, Burcu HUDAVERDİ UCER²

S91- ISO BASED QUALITY MANAGEMENT SYSTEM: MONITORING AND REPORTING OF QUALITY IN STATISTICS
Laura Lukšaitė BALAKAUSKIENĖ¹

S92- STATISTICAL DISCLOSURE CONTROL – BENEFITS AND CHALLENGES
Arijana Amina RAMİĆ¹

S93- GOOD QUALITY OF STATISTICAL TRAINING - GOOD QUALITY OF STATISTICAL INFORMATION
Vadym PISHCHEYKO¹, Arijana Amina RAMİĆ²
S94- HUMAN DEVELOPMENT INDEX ESTIMATION AND CLASSIFICATION BY USING ARTIFICIAL NEURAL NETWORKS AND PATH ANALYSIS
İsmail GÜRK, Hüseyin TATLIDİL

S95- SURVEY QUALITY MEASUREMENTS AT UNIVERSITIES AND STATISTICAL OFFICE IN MONTENEGRO
Gordana RADOJEVİĆ, Vuk CADJENOVİĆ

S96- ACADEMIC AND OFFICIAL STATISTICS COLLABORATION: EXPERIENCES FROM BOSNIA AND HERZEGOVINA
Fahir KANLIČ

S97- OPTIMIZATION OF MULTISTATIC SENSOR LOCATIONS FOR POINT COVERAGE PURPOSES
Mumtaz KARATAS, Emily CRAPARO, Ertan YAKICI

S98- THRESHOLD ACCEPTING ALGORITHM FOR THE CAPACITATED P-MEDIAN PROBLEM
İslam ALTIN, Aydin SİPAHİOĞLU

S99- PERFORMANCE EVALUATION OF METAHEURISTIC METHODS IN MAXIMUM LIKELIHOOD ESTIMATION FOR DISTRIBUTION PARAMETERS
Aynur SAHİN, Nimet YAPICI PEHLİVAN

S100- EMPIRICAL LIKELIHOOD ESTIMATION FOR LINEAR REGRESSION MODEL USING MM ESTIMATION METHOD
Senay ÖZDEMİR, Olcay ARSLAN

S101- A NEW GOODNESS OF FIT TEST FOR NORMALITY
Deniz ALPTEKİN, Süleyman GÜNAY

S102- STATISTICAL INFERENCE FOR THE MULTIPLE LINEAR REGRESSION MODEL: SKEW-T ERROR TERMS
İklim GEDİK BALAY, Birdal ŞENOĞLU
S103- ESTIMATION OF PARAMETERS OF SKEW-T DISTRIBUTION UNDER PROGRESSIVE CENSORING VIA MAXIMUM LIKELIHOOD USING GENETIC ALGORITHM
Abdullah YALÇINKAYA¹, Ufuk YOLCU², Birdal ŞENOĞLU¹

S104- A DISTRIBUTION FREE TEST FOR SYMMETRY COMBINING RUNS AND SIGNS
Ayman BAKLIZI¹

S105- PARAMETER ESTIMATION FOR THE MULTIPLE LINEAR REGRESSION MODEL UNDER SKEW NORMAL ERROR DISTRIBUTION: AN APPLICATION TO ENGINEERING DATA
İlçim GEDİK BALAY¹, Birdal ŞENOĞLU¹

S106- COMPARISION OF NORMALITY TEST USED COMMONLY
Murat GÜL¹

S108- SCORE TEST FOR HOMOGENEITY OF INVERSE GAUSSIAN SCALE PARAMETERS
Esra GÖKPINAR¹, Gamze GÜVEN², Fikri GÖKPINAR¹

S109- FIDUCIAL BASED APPROACH FOR TESTING THE EQUALITY OF THE TREATMENT MEANS IN ONE-WAY ANOVA WHEN THE USUAL NORMALITY AND HOMOGENEITY OF VARIANCES ASSUMPTIONS ARE NOT MET
Gamze GÜVEN¹, Özge GÜRER², Hatice ŞAMKAR¹, Birdal ŞENOĞLU²

S110- STATIONARY BOOTSTRAP BASED MULTI-STEP FORECASTS FOR UNRESTRICTED VAR MODELS
Ufuk BEYAZTAŞ¹

S111- A SIMULATION STUDY ON THE PERFORMANCE OF WEIGHTED BOOTSTRAP IN THE PRESENCE of Outliers
Uğur BİNZAT¹, Engin YILDIZTEPE¹

S112- MEASURING SYSTEMIC RISK USING COVAR-COPULA APPROACH
Tolga YAMUT¹, Burcu ÜÇER¹
S113- A NOTE ON COMPARING FIVE ROBUST REGRESSION METHODS UNDER VARIANCE HETEROGENEITY AND NONNORMALLITY

Gözde NAVRUZ1, A. Fırat ÖZDEMİR1

S114- ROBUST CONFIDENCE INTERVALS FOR THE DIFFERENCE OF TWO POPULATION VARIANCES

Hayriye Esra AKYÜZ1, Hamza GAMGAM2

S115- ROBUST VARIABLE SELECTION IN THE PRESENCE OF CELLWISE AND CASEWISE OUTLIERS

Onur TOKA1, Meral ÇETİN1, Olcay ARSLAN2

S116- RSSAMPLING: A PIONEER PACKAGE FOR RANKED SET SAMPLING WITH R

Bekir CETİNTAV1, Büşra SEVİNÇ2, Melek ESEMEN2, Selma GÜRLER3

S117- A NEW FAMILY OF EXPONENTIAL TYPE ESTIMATORS FOR THE POPULATION MEAN IN STRATIFIED RANDOM SAMPLING

Ceren ÜNAL3, Cem KADILAR1

S118- A NEW CLASS OF EXPONENTIAL TYPE OF MEDIAN ESTIMATOR USING AUXILIARY INFORMATION

Sibel AL1, Hulya CINGI1

S119- ESTIMATION OF POPULATION PROPORTION FOR BINARY VARIABLE IN MODIFIED RANKED SET SAMPLING METHODS

Aylin GÜÇOĞLU1, Neslihan DEMİREL 2

S120- ESTIMATION OF THE DISTRIBUTION FUNCTION USING PERCENTILE RANKED SET SAMPLING

Yusuf Can SEVİL1, Tuğba ÖZKAL YILDIZ1

S121- MODEL BASED INFERENCE USING JUDGMENT POST STRATIFIED SAMPLES

Konul BAYRAMOGLU KAVLAK1, Omer OZTURK2
S123- LOCAL TREND ALTERNATIVES
Sercan GÜR¹, Klaus PÖZELBERGER¹

S124- COMPARISON OF BIDDING STRATEGIES IN DOUBLE AUCTION MARKET WITH MULTI-AGENT SIMULATIONS
Baki UNAL¹, Cagdas Hakan ALADAG²

S125- TRAFFIC SIMULATION OF A CROSSROAD IN GAZIEMIR DISTRICT DURING RUSH HOURS
Sercan MADANLAR¹, Özgür YALÇINKAYA², Burcu FELEKOĞLU²

S126- COMBINING VALUE STREAM MAPPING AND SIMULATION FOR FURNITURE INDUSTRY: A CASE STUDY
Hacer GÜNER GÖREN¹

S127- A SIMULATION-BASED RESPONSE SURFACE STUDY FOR AN EMERGENCY DEPARTMENT OF A HOSPITAL
Ceren GEÇ¹, Özgür YALÇINKAYA³

S128- EVALUATION OF THE EFFECT OF QUALITY DIMENSIONS ON THE SATISFACTION LEVEL OF THE PATIENTS
Aydin TEYMOURIFAR¹, Menderes TARCAN², Onur KAYA¹, Gurkan OZTURK³

S129- GENERALIZED ESTIMATING EQUATIONS APPROACH FOR THE RISK ASSESSMENT OF FRAGILE STATE INDEX
Harun YONAR¹, Neslihan İVİT¹

S130- THE RELATION BETWEEN THE LEVEL OF SATISFACTION OF PATIENTS, AND THEIR PERCEIVED QUALITY AND LOYALTY IN A HOSPITAL IN TURKEY
Aydin TEYMOURIFAR¹, Menderes TARCAN², Onur KAYA¹, Gurkan OZTURK³

S131- EXAMINING THE PSYCHOMETRIC PROPERTIES OF THE DEPRESSION ANXIETY AND STRESS SCALE VIA A RASCH MEASUREMENT MODEL: A NON-CLINICAL EXPERIMENT
Meral YAY¹
S132- GENERALIZED LINEAR MODEL BASED CONTROL CHARTS WITH POISSON RESPONSE  
Ulduz MAMMADOVA¹, Mahmude Revan ÖZKALE ATICIOĞLU¹

S133- BI-OBJECTIVE OPTIMIZATION OF CUSCRORE CONTROL CHART PARAMETERS: A CASE STUDY FOR IMA (1,1) VISCOSITY DATA  
Ezgi Aktaş DEMİRTAŞ¹

S134- ROBUST ALTERNATIVES TO THE TUKEY’S CONTROL CHART  
Hayriye Esra AKYÜZ¹, Moustafa Omar Ahmed Abu-SHAWIESH²

S135- A RESEARCH ON ECONOMIC LITERARY LEVELS OF FACULTY OF ECONOMIC AND ADMINISTRATIVE SCIENCES STUDENTS  
Ebru ÖZGÜR GÜLER¹, Dilek VEYSİKARAN²

S136- BUFFER DESIGN FOR FAILURE PRONE PRODUCTION LINES UNDER FINANCIAL CONSIDERATIONS  
Tuğçe DABANLI¹, Seren Özmehtem TAŞAN², Derya Eren AKYOL²

S137- SHORT-TERM PROBABILITY FORECASTING BY HIDDEN MARKOV MODELS: AN APPLICATION TO A TEAM PERFORMANCE IN SOCCER  
Özgür DANISMAN¹, Umay UZUNOĞLU KOÇER¹

S138- THE EXACT DISTRIBUTION AND GEOMETRIC FUNCTION OF A GEOMETRIC PROCESS WITH ERLANG INTERARRIVAL TIMES  
Mustafa Hilmi PEKALP¹, Halil AYDOĞDU¹

S139- THE EFFECTS OF CONTRACT MECHANISMS BETWEEN PRIVATE HOSPITALS AND THE GOVERNMENT ON THE PUBLIC BENEFIT  
Aydin TEYMOURIFAR¹, Onur KAYA¹, Gurkan OZTURK¹

S140- ESTIMATION OF DURATION OF MAXIMUM DRAWDOWN IN OIL PRICES  
Merve Salçı BİLİÇİ¹, Ceren VARDAR ACAR¹, Fatma Pınar ERDEM KÜÇÜKBIÇAKÇI², İbrahim ÜNALMIŞ²
S141- PLUG-IN ESTIMATOR FOR THE MEAN VALUE FUNCTION OF TYPE II COUNTER PROCESS WITH CONSTANT LOCKING TIME
Mustafa Hilmi PEKALP1, Halil AYDOĞDU1

S142- A GENERALIZED CORRELATED RANDOM WALK APPROXIMATION TO FRACTIONAL BROWNIAN MOTION
Buket COSKUN1, Ceren VARDAR ACAR1

S145- A NEW DISTRIBUTION WITH FOUR PARAMETERS; PROPERTIES AND APPLICATIONS
Kadir KARAKAYA1, İsmail KINACI1, Coşkun KUŞ1, Yunus AKDOĞAN1

S146- A NEW GENERALIZATION OF THE LOMAX DISTRIBUTION: TWO-SIDED LOMAX DISTRIBUTION
Fatih SAHİN1, Yunus AKDOĞAN1, Kadir KARAKAYA1, Mustafa Çağatay KORKMAZ2

S147- UNIBIMODAL ASYMMETRIC DISTRIBUTIONS
Mehmet Niyazi ÇANKAYA1

S149- POWER DAGUM DISTRIBUTION: PROPERTIES AND APPLICATIONS
Merve COKBARLı1, Caner TANIŞ2, Buğra SARAÇOĞLU3

S150- ON FUNCTIONS BOUNDED BY KARAMATA FUNCTIONS
Edward OMEY1

S151- EXPONENTIED-MIXED EXPONENTIONAL DISTRIBUTION
Kadir KARAKAYA1, İsmail KINACI1, Yunus AKDOĞAN1, Coşkun KUŞ1

S152- A NEW FAMILY OF DISTRIBUTIONS: EXPONENTIAL POWER-X FAMILY OF DISTRIBUTIONS AND ITS SOME PROPERTIES
Noorsl Ahmed ZEENALABIDEN1, Buğra SARAÇOĞLU2
S153- TOPP-LEONE NORMAL DISTRIBUTION WITH APPLICATION TO INCREASING FAILURE RATE DATA

Vikas Kumar SHARMA

S154- PARAMETER ESTIMATION FOR SKEW NORMAL DISTRIBUTION UNDER HYBRID CENSORING

Nuri ÇELİK, Birdal ŞENOĞLU

S155- TESTING FOR EQUALITY OF TWO COPULAS BASED ON EXCEEDANCE STATISTICS

Ayşegül EREM

S156- GENERALIZED CUBIC TRANSMUTED FAMILY OF DISTRIBUTIONS AND ITS STATISTICAL PROPERTIES

Caner TANIS, Buğra SARAÇOĞLU

S157- A MODIFIED LINDLEY DISTRIBUTION

Coşkun KUŞ, Mustafa Çağatay KORKMAZ, İsmail KINACI

S158- COMPARISON OF THE POWER OF SOME NORMALITY TESTS: A DECISION SUPPORT SYSTEM USING SIMULATION METHOD

Metin ÖNER, Yağmur USTA

S159- CHANGE POINT DETECTION IN VEHICLE TIME HEADWAYS

Aylin ALIN, Almila HACIOĞLU, Serhan TANYEL, Ufuk BEYAZTAŞ

S160- MODE MIXING IN EMPIRICAL MODE DECOMPOSITION (EMD)

Muhammed IJAİTEH, M. Akif BAKIR, Atacan ERDİŞ

S161- COMPARISON OF CART, CART WITH BAGGING AND RANDOM FORESTS ON SEASONAL TIME SERIES FORECASTING USING SIMULATION

Mevlüt TÜRE, İmran KURT ÖMÜRLÜ, Hakan ÖZTÜRK, Fulden CANTAŞ
S162- BOOTSTRAPPED FUZZY TIME SERIES MODEL BASED ON PARTICLE SWARM OPTIMIZATION  
Ozge CAGCAG YOLCU1, Erol EGRİOGLU2, Eren BAŞ3, Ufuk YOLCU3

S163- AN INTUITIONISTIC FUZZY TIME SERIES FORECASTING MODEL BASED ON BAT ALGORITHM  
Ufuk YOLCU1, Eren BAŞ3, Erol EGRİOGLU2, Ozge CAGCAG YOLCU3

S164- INTUITIONISTIC HIGH ORDER FUZZY TIME SERIES METHOD BASED ON REGRESSION AND PRINCIPAL COMPONENT ANALYSIS  
Cem KOCAK1, Erol EGRİOGLU2, Eren BAŞ3, Ufuk YOLCU3

S165- COMPARING ETHEREUM RETURN AND USD EXCHANGE RATE RETURN WITH TIME SERIES ANALYSIS  
Batuhan ÖZKAN1, Coşkun PARİM1, Erhan ÇENE1

S166- PARAMETER ESTIMATION OF THE AUTOREGRESSIVE MODELS WITH SKEW DISTRIBUTED INNOVATIONS USING EM ALGORITHM  
Yetkin TUAC1, Olcay ARSLAN1

S167- THE USE OF DATA MINING AND NEURAL NETWORKS FOR FORECASTING PATIENT VOLUME IN AN EMERGENCY DEPARTMENT  
Ceren Öcal TASAR1, Gökem SARIYER2

S168- BOOTSTRAPPED HOLT METHOD WITH AUTOREGRESSIVE COEFFICIENTS BASED ON HARMONY SEARCH ALGORITHM  
Eren BAŞ1, Erol EGRİOGLU3, Ufuk YOLCU3

S169- Statistical Analysis in Assembly Line Balancing: Lean Manufacturing Application in Energy Sector  
Burcu Ozcan1, Kubra Mutlu1, İpek Dolu1

S170- BAYESIAN ESTIMATION FOR INVERSE GAUSSIAN DISTRIBUTION USING DIFFERENT APPROXIMATION METHODS  
İlhan USTA1, Merve AKDEDE2
S171- Supervised Term Weighting Schemes for Opinion Spamming

Aytuğ Onan

S172- Estimation of Censored Regression Model with Maximum Entropy Distributions

Ilhan USTA, Yeliz MERT KANTAR, Ismail YENILMEZ

173- Confidence Interval for the Concentration Parameter of Von Mises Distribution

NOR HAFIZAH MOSLIM, YONG ZULINA ZUBAIR, ABDUL GHAPOR HUSSIN, SITI FATIMAH HASSAN, NURKHAIRANY AMYRA MOHTAR

POSTER PRESENTATIONS

P1- COMBINING LOGISTIC REGRESSION AND RANDOM FOREST METHODS FOR THE ASSESSMENT OF ALBUMINURIA IN DIABETES MELLITUS PATIENTS

Aslı SUNER, Banu SARAR YÜREKLİ, Timur KÖSE, Mehmet Nurullah ORMAN

P2- A BAYESIAN APPROACH TO INFER THE OPTIMAL TENSION VALUE OF THE EXPONENTIAL CUBIC SPLINE TO ESTIMATE THE PROFILE GRADIENT OF A NOISY SYSTEM

Kushani De SILVA, Adom GIFFIN

P3- INVESTIGATION OF SOME INDIVIDUAL FACTORS AFFECTING THE SUCCESS OF MATHEMATICS WITH MULTIPLE CORRESPONDENCE ANALYSIS

Mehmet Tahir HUYUT, Siddik KESİN

P4- AN EMPIRICAL STUDY ON CLASSIFICATION PERFORMANCES OF ARTIFICIAL NEURAL NETWORKS, LOGISTIC REGRESSION, ANFIS, K-NEAREST NEIGHBOR ALGORITHM AND BAGGING CART

İmran KURT ÖMÜRLÜ, Fulden CANTAŞ, Mevlüt TÜRE, Hakan ÖZTÜRK

P5- POINT SYMMETRY MODELS IN SQUARE CONTINGENCY TABLES

Gökçen ALTUN, Serpil AKTAS ALTUNAY
P6- A SENTIMENT ANALYSIS ON THE SOCIAL MEDIA POSTS OF TURKISH FOOTBALL FANS.
Ismail Hakki KINALIOGLU¹, Coşkun KUŞ², Ismail KINACI²

P7- MORE THAN A DATA: MIGRATION PROBLEM AND ITS IMPACTS IN TURKEY
Ozan EVKAYA¹

P8- THE CIRCULAR MARKET FLOW AS AN APPROACH TO EXPLAIN THE VALUE OF OFFICIAL STATISTICS TO USERS
Florabela CARAUSU¹, Margarita ROHR²

P9- CALCULATION OF LOSS PROBABILITY IN THE GI/M/2/K QUEUEING MODEL WITH ORDERED ENTRY
Hanifi Okan ISGUDER¹

P10- ESTIMATION BASED ON PROGRESSIVE FIRST-Failure CENSORING FROM BURR XII DISTRIBUTION
İlhan USTA¹, Nihat ÜNAL¹

P11- INVESTIGATION OF DATA FROM VARIOUS AIR QUALITY MONITORING STATIONS IN ANKARA USING QUANTILE REGRESSION ANALYSIS
Nur Efsan TIĞLİ¹, Şengül CANGÜR¹

P12- MEASUREMENT OF WORKLOAD FOR GERIATRIC PATIENTS BASED ON TRIAGE CODES IN EMERGENCY DEPARTMENTS USING SIMULATION
Asım LEBLEBİÇİ¹, Ahmet Cagdas ACARA², Ebru USTA³, Femia YALCİN⁴

P13- ON A NEW GENERALIZATION FOR INVERTED EXPONENTIAL DISTRIBUTION
Gamze ÖZEL¹, Selen ÇAKMAKYAPAN², Ceren ÜNAL¹

P14- COMPARISON OF FOUR DIFFERENT TEST STATISTICS FOR PERCENT CHANGES: A SIMULATION STUDY
Handan ANKARALI¹, Özge PASİN², Seyit ANKARALI³
P15- SAMPLING DISTRIBUTIONS OF FOUR TEST STATISTICS FOR POSTTEST - PRETEST RATIO: A SIMULATION STUDY
Handan ANKARALI¹, Ozge PASIN², Seyit ANKARALI³

P16- THE RELATIONSHIP BETWEEN METEOROLOGICAL VARIABLES AND AIR POLLUTANTS: A CAUSALITY APPROACH
Cem KADILAR¹, Gamze ÖZEL¹
Dear Participants of the 4th International Researchers, Statisticians and Young Statisticians Congress,

I am proud to announce that we have successfully completed our congress. The purpose of IRSYSC is to bring together the leading experts, researchers and graduate students sharing the latest research in theory and application of statistics. I am happy to announce that 194 participants from 17 different countries and 44 different cities in Turkey joined us for this event. Over the course of 2 and a half day program, we heard 4 keynote addresses delivered by distinguished scientists Prof. Richard De Veaux, Prof. Narayanaswamy Balakrishan, Prof. Ismihan Bayramoglu and Prof. Metin Sezgin. We had 197 oral presentations in total, separated into 8 different contributed sessions each with 6 parallel sessions, and one poster session with 25 posters, all covering a variety of topics such as stochastic processes, linear models, official statistics, biostatistics, bioinformatics, financial mathematics and machine learning among others.

The jury, composed of members of scientific committee, evaluated the extended abstracts and the presentations of the graduate students applied for the best poster and oral presentation awards. PhD student Yetkin TUAC from Ankara University Department of Statistics won the best oral presentation award and was awarded with 250$. Unfortunately, none of the candidates were qualified for the best poster award. We would like to congratulate all the contestants for their courage and we wish them all continued success in their future endeavors.

I would like to remind you that the papers presented at the congress can be submitted to high profile journals “Hacettepe Journal of Mathematics and Statistics” or “Turkish Journal of Fuzzy Systems” depending on the topic of the paper. All submissions will go under standard evaluation process by the journals. The papers accepted by the editor and the referees will be published in an issue in 2019. For detailed information please check our web site [http://www.irsysc2018.com/call-for-papers](http://www.irsysc2018.com/call-for-papers).

I would like to express my sincere thanks to the all the members of organizing and scientific committees, all keynote speakers and all participants for making this event possible. I hope you all had a great and fruitful congress both academically and socially. Looking forward to meeting you again at a future event….

Prof. Aylin ALIN
Chair
CONGRESS CHAIR
Aylin ALIN, Dokuz Eylul University, Turkey

CONGRESS VICE CHAIRS
Burcu ÜÇER, Dokuz Eylul University, Turkey
Neslihan DEMİREL, Dokuz Eylul University, Turkey
Umay UZUNOĞLU KOÇER, Dokuz Eylul University, Turkey

HONORARY COMMITTEE
Erdal ÇELİK, Rector of Dokuz Eylul University, Turkey
Banu Esra ASLANERTİK, Vice Rector of Dokuz Eylul University, Turkey
Hikmet ÇATAL, Vice Rector of Dokuz Eylul University, Turkey

ORGANIZING COMMITTEE
A. Fırat ÖZDEMİR , Dokuz Eylul University, Turkey
Almila HACIOĞLU, Dokuz Eylul University, Turkey
Bahar YALÇIN, Dokuz Eylul University, Turkey
Berhan ÇOBAN, Dokuz Eylul University, Turkey
Büşra SEVİNÇ, Dokuz Eylul University, Turkey
Engin YILDIRIZTEPE, Dokuz Eylul University, Turkey
Filiz KARADAĞ, Dokuz Eylul University, Turkey
Gözde NAVRUZ, Dokuz Eylul University, Turkey
H. Okan İŞGÜDER, Dokuz Eylul University, Turkey
Hanife SELAMLAR, Dokuz Eylul University, Turkey
İdil YAVUZ, Dokuz Eylul University, Turkey
Mahmut Sami ERDOĞAN, Dokuz Eylül University, Turkey
ÖZGÜL VUPA ÇİLENGİROĞLU, Dokuz Eylül University, Turkey
ÖZGÜR DANIŞMAN, Dokuz Eylül University, Turkey
ÖZLEM EGE ORUÇ, Dokuz Eylül University, Turkey
Pervin BAYLAN İRVEN, Dokuz Eylül University, Turkey
Sedat ÇAPAR, Dokuz Eylül University, Turkey
Selim Orhun SUSAM, Dokuz Eylül University, Turkey
Selma GÜRLER, Dokuz Eylül University, Turkey
Senem VAHAPLAR, Dokuz Eylül University, Turkey
Tuğba YILDIZ, Dokuz Eylül University, Turkey
Uğur BİNZAT, Dokuz Eylül University, Turkey

SCIENTIFIC COMMITTEE
Anthony C. ATKINSON, London School of Economics and Political Science, UK
Arijana Amina Ramic, Germany
Augustyn MARKIEWICZ, Poznań University of Life Sciences, Poland
Aydin ERAR, Mimar Sinan University, Turkey
Ayşen APAYDIN, Turkish Statistical Association, Turkey
Berna YAZICI, Anadolu University, Turkey
Birdal ŞENOĞLU, Ankara University, Turkey
Çağdaş Hakan ALADAĞ, Hacettepe University, Turkey
Cem KADILAR, Hacettepe University, Turkey
Cengiz ÇELİKOĞLU, Dokuz Eylül University, Turkey
Claudio AGOSTINELLI, University of Trento, Italy
### Scientific Program
Saturday, April 28

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-18:00</td>
<td>Registration</td>
</tr>
<tr>
<td>10:00-10:30</td>
<td>Opening Ceremony</td>
</tr>
<tr>
<td>10:30-11:15</td>
<td>Keynote Speaker Chair: Melin Sozgan <strong>The Seven Deadly Sins of Big Data -- and How to Avoid Them</strong> Richard De Veaux</td>
</tr>
</tbody>
</table>

#### 11:15 - 11:30 Coffee Break

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Room Kasara-1 Chair: Coşkun Kus</th>
<th>Room Karina Chair: A. Fırat Özdemir</th>
<th>Room Borda Chair: Aylin Akin</th>
<th>Room Patalya Chair: Berdal Senoglu</th>
<th>Room Alabanda Chair: Ayşu Öner</th>
<th>Room Elanca Chair: Burcu Uçer</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:30</td>
<td>Generalized cubic transmuted family of distributions and its statistical properties Caner Tanoğlu, Buğra Sarıoğlu</td>
<td>Robust confidence intervals for the difference of two population variances Hayriye Esra Akyüz, Hanza Gamgam</td>
<td>Optimization of multistatic sensor locations for point coverage purposes Mümduz Karatas, Emily Craparo, Ertan Yalçın</td>
<td>A comparative study on modelling claim frequency in the presence of zero-inflation Övgücan Karadağ Erdemir, Özge Karadağ</td>
<td>Hesitant fuzzy engineering economic analysis of solar power plants Veyşil Çoban, Sezi Çevik Önar</td>
<td>Evaluation of performances of unsupervised anomaly detection algorithms Beyza Kızılkaya, Engin Yıldıztepe</td>
</tr>
</tbody>
</table>

#### 12:30 - 13:30 Lunch

**Contributed Session 2**

<table>
<thead>
<tr>
<th>Room Kasara-1 Chair: Nuri Celik</th>
<th>Room Karina Chair: Racer Güner Gönül</th>
<th>Room Borda Chair: Ufuk Boyaztas</th>
<th>Room Patalya Chair: Murat Sağır Özdemir</th>
<th>Room Alabanda Chair: Veyşil Çoban</th>
<th>Room Elanca Chair: Mümduz Karadağ</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:20</td>
<td>Parameter estimation for skew normal distribution under hybrid censoring Nuri Celik, Bidal Senoğlu</td>
<td>Combining value stream mapping and simulation for furniture industry: a case study Hacer Güner Gören</td>
<td>RSSampling: A pioneer package for ranked set sampling with R Bekir Çeliknav, Bülent Sevinç, Melih Esermen, Selma Güner</td>
<td>Confidence Interval for the concentration parameter of von Mises distribution Nor Hafizhah Motlim, Yong Zuilna Zulain, Abdul Ghafor Hussin, Siti Fatimah Hessen, Nurhakimamy Amyra Mokhtar</td>
<td>Testing for equality of two copulas based on exceedance statistics Aysegül Erem</td>
</tr>
<tr>
<td>14:40</td>
<td>Supervised term weighting schemes for opinion spamming Ayşu Öner</td>
<td>A Poisson type ordering for importance sampling and a comparison with pseudorandom and delayed acceptance MCMC Jordan Franks, Matti Viholainen</td>
<td>Estimation of population proportion for binary variable in modified ranked set sampling methods Aylin Göççüoğlu, Neslihan Demiriel</td>
<td>Assessing the performance of modified ridge estimator in simultaneou s models Selma Tokar, Nimet Özbay</td>
<td>Comparison of the power of some normality tests: A decision support system using simulation method Motin Öner, Yaşar Usta</td>
</tr>
</tbody>
</table>
15:00
Random forest and boosting tree algorithms in the prediction and dimensional reduction modeling of biomass energy system
Hali Aksay, Gultakin Odemir

Estimation of the distribution function using percentile ranked set sampling
Yusuf Caa Cevi, Tugba Ozkal Yıldız

Comparison of simple methods used to determine the solar radiation on clear sky on the basis of fuzzy logic
Yevsey Cobi, Sez Çevik Onur

15:40
A new family of distributions: Exponential Power-X Family of distributions and its some properties
Noorul Ahamed, Zeenalabdien, Bulstra Saracoğlu

Survey quality measurements at universities and Statistical Office in Montenegro
 Gordana Radijevic, Vuk Cadjenovic

Robust alternatives to the Tukey’s control chart
Hayriye Ensa Akyuz, Mustafà Omar Ahmed Abu-Shawish

16:00
A modified Lindsey distribution
Coşkun Kus, Mustafa Çağdas Hakan Alabudak

Good quality of statistical testing - good quality of statistical information
Vadym Pishchukyo, Arñana Amina Ramic

Optimization of multi-item vendor machine operation
Ertan Yakici, Hakan Mardancıoğlu, Osman Tufan Doğan, Serhat Duran

16:20
On functions bounded by Karamata functions
Edward Ormay

ISO based quality management system: Monitoring and reporting of quality in statistics
Laura Lukášová, Balázskaunkeinen

Solving nonlinear programming problems with particle swarm optimization
Bilent Alptekin, Çağdas Hakan Alabudak

16:40
Exponential-mixed exponential distribution
Kadir Karayaka, Ismail Kınıç, Yimus Açıklan, Coşkun Kus

Score test for homogeneity of inverse Gaussian scale parameters
Esra Gökşen, Gamze Gönven, Piler Gökşen

An integrity bus scheduling problem with central city location
Hasan Türkkü, Serhat Berke, Alper Öztürk, Emre Kargın, Muhammet Karatas

17:00
A new log product exponential type estimator in sample surveys
Sachin Malvi, Rajesh Singh

Retrospective reconstruction of economic time series
Margareta Rohr

A combined customer and supplier logistics problem for a company: A mathematical model
Müjgan Sağır Özdemir

17:20 - 17:40
Coffee Break

17:40
Evaluation of selected some environmental factors affecting the success of mathematics with multiple correspondence analysis
Mehtem Tahir Huyut, Sabil Köksal

Comparing Ethereum return and USD exchange rate return with time series analysis
Batuhan Oktav, Coşkun Parım, Erhan Çene

A research on economic literacy levels of faculty of economic and administrative sciences students
Ebru Özgür Gölün, Dilber Yekanlar

18:00
Differential network analysis on mouse brain parts using PSIR Asilşen, Ayça Özlem, Gökhlan Karakulak, Aslı Suner

Comparison of normality test used commonly
Munir Geli

A comparison of page and Jonckheere tests for repeated measures in randomized block designs
Hatice Tuğba Akkuş, Deniz Özön, Hulya Bayrak

Using particle swarm optimization for estimating the parameters of kappa distribution via maximum likelihood method: An application to hydrology
Sokru Atil, Cagdas Hakan Alabudak, Birdal Senoglu
<table>
<thead>
<tr>
<th>Time</th>
<th>Session Name</th>
<th>Room</th>
<th>Chair</th>
<th>Speaker(s)</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>Coffee Break</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:40</td>
<td>Contributed Session 1</td>
<td>Room Kara</td>
<td>Chair: Hülya Cingi</td>
<td>Speaker: Naima Almas, Mehmud Niyazi Cankaya</td>
<td>A criterion to evaluate the power of the test Mustafa Cavus, Berna Yazar, Ahmet Sezer</td>
</tr>
<tr>
<td>10:40</td>
<td></td>
<td>Room Kasa</td>
<td>Chair: Aslı Süner</td>
<td>Speaker: Naima Almas, Mehmud Niyazi Cankaya</td>
<td>Some new functional forms of calibration estimator Naima Almas, Muhammad Harif</td>
</tr>
<tr>
<td>11:00</td>
<td></td>
<td>Room Borda</td>
<td>Chair: Mehmet Niyazi Cankaya</td>
<td>Speaker: Ali İhsan Boyacı, Kasım Bayınalı</td>
<td>Optimization of channel cover seal production parameters using response surface methodology Ali İhsan Boyacı, Kasım Bayınalı</td>
</tr>
<tr>
<td>11:00</td>
<td></td>
<td>Room Patayla</td>
<td>Chair: Erşin Ergiliği</td>
<td>Speaker: Eren Bas, Erol Ergiliği, Ufuk Yolcu</td>
<td>The use of data mining and neural networks for forecasting patient volume in an emergency department Ceren Özlü Taşar, Gökdem Sanyer</td>
</tr>
<tr>
<td>11:20</td>
<td></td>
<td>Room Anabanda</td>
<td>Chair: Elif Doğru</td>
<td>Speaker: Eren Bas, Erol Ergiliği, Ufuk Yolcu</td>
<td>On estimation of Kendall’s distribution function using Bernstein polynomial approach for some Archimedean copula functions Selim Orhun Sukan, Burcu Hudaverdi Ucer</td>
</tr>
<tr>
<td>11:20</td>
<td></td>
<td>Room Blanca</td>
<td>Chair: Nur Cevik</td>
<td>Speaker: Ilhan Usta, Yeliz Mert Kantar, Ismail Yenilmez</td>
<td>Threshold accepting algorithm for the capacitated p-median problem Islam Altın, Ayda Şipalakliğlu</td>
</tr>
<tr>
<td>11:40</td>
<td></td>
<td>Room Tüylen</td>
<td>Chair: Ceyhan Kaya</td>
<td>Speaker: Sibel Toklu, Emel Cankaya</td>
<td>An intuitivistic fuzzy time series forecasting model based on bat algorithm Ufuk Yolcu, Eren Bas, Erol Ergiliği, Ozge Cagacag Yolcu</td>
</tr>
<tr>
<td>11:40</td>
<td></td>
<td>Room Tüylen</td>
<td>Chair: Hüseyin Tali</td>
<td>Speaker: Ceren Göge, Özgür Yalçınkaya</td>
<td>The use of data mining and neural networks for forecasting patient volume in an emergency department Ceren Özlü Taşar, Gökdem Sanyer</td>
</tr>
<tr>
<td>11:40</td>
<td></td>
<td>Room Tüylen</td>
<td>Chair: Hüseyin Tali</td>
<td>Speaker: Ceren Göge, Özgür Yalçınkaya</td>
<td>The use of data mining and neural networks for forecasting patient volume in an emergency department Ceren Özlü Taşar, Gökdem Sanyer</td>
</tr>
<tr>
<td>12:00</td>
<td></td>
<td>Room Tüylen</td>
<td>Chair: Hüseyin Tali</td>
<td>Speaker: Ceren Göge, Özgür Yalçınkaya</td>
<td>The use of data mining and neural networks for forecasting patient volume in an emergency department Ceren Özlü Taşar, Gökdem Sanyer</td>
</tr>
</tbody>
</table>

12:30 - 13:30 Lunch

13:40-14:25 Keynote Speaker
Chair: Narayanaswamy Balakrishnan
Bivariate and multivariate data in reliability and survival analysis
Ismini Bayramoğlu
### Contributed Session 2

**Chair:** Ayman Baklizi

<table>
<thead>
<tr>
<th>Time</th>
<th>Room</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30-14:45</td>
<td>Room Kasara</td>
<td>Examination of insulin resistance in rheumatoid patients with logistic regression analysis Murat Gül, Ilker İlhanlı, Yunus Durmaz</td>
</tr>
<tr>
<td>14:30-14:45</td>
<td>Room Karina</td>
<td>A new goodness of fit test for normality Deniz Alptekin, Soleyman Gilay</td>
</tr>
<tr>
<td>14:45-15:00</td>
<td>Room Borda</td>
<td>The relation between the level of satisfaction of patients, and their perceived quality and loyalty in a hospital in Turkey Aydin Teymourifar, Menderes Tanrancı, Onur Kay, Gurkan Oztruk</td>
</tr>
<tr>
<td>14:45-15:00</td>
<td>Room Patarya</td>
<td>Dynamic modeling of measures of credit quality Laura Vona</td>
</tr>
<tr>
<td>15:00-15:15</td>
<td>Room Alabanda</td>
<td>Dynamic lot-sizing decisions under fuzzy demand and the comparison with heuristic methods Metin Oner, Yaşmur Usta</td>
</tr>
<tr>
<td>15:00-15:15</td>
<td>Room Bianca</td>
<td>Bayesian meta-analysis of prevalence: Alzheimer’s disease in Europe Esin Avcı</td>
</tr>
<tr>
<td>15:15-15:30</td>
<td>Room Kasara</td>
<td>A generalized bivariate copula for flood analysis in Peninsular Malaysia Iziz Fakhrudin Kamaruzaman, Wan Zawiah Wan Zin, Noraziah Mohd Anif</td>
</tr>
<tr>
<td>15:15-15:30</td>
<td>Room Karina</td>
<td>The performance of different propensity score methods for estimating the Area Under the Curve (AUC) Hadiza Galadima, Donna McCarthy</td>
</tr>
<tr>
<td>15:15-15:30</td>
<td>Room Borda</td>
<td>Examining the psychometric properties of the Depresssion Anxiety and Stress scale via a Rasch measurement model: A non-clinical experiment Meral Yay</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td>Room Patarya</td>
<td>Rigid and exponential weights for multivariate agreement studies in latent class models Serpil Aktaş Altunay, Ayfer Ezgi Yilmaz</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td>Room Alabanda</td>
<td>The life satisfaction of Turks: A comparison of multiple correspondence analysis and joint correspondence analysis Elif Coker</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Room Bianca</td>
<td>Topso-Leonie normal distribution with application to increasing failure rate data Vikas Kumar Sharma</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Room Kasara</td>
<td>Comparison of piecewise and polynomial regression analyses in simulation and real data sets Bulqas Varol, İmran Kurt Ömürlioğlu, Mevlut Türe</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Room Karina</td>
<td>Parameter estimation for the multiple linear regression model under skew normal error distribution: An application to engineering data İklim Gedik Bayal, Birdal Şenolğlu</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Room Borda</td>
<td>Disease diagnosis with support vector machines: An application for diagnosis of diabetes Melike Kılıç, Gürvenc Anslan</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Room Patarya</td>
<td>Unimodal asymmetric distributions Mehmet Niyazi Çankaya</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Room Alabanda</td>
<td>Fiducial based approach for testing the equality of the treatment means in One-way ANOVA when the usual normality and homogeneity of variances assumptions are not met Gamze Güven, Özge Gürel, Hatice Simkan, Birdal Şenolğlu</td>
</tr>
</tbody>
</table>

### Contributed Session 3

**Chair:** Arjana Armina Ramić

<table>
<thead>
<tr>
<th>Time</th>
<th>Room</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:10-16:20</td>
<td>Room Kasara</td>
<td>Comparison of wavelet subbands for epileptic seizure detection using EEG signals Hakan Öztürk, Mevlut Türe, İmran Kurt Ömürlioğlu, Fuðan Cantas</td>
</tr>
<tr>
<td>16:10-16:20</td>
<td>Room Karina</td>
<td>A study of local influence of correlated binomial data Daniela Trentin Nova, Fernanda De Bastiani, Miguel Angel Urbic-Dioza, Orietta Nicolas, Manuel Gala</td>
</tr>
<tr>
<td>16:10-16:20</td>
<td>Room Borda</td>
<td>Generalized linear model based control charts with Poisson response Ulduz Mumadzova, Mahmudu Rehan Oztalek</td>
</tr>
<tr>
<td>16:20-16:30</td>
<td>Room Patarya</td>
<td>Fiducial based approach for testing the equality of the treatment means in One-way ANOVA when the usual normality and homogeneity of variances assumptions are not met Gamze Güven, Özge Gürel, Hatice Simkan, Birdal Şenolğlu</td>
</tr>
<tr>
<td>16:20-16:30</td>
<td>Room Alabanda</td>
<td>Unimodal asymmetric distributions Mehmet Niyazi Çankaya</td>
</tr>
<tr>
<td>16:20-16:30</td>
<td>Room Bianca</td>
<td>Fiducial based approach for testing the equality of the treatment means in One-way ANOVA when the usual normality and homogeneity of variances assumptions are not met Gamze Güven, Özge Gürel, Hatice Simkan, Birdal Şenolğlu</td>
</tr>
</tbody>
</table>

### Coffee Break

16:10-16:30 Coffee Break
17:50-19:30
Poster Presentations

Gökşen Altun, Serpil Aktaş Altunay
Point symmetry models in square contingency tables

Ozan Evkaya
More than a data: Migration problem and its impacts in Turkey

İmren Kurt Ömürli, Fülden Cantas, Mevlüt Türe, Hakan Öztürk
An empirical study on classification performances of artificial neural networks, logistic regression, ANFIS, k-nearest neighbor algorithm and bagging CART

Asıl Suner, Baru Sarar Yorkeyli, Timur Köse, Mehmet Nurullah Orman
Combining logistic regression and random forest methods for the assessment of albuminuria in diabetes mellitus patients

Cem Kadir İ, Gamze Özel
The relationship between meteorological variables and air pollutants: A causality approach

Kushani De Silva, Adam Griffin
A Bayesian approach to infer the optimal tension value of the exponential cubic spline to estimate the profile gradient of a noisy system

Mehmet Tahsin Huyut, Savaş Keskin
Investigation of some individual factors affecting the success of mathematics with multiple correspondence analysis

Florabela Carausu, Margarita Rehr
The circular market flow as an approach to explain the value of official statistics to users

Handan Ankarali, Özge Pasin, Seyit Ankarali
Sampling distributions of four test statistics for posttest – pretest ratio: A simulation study

Handan Ankarali, Özge Pasin, Seyit Ankarali
Comparison of four different test statistics for percent changes: A simulation study

Gamze Özel, Selen Calmakyapay
On a new generalization for inverted exponential distribution

İsmail Hakki Kinalıoğlu, Coskun Kus, İsmail Kinaci
A sentiment analysis on the social media posts of Turkish football fans

Nur Elhan Tığlı, Sengül Çorgür
Investigation of data from various air quality monitoring stations in Ankara using quantile regression analysis

Hanife Okan Işgider
Calculation of loss probability in the GI/M/2/K queuing model with ordered entry

Asım Leblebiç, Ahmet Çağdas Acara, Ebru Usta, Farnin Yalçın
Measurement of workloadfortriagepatientsbasedonatriagedecisionsinemergencydepartmentssimulation

Gamze Özel, Selen Calmakyapay, Ceren Ünal
On a new generalization for inverted exponential distribution
### Contributed Session 1

<table>
<thead>
<tr>
<th>Time</th>
<th>Room Kairem-1</th>
<th>Room Karina</th>
<th>Room Borda</th>
<th>Room Patalaya</th>
<th>Room Alabanda</th>
<th>Room Blanca</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00</td>
<td>Traffic simulation of a crossroad in Gazıemir district during rush hours: Sercan Madanlar, Ozgur Yalcinkaya, Burcu Felekoglu</td>
<td>Large scale pairwise bipartite ranking using fast gradient descent: Engin Tas</td>
<td>A hybrid classification method on the prediction of football matches result: Ismail Hakki Kinalioğlu, Ceksin Kus, Ismail Knaci</td>
<td>Robustness of analysis of covariances (Anova) under of distributions assumptions and variance homogeneity: Mustafa Agah Tekkok, Ozlem Guiltu Kaynag, Can Alie, Beyza Doganay Erdogan</td>
<td>Plug-in estimator for the mean value function of Type II counter process with constant locking time: Mustafa Hilmi Pekalp, Hali Aydogdu</td>
<td>Stationary bootstrap based multi-step forecasts for unrestricted VAR models: Ufuk Beyaztas</td>
</tr>
</tbody>
</table>

### 12:00 - 12:20 Coffee Break

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:20</td>
<td>Closing Ceremony</td>
</tr>
</tbody>
</table>

---

**Keynote Speaker**

Chair: Richard De Vaux  
Psychology, Statistical Models and Intelligent Interaction  
Mein Sezgin  

---

**Organizational Sponsor**

VakifBank

**Contact Information**

MOTTO  
www.motto.tc  
0232 446 06 10  
info@motto.tc
KEYNOTE SPEAKERS
Cure Models

Narayanaswamy BALAKRISHNAN

McMaster University Hamilton, Canada.

With the advancement in medical technology, more and more patients, after going through a cancer treatment (or for other serious diseases, as a matter of fact), tend to live a long life, free of cancer. If one were to model the lifetimes of these individuals after the treatment, then one has to duly account for such long-term survivors. Models that are intended for this purpose are called "cure models" or "long-term survival models." Such models have become very popular in the last decade or two, and have been used in many different contexts, including biostatistics, finance, demography, reliability theory, etc.

In this talk, I will first introduce a basic mixture cure rate model as it was originally introduced in 1950s. After that, I will formulate cure rate model in the context of competing risks and present some flexible families of cure rate models. I will then describe various inferential results for these models. Next, as a two-stage model, I will present destructive cure rate models and discuss inferential methods for it. In the final part of the talk, I will discuss various other extensions and generalizations of these models and the associated inferential methods. All the models and inferential methods will be illustrated with simulation studies as well as some well-known melanoma data sets from the literature. While discussing inference, likelihood method based on EM algorithm will be focused as a natural method of fit of these ismi
Devices are getting faster, more affordable and smarter. Intelligent user interfaces that power these devices make heavy use of machine learning to enable natural human-device communication.

T. Metin Sezgin

Associate Professor Koc University College of Engineering Istanbul, Turkey.

Speech is the dominant modality in human-human communication. It is supported in subtle ways through other communicative cues (e.g., gestures, eye-gaze, and haptics). These cues, although subtle, play a major role in enriching human-human interaction by communicating complementary information. In this talk, I will present case studies that demonstrate the wide range of information that can be extracted from subtle cues, and will show examples of how human-computer interaction in general, and human-robot interaction in particular, can be enhanced with strategic use of subtle communicative cues. The examples come from multimodal interfaces using eye-gaze tracking, and pen-based interaction.
Bivariate and multivariate data in reliability and survival analysis.

İsmihan Bayramoğlu

Economy Üniversty İzmir, Turkey.

I will talk on modeling of bivariate data in survival and reliability analysis. The bivariate and multivariate data appearing in models and processes dealing with death in biological organisms and failure in mechanical systems are usually statistically dependent. In this context, the dependency properties and stochastic ordering properties of bivariate and multivariate data and distributions will be discussed. I will talk also on copula models having important applications in reliability engineering.

Some new results concerning problems of stochastic mathematics also will be presented.
The Seven Deadly Sins of Big Data -- and How to Avoid Them

Richard De Veaux

As we are all too aware, organizations accumulate vast amounts of data from a variety of sources nearly continuously. Big data advocates promise the moon and the stars as you harvest the potential of all these data. There is certainly a lot of hype. There’s no doubt that savvy organizations are fueling their strategic decision making with insights from data mining, but what are the challenges?

Much can wrong in the data mining process, even for trained professionals. In this talk I'll discuss a wide variety of case studies from a range of industries to illustrate the potential dangers and mistakes that can frustrate problem solving and discovery -- and that can unnecessarily waste resources. My goal is that by seeing some of the mistakes I have made, you will learn how to take advantage of data mining insights without committing the "Seven Deadly Sins."
ORAL PRESENTATIONS
S1- A COMPARATIVE STUDY ON MODELLING CLAIM FREQUENCY IN THE PRESENCE OF ZERO-INFLATION

Övgücan KARADAĞ ERDEMİR¹, Özge KARADAĞ²

¹Department Of Actuarial Science, Hacettepe University, Ankara, Turkey, ²Department Of Statistics, Hacettepe University, Ankara, Turkey.

In non-life insurance mathematics, there are two components of claim losses, claim frequency and claim severity. The claim frequency is the number of claims in a portfolio of insurance policies over a period of time, while claim severity shows the monetary losses of insurance claims. To model claims in non-life insurance the most used models are generalized linear models, depending on the distribution of claims. The distribution of the claim frequency is generally assumed Poisson, however insurance claim data contains zero counts which effects the total loss, calculated by both the claim frequency and the claim severity. Observations departure on the majority of the data, these zero counts, have a strong influence on standard techniques. In the presence of excess zero, there are more appropriate distributions for the claim frequency such us Negative Binomial, Zero-Inflated Poisson, Zero-Inflated Negative Binomial and Hurdle Models instead of a standard Poisson distribution. In this study, using a real annual automobile insurance data of an insurance company, different distributions are compared for claim frequency which contains zero counts.

KEYWORDS: claim frequency, generalized linear model, zero-inflation, count data models
S3- PREDICTION OF CLAIM PROBABILITY USING TREE-BASED METHODS

Aslıhan ŞENTÜRK ACAR

Hacettepe University.

Machine learning (ML) is a growing field that uses the learning algorithms to construct relationship between an outcome and predictors as an intersection of computer science and statistics. Classification trees are the basic ML methods for the classification tasks. Although classification trees have many advantages like easy visual interpretation, they have some disadvantages such as high variance. Therefore, new methods are developed to overcome the drawbacks of classification trees. Some of these methods are as bagging, random forest and boosting. Probability of claim occurrence is one of the main components for the estimation of insurance risk premium. This probability is generally estimated using classical logistic regression model that is a parametric approach. Objective of this study is to use tree-based ML methods for the prediction of claim probability and to compare the predictive performances with logistic regression using a health insurance data set from a Turkish insurance company. Classification trees, bagging, random forest and boosting methods are used as alternatives to logistic regression. Predictive performances of different methods are compared and the results are discussed.

KEYWORDS: claim probability, classification trees, machine learning, prediction, insurance
S4- COMPARISON BETWEEN MULTINOMIAL LOGISTIC REGRESSION AND ARTIFICIAL NEURAL NETWORKS: AN APPLICATION

Deniz ALPTEKİN¹, Bülent ALPTEKİN², Çağdaş Hakan ALADAĞ¹

¹Department Of Statistics, Hacettepe University, Ankara, Turkey, ²Department Of Statistics, Middle East Technical University, Ankara, Turkey.

Multinomial logistic regression and artificial neural networks have employed to explore unknown patterns and prediction in the various field. Both methods share common roots in statistical pattern recognition. However, multinomial logistic regression model complexity is low, there is always risk of overfitting when few interaction terms and variable transformations are used. This may cause low model flexibility. Compare to multinomial logistic regression, artificial neural networks models are more flexible. The aim of this study is to compare the classification performance of multinomial logistic regression and artificial neural networks for predicting kyphosis and scoliosis.

KEYWORDS: Classification, kyphosis, scoliosis
S5- AN ALTERNATIVE ROBIT REGRESSION MODEL TO PROBIT REGRESSION MODEL ON BAYESIAN APPROACH

Emre KOCABALKAN1, Burcu MESTAV1

1Canakkale Onsekiz Mart University.

In regression, if response variable comes from exponential family, then Generalized Linear Models is an alternative to weighted Least Square method. Probit models are special cases of GLMs. Probit model is a standard probabilistic statistical classification model that has been extensively used. Probit model is flexible and convenient for modeling binary data, but it can run into problems with outliers due to the light tails of normals. Recently, some robust models which obtain more consistent results have been developed to reduce deviations in parameter estimates of the dataset with outliers. The T distribution provides a useful extension of the Normal for statistical modeling of data sets involving errors with longer than Normal tails. Therefore, we use a more robust model which replaces the normal distribution for \( e_i \) with T distribution. This model is called Robit model. It is aim of this paper that the robit model is a useful robust alternative to the probit model for analyzing binary response data with Bayesian approach. The data were analyzed using the R2WinBUGS package in the R-project. According to the results of analysis of the data with Probit and Robit models; (i) the parameter estimates obtained by the classical approach are similar in the Bayesian approach, (ii) the known degree of freedom Robit model is more resistant than the Probit model; (iii) Robit model alternative to the Probit model according to the DIC values; (iv) the Robit model is the best model in the presence of contradictory observations in the data set.

KEYWORDS: Bayesian Approach, Probit Model, Robit Model, Robust Regression
Humans have recorded their knowledge, thoughts and observations to inherit their minds and souls to further generations for ages. To achieve this, writing was used as the main tool in the past. However, the improvements and innovations in the science and technology have forced them to find new efficient and rapid ways to record data such as Data Lakes and Cloud Storages. These rapid ways led an extremely enormous, various and fast updated data which is called Big Data. On the other hand, the ongoing improvements also led another phenomenon to occur: Industry 4.0. Nowadays, the journey of the industry starting with steam engines in 18th century is proceeding with Internet of Things (IoT), Artificial Intelligence (AI), and so on. It is undoubtedly clear that Industry 4.0 will change the world in a positive way. However, without the accurate data, is it possible? This paper describes the importance of the Big Data Analytics for Industry 4.0. Moreover, a Big Data Analytics framework is proposed for collection, analysis and integration of all data in this context.

**KEYWORDS:** Big Data, Industry 4.0, Data Analytics
Advances in mass spectrometry-based proteomics allow the quantitative profiling of an increasingly large number of biological samples and analytes. This introduces new challenges for fully automated, longitudinal monitoring of these experiments for system suitability and quality control (QC). On one hand, automated systems must detect different types and patterns of deviations from optimal performance. On the other hand, they must integrate these patterns across multiple metrics (such as retention time, peak area, mass accuracy etc.) of multiple analytes. Although some of these goals can be achieved with expertise, or with simple statistical summaries, this is typically insufficient. Specialized methods and software are needed to characterize the stochastic properties of the data in a way that reflects the details of sample preparation and spectral acquisition, and to implement the statistical analysis workflows in a way that is useful for both experimentalists and statisticians. MSstatsQC and its graphical user interface, MSstatsQCgui, are two R/Bioconductor packages for statistical longitudinal system suitability and quality control (QC) of mass spectrometry-based proteomics experiments. In this study, we demonstrated the importance of monitoring longitudinal system suitability and quality control performance for reproducible and reliable results by implementations with MSstatsQC.

**KEYWORDS:** proteomics, mass spectrometry, process monitoring
S8- A GENERAL OVERVIEW FOR ASSESSING BIOLOGICAL TIME SHIFTS IN TIME SERIES GENE EXPRESSION DATA

Semiha ÖZGÜL¹, Aslı SUNER¹

¹Biostatistics And Medical Informatics Department, Ege University, Izmir.

Gene expression time series are sequences of gene expression levels collected at specific time points. The gene expression time-series of biological processes such as developmental, cell cycle and dose-response may vary in sampling times or rates of progression in different units although they typically follow a typical pattern. The size, direction, and pattern of these timing differences can have many kinds of different and significant effects; therefore they are crucial and are also a common phenomenon in biology. The aim of this study was to provide a general overview of different approaches used to analysis of time shifts such as fast Fourier transform, Hidden Markov Models, Gaussian Process and alignment algorithms, and to identify potential time shifts between mouse and human fetal lung development with the dynamic time-warping algorithm, a simulation procedure. For all statistical analyzes, the R 3.2.2 program was used and the level of significance was α = 0.05. Our results showed that the genes associated with RNA metabolism and processing were expressed at significantly different rates during the lung development between mouse and human. Additionally, 430 (2.77%) of 15474 homologous genes showed significant time shifts. Of the significant genes, 158 (36.74%) were the exact match; 101 (23.49%) were positive and 72 (16.75%) were negative time shifts, but no pattern was observed in 99 (23.02%) of them. In conclusion, the methods used to estimate time shifts provide great ease in bringing out the similarities and differences in the biological events at different speeds at the genome level.

S9- IS THE OUTLIER DETECTION APPROPRIATE FOR PROTEIN-PROTEIN INTERACTION DATA?

Ezgi AYYILDIZ¹, Vilda PURUTÇUOĞLU¹

¹Department Of Statistics, Middle East Technical University, Ankara, Turkey.

In systems biology, in order to describe the activation flow of the transaction pathways, the interaction between systems’ components such as genes and proteins can be represented by a network. But, the mathematical modelling of the biological networks is challenging because of their dependent, high dimensional and sparse structures. There are a number of methods to present these complexities under distinct assumptions. Whereas the performance of these models can be dependent on the quality of the data in the sense that the data have any outliers or not. Hereby, to discard the effect of outliers and other erroneous signals, different data use distinct approaches. For instance, the microarray, Chip-seq and RNA-seq datasets, also commonly used in modelling, apply the normalization procedure to clean the batch effects and outliers. Hence, in this study, we are particularly interested in protein-protein interaction data and investigate whether an outlier detection approach is necessary as a pre-processing step before modelling so that we can improve the fitted model and also whether there is any appropriate method to be used for this purpose considering the challenges in the structure of systems. Accordingly, we use various benchmark networks and perform different outlier detection methods designed for univariate, multivariate and high dimensional cases and finally, check the accuracy of estimated systems under 3 network models, namely, GGM, MARS and CMARS. From the results, it is seen that such a pre-processing step cannot improve our accuracy and there are only a few methods applicable for sparse and high-dimensional systems.

KEYWORDS: Outlier detection, protein-protein interaction networks, multivariate adaptive regression spline, Gaussian graphical model, conic multivariate adaptive regression spline
S10- COMPARISON OF WAVELET SUBBANDS FOR EPILEPTIC SEIZURE DETECTION USING EEG SIGNALS

Hakan ÖZTÜRK¹, Mevlüt TÜRE¹, İmran KURT ÖMÜRLÜ¹, Fulden CANTAŞ⁶

¹Biostatistics, Adnan Menderes University, Aydın, Turkey.

In this study, it is aimed to compare the performances of the frequency sub-bands for epileptic seizure detection using Electroencephalogram (EEG) signals. Discrete wavelet transform (DWT) was performed to decompose the EEG signals into detail (D1-D5) and approximation (A5) sub-bands. Four statistical features such as absolute mean, standard deviation, energy and entropy of wavelet coefficients were calculated for each frequency sub-bands and used as inputs to the Multi-Layer Perceptron (MLP) model. MLP was performed with 10-fold cross validation to calculate the prediction probabilities of each sub-bands for seizure detection. The Receiver Operating Characteristic (ROC) analysis was applied to prediction probabilities for each sub-bands. Then area under the curves (AUCs), sensitivity, specificity and accuracy rates were calculated for each sub-bands. AUC was used as comparison criterion of sub-bands for seizure detection. AUCs for each sub-bands were statistically significant (p<0.001). D5 was achieved the highest AUC=0.994 and significantly higher than other sub-bands (p<0.01). The AUCs for D2, D4, D3, D1 and A5 were obtained, 0.987, 0.986, 0.981, 0.967 and 0.961, respectively. In conclusion, we got the most valuable information from 2.7 - 5.4 Hz frequency range (D5) for epileptic seizure detection.

KEYWORDS: EEG, Discrete Wavelet Transform, Multi-Layer Perceptron, ROC
S11- BAYESIAN META-ANALYSIS OF PREVALENCE: ALZHEIMER’S DISEASE IN EUROPE

Esin AVCI

1Giresun University.

Statistically, the prevalence is defined as the frequency of disease at a given time in a particular population. To estimate more precise prevalence meta-analysis can be useful method. Meta-analysis is a statistical method that combine the results of studies on a determined topic to derive an average estimate. In medical and statistical literature meta-analysis of prevalence is often considered in frequentist approach rather than Bayesian approach. Frequentist meta-analysis used two main statistical models; Fixed-effect and random-effect model. Under the fixed-effect model the effect size is assumed to be same for all studies and all differences in observed effects are due to sampling error. By contrast, under the random-effects model true effect could be change from study to study. In Bayesian meta-analysis, one has to gather the data from the selected studies, choose informative or non-informative prior, model the posterior and run simulations to asses parameters of the posterior distribution. Besides direct probability statements on different scales and predictions, the conflict between fixed- and random-effects meta-analysis are handled by Bayesian approach. To determine prevalence of Alzheimer’s Disease in Europe, frequentist and Bayesian meta-analysis applied to data obtained from Niu et al. (2017). As a result, Bayesian approach given a narrower confidence interval and smaller relative error than frequentist approach. Hence, more accurate prevalence estimates are derived from the Bayesian approach.

KEYWORDS: Meta-analysis, Bayesian approach, Prevalence, Alzheimer’s Disease
S12- PERFORMANCE EVALUATION OF SUPERVISED MACHINE LEARNING ALGORITHMS FOR PREDICTING DIABETES MELLITUS

Yüksel ÖZKAN¹, Banu SARAR YÜREKLİ², Aslı SUNER¹

¹Ege University, Faculty Of Medicine, Department Of Biostatistics And Medical Informatics, İzmir, Turkey, ²Ege University, Faculty Of Medicine, Department Of Endocrinology, İzmir, Turkey.

This study was aimed to compare the performances of machine learning methods increasingly used in health sciences for the classification of diabetes mellitus. Diabetes mellitus (healthy, pre-diabetes and diabetes) was classified with the supervised machine learning methods; including naïve Bayes, Bayes network, random forest, decision tree, support vector machine, k-nearest neighbors, artificial neural network and multinomial logistic regression. The application data was obtained from 232 patients aged ≥18 who were admitted to the endocrinology department of a public hospital from June to September 2013. Three different approaches were used for the classification. Firstly, 39 variables in the dataset were included in all models. Secondly, 18 variables which found statistically significant in the univariate logistic regression method were utilized. Lastly, all models were built with 21 clinically significant variables which were determined by the expert. The performance of the models was evaluated in terms of their accuracy, kappa statistics, mean absolute error, root mean squared error, relative absolute error, sensitivity, specificity, precision, F-measure, Matthews correlation coefficient, area under the ROC curve and Youden index. 10-folds cross-validation method was applied to test the performance of the models. Additionally, run times of each method were calculated. All analyses were performed with software packages WEKA 3.8.2 and R Studio 1.1.383. As a result, for all three approaches, random forest method had the best performance and the accuracy 82.33%, 84.48%, and 81.90%, respectively. In conclusion, when choosing the best model, it would be useful to take into account statistical significance as well as clinical significance.

KEYWORDS: Diabetes Mellitus Diagnosis, Classification, Machine Learning Algorithms, Supervised Learning
S13- COMPARISON OF PIECEWISE AND POLYNOMIAL REGRESSION ANALYSES IN SIMULATION AND REAL DATA SETS

Buğra VAROL¹, İmran KURT ÖMÜRLÜ¹, Mevlüt TÜRE¹

¹Adnan Menderes University Medical Faculty Department Of Biostatistics, Aydın, Türkiye.

Piecewise regression, which one or more pieces are combined in breakpoints, is widely used as a statistical technique. It was aimed to compare piecewise regression analyses and polynomial regression analysis using both simulated data and real data sets. In the application step of the study, algorithms were created by using R software for simulation practice. Polynomial and piecewise regression analysis methods were compared using data sets with n=100 units and 1000 times running simulation. In addition, estimation performances of piecewise and polynomial regression built by using the data sets which contained in the number of tuberculosis cases according to age in 2010 year and the number of measles cases from 1993 to 2015 years in Turkey were compared. It was found that there was a significant difference between all of the polynomial and piecewise regression models (p<0.001). The R^2 values of piecewise regression models were higher than polynomial regression models; MSE, AIC and BIC values were observed to be lower. According to the result of both simulation and real data set applications, piecewise regression models which were generated according to optimal knots were found to have better estimation performance than polynomial regression models according to R^2, MSE, AIC and BIC criteria. This study revealed that data analysis with piecewise regressions having optimal knots provided statistically superiority, although polynomial regression methods are preferred in the field of health studies mostly.

KEYWORDS: Piecewise Regression, Simulation, Polynomial, Tuberculosis, Measles, Knot
S14- DETERMINATION OF RISK FACTORS FOR ICU MORTALITY WITH SINGLE AND SERIAL BIOMARKER VALUES

Naime MERİC KONAR1, Eda KARAİSMAİLOGLU2, Ahmet Ergun KARAAGAOGLU3

1Department Of Biostatistics And Medical Informatics, Ahi Evran University, Kirsehir/Turkey, 2Department Of Biostatistics, Kastamonu University, Kastamonu / Turkey, 3Department Of Biostatistics, Hacettepe University, Ankara / Turkey.

C-Reactive Protein(CRP) and Procalcitonin(PCT) biomarkers are widely utilized in Intensive Care Units(ICU). Higher levels of these biomarkers indicate risk of a disease. The aims of this study are, • To determine the important risk factors for ICU mortality with single and serial CRP and PCT values. • To identify the models with better model fit and better predictive accuracy. Cox Regression Model and Joint Modeling are performed in determination of the factors, while goodness-of-fit statistics along with residual plots are used to evaluate the model fit and time-dependent AUC is calculated for the predictive accuracies. Patients that were admitted to Hacettepe University Adult Intensive Care Units between January 2015 and March 2017 were included in the study. Serial CRP and PCT measurements of the patients are recorded with age, gender and the unit variable where patients stay. The follow-up time is planned as the first 30 days in ICU. At the end of the analyses, both single and serial PCT values are found as statistically significant, while all other covariates are found to have no effect on ICU mortality in both cases. In single-measurement case, all of the covariates and single CRP value are not found as statistically significant, however in serial-measurement case, serial CRP values and unit variable are found to have an effect on ICU mortality. Joint model for CRP gave better model fit, whereas serial PCT values are found to have high diagnostic accuracy over the follow-up. In conclusion, serial PCT are recommended to predict ICU mortality, whereas joint model of CRP should be utilized to calculate the risk scores for ICU mortality.

KEYWORDS: joint modeling, predictive accuracy, model fit
S15- COMPARING THE PERFORMANCE OF DIFFERENT STATISTICAL SOFTWARE PACKAGES FOR POWER ANALYSIS IN CLINICAL TRIALS

Gülden HAKVERDİ, Yüksel ÖZKAN, Aslı SUNER

1Department Of Biostatistics And Medical Informatics, Faculty Of Medicine, Ege University, Izmir, Turkey.

In clinical trials, power analysis is of great importance because it directly affects the outcome of the study. Power analysis, often preferred to check the reliability of the decisions made in a study, is a statistical method used to determine the optimum sample size required for the success of a given research project. Statistical software packages help researchers to make power analysis more accurate, and easy to perform. In this study, we aimed to apply power analysis with different statistical software packages and to compare the performances of these programs. Besides, a resource has been created to assist researchers by explaining the advantages and disadvantages of these packages. In application, the power calculation for two and more than two independent samples comparison groups of serum osmolality to predict dehydration in elderly people data downloaded from a published study in the UK in 2015. All the statistical power analyzes were performed with trial versions of NCSS, Stata, Gpower, IBM SPSS SamplePower V3.0.0 and R 3.3.0 statistical software packages. The power obtained for two and more than two independent samples were found 0.99 and 0.95, respectively. Although the power calculations in all packages are similar, Gpower had the most detailed information and the best performance.

KEYWORDS: Power analysis, clinical trials, statistical package programs
S16- ROBUSTNESS OF ANALYSIS OF COVARIANCES (ANCOVA) UNDER OF DISTRIBUTIONS ASSUMPTIONS AND VARIANCE HOMOGENEITY

Mustafa Agah TEKİNDAL¹, Özlem Güllü KAYMAZ², Can ATEŞ³, Beyza DOĞANAY ERDOĞAN⁴

¹Department Of Biostatistics, Faculty Of Veterinary Medicine, Selçuk University, Konya, Turkey , ²Department Of Statistics, Faculty Of Sciences, Ankara University, Ankara, Turkey , ³Department Of Biostatistics, Faculty Of Medicine, Van Yüzüncü Yıl University, Van, Turkey , ⁴Department Of Biostatistics, Faculty Of Medicine, Ankara University, Ankara, Turkey.

Analysis of covariance (ANCOVA) is used when there are explanatory variables measured at the third measurement level as well as factors measured at the classification level. The explanatory variables at the third level of measurement are called covariates, or concomitant variables, or supplementary variables. As in all parametric methods, the ANCOVA method assumes that normal distributions of errors, homogeneity of variances, and error terms are independent of each other. However, unusual distributions in practice are more common than normal distribution. In this study, we examined the robustness of ANCOVA method when the assumptions of normality and homogeneity of variances are violated. For this purpose, a simulation studies under different scenarios was conducted. Random numbers were generated from Gamma, Beta and Normal distributions considering different groups and different sample sizes. In the simulation studies, 10000 replications were run under the null hypothesis of no group differences and type-I error rates were calculated for each scenario. The results showed that type-I error rate is affected by skewness of the distribution, sample size and homogeneity of variance. Further work can be extended by simulation studies under different distributions and parameter values.

KEYWORDS: Analysis of Covariance, Type-I error, Robustness, Gamma, Beta, Normal
S17- EVALUATION OF SELECTED SOME ENVIRONMENTAL FACTORS AFFECTING THE SUCCESS OF MATHEMATICS WITH MULTIPLE CORRESPONDENCE ANALYSIS

Mehmet Tahir HUYUT1, Sıddık KESKİN2

1Department Of Medical Services And Techniques, Vocational School, Bitlis Eren University, Bitlis, Turkey, 2Department Of Bioistatistic, Medical Faculty, Van Yüzüncü Yıl University, Van, Turkey.

It is thought that it is very important to apply multivariate statistical methods for multivariate data and their analysis, since it is imperative to evaluate all aspects of the investigative events as much as possible in order to ensure valid and reliable results. In this study, it was aimed to determine the environmental factors affecting the success of Mathematics in the ninth grade of secondary education and the relation between them by Multiple Correspondence Analysis. In the preliminary study, 35 different questionnaires were determined and these questions were enforced to 471 students from 7 different high school. Multiple correspondence analysis method was used to evaluate the obtained data. In the study, it was found that students who found the language of the teacher more understandable were more likely to succeed and those who did not had a working desk and study room less enjoyed from the mathematics. In addition, it was found that the physical possibilities of the school were positively related to the success category. Mother and father's literacy showed that does not have a positive effect on the success of the student. There was no high correlation between the settlement site factor and success. Also it was found that friends and teachers among the environmental factors were found to be more effective and female students were more adversely affected by these factors. The results revealed that multiple correspondence analysis offered significant satisfactory conclusions when the relationships between multiple categorical variables were examined with all aspects.

KEYWORDS: Key words: Multiple correspondence analysis, environmental factors affecting mathematics success, categorical variables, dimension reduction
S18- LONGITUDINAL MODELING OF GENETIC INFLUENCES BY USING LATENT CLASS GROWTH ANALYSIS.

Özge KARADAĞ 1, Serpil AKTAŞ ALTUNAY 2

1Hacettepe University, Department Of Statistics, 2Hacettepe Universtiy, Department Of Statistics.

Latent-variable models have been used for multivariate longitudinal analyses for many years and genetics has been using path analyses with unobserved variables for almost as long. Standard growth models are a useful tool for studying the research questions for which all individuals in a given sample are expected to change varying between people. However, latent class mixed model (LCMM) and latent class growth models (LCGM) identify district subgroups of individuals following a similar pattern of change over time. LCGM is a special type of growth mixture modeling, whereby the variance and covariance estimates for the growth factors within each class are assumed to be fixed to zero. By based on this assumption, all individual’s growth trajectories within a class are homogeneous. In this study it is aimed to analyze multivariate longitudinal data which includes genetic components by using latent class and growth mixture modelling techniques. A real data set is used to illustrate how these models can be applied in the presence of heterogeneously distributed trajectories depending to the different ethnic populations.

KEYWORDS: Latent-class models, genetic component, repeated measurements
S19- EXAMINATION OF INSULIN RESISTANCE IN RHEUMATIC PATIENTS WITH LOGISTIC REGRESSION ANALYSIS

Murat GÜL, İlker İLhanlı, Yunus Durmaz

1Department Of Statistics, Giresun University, Giresun,Turkey, 2Department Of Physical Medicine And Rehabilitation,Giresun University, Turkey, 3Department Of Physical Medicine And Rehabilitation, Mehmet Akif İnan Training And Research Hospital, Şanlıurfa, Turkey.

The purpose of this study is to evaluate logistic regression model which is able to define the relation between dependent variable that is binary outcome variable and the set of independent variables that contains both continuous and discrete variables. To describe the application of logistic regression analysis, it was studied on rheumatic diseases data to determine important risk factors of affecting insulin resistance. 469 patients and their results were examined and a model was created. At the end of the study, it was understood that some independent variables had an effect on insulin resistance. Logistic regression model, 98.1% of those without insulin resistance, and 97.6% of those with insulin resistance was correctly estimated.

KEYWORDS: Logistic Regression, insulin resistance, rheumatic diseases
S20- RISK ADJUSTED HOSPITAL MORTALITY PREDICTION MODEL: A CASE STUDY IN A TURKISH TRAINING AND RESEARCH HOSPITAL

Fatma GÜNTÜRKÜN, Özgül VUPA ÇİLENGİROĞLU

1Statistics; Dokuz Eylul University; Izmir; Turkey.

In today’s world, health organizations give much importance to quality and patient safety. To this end, conservation of life and prevention of excessive deaths are one of the vital objectives for health services in all countries. Although the main function of hospitals is to save lives, there is a little attention to hospital mortality. In this context; generating reliable mortality statistics and then monitoring them is a prerequisite for improvement in care and development in patient safety. In this study; a risk adjusted hospital mortality prediction model was developed by using some popular data mining techniques; logistic regression, decision trees, random forests and artificial neural networks. The data from 30182 inpatients of one of the Turkish training and research hospitals with 1155 beds were used. The data collected from inpatients whose discharge period was January to November in 2014. At the end, the performance of these approaches were compared.

KEYWORDS: Risk adjusted mortality, logistic regression, decision trees, random forest, artificial neural networks
S21- META-ANALYSIS OF MEDIAN AND RANGE REPORTED STUDIES: DIFFERENCE BETWEEN GENDER IN TERMS OF INTERNET ADDICTION SCALE SCORES IN TURKEY

Esin AVCI

1Giresun University.

Meta-analysis, combine several studies and consider more data in a statistical point of view. In medical research, it is common to compare the means between two groups via meta-analysis. To perform meta-analysis of continuous data, the mean and variance (or standard deviation) are needed. However, sometimes the studies only report the median, range and the sample size. To combine results from these studies, Hozo (2005) and Wan (2015) proposed two different methods to estimate the sample mean and standard deviation from these quantities. This study aim to determine the difference between gender in terms of internet addiction scale scores in Turkey. Using “internet addiction scale” and “IAS” keywords, 9 electronic bibliographic databases, reference lists of relevant studies and searches of the internet was performed to identify articles related to internet addiction scale. Cross-sectional design that published in English and Turkish language studies were considered for inclusion criteria if they reported mean/standard deviation or median/range for internet addiction scale according to gender. A total of 77 studies were located, with 10 meeting the inclusion criteria. Two of them were reported in median/range and the rest in mean/standard deviation. To obtain mean and standard deviation from two studies both of method (Hozo and Wan) were applied. The standardized mean difference of internet addiction scale score were statistically significant for gender for both methods. Wans’ method given the narrowest confidence interval for standardized mean difference.

KEYWORDS: medyan, range, meta-analysis, internet addiction scale, IAS
S22- DIFFERENTIAL NETWORK ANALYSIS ON MOUSE BRAIN PARTS USING PLSR

Aylin ALIN¹, Ayça ÖLMEZ², Gökhan KARAKÜLAH³, Ash SUNER⁴

¹Faculty Of Science, Department Of Statistics, Dokuz Eylül University, İzmir, Turkey, ²The Graduate School Of Natural And Applied Science, Department Of Statistics, Dokuz Eylül University, İzmir, Turkey, ³Izmir International Biomedicine And Genome Institute, Dokuz Eylül University, İzmir, Turkey, ⁴Medical Faculty, Department Of Biostatistics And Medical Informatic, Ege University, İzmir, Turkey.

Complex diseases, such as cancer and Alzheimer, are mainly caused by multiple genes that have altered expression levels. Gene co-expression network analysis is useful for detecting disease associated gene modules. A simple gene co-expression network is usually constructed under a certain biological condition. However, biological interactions are dynamic. Gene expression behaviors show differences in one condition to another, and they can form in distinct network topologies. Differential network analysis (DiNA) that is used to expose the dynamic nature of network interactions aims to detect differences between networks under different conditions. It is one of the most active research areas in computational biology. DiNA consists of two main parts; forming a gene network for each condition, and testing for statistically significant differences between networks. The aim of this study is to perform computational DiNA to identify the differences among gene networks of fore-, mid- and hindbrain of a developing mouse brain. Computational approaches for gene network construction are based on high-throughput gene expression data where the number of samples is much smaller than the number of genes (n << k). Partial Least Squares Regression (PLSR) is a very popular method, which allows studying with such data sets. In this study, we will use PLSR for constructing gene networks for each brain region of a developing mouse brain, and use distance-based approach to test for differences between network structures.

KEYWORDS: Gene Network, Differential Network Analysis, Partial Least Square Regression, NIPALS, Next Generation Sequence.
S23- COMPARISON OF NEURAL NETWORKS AND MULTIVARIATE NONLINEAR REGRESSION: CASE STUDY FOR ANIMAL SCIENCE

Ash AKILLI¹, Ozkan GORGULU²

¹Department Of Animal Science, Ahi Evran University, Kirsehir, Turkey, ²Department Of Biostatistics And Medical Informatics, Ahi Evran University, Kirsehir, Turkey.

In various applications in the field of animal breeding and nutrition, dependent variables can change with time and exhibit a curvilinear structure. The modelling of egg performance curve is one of the most common problems encountered in animal husbandry. Numerous different model structures are included in the literature in the context of nonlinear regression analysis. Artificial neural networks have been used as an alternative to nonlinear regression analysis in recent years and this method is one of the most successful methods in modelling nonlinear data structures. In this study, multivariate nonlinear regression analysis and artificial neural network methods were used to investigate the dependency structure of egg performance and egg weight variables. Artificial neural networks have performed successful estimates with quite low error rates compared to nonlinear regression analysis. The results of analysis were evaluated with mean squared error and mean absolute percentage error criteria. The study concluded that artificial neural networks could be used as an alternative method when determining the effects of research applications on the egg production process or how it changes over time.

KEYWORDS: Egg production curve, neural networks, nonlinear regression, multivariate analysis.
S24- USE OF ARTIFICIAL INTELLIGENCE METHODS IN 305-DAY MILK YIELD ESTIMATION

Ozkan GORGULU1, Ash AKILLI2

1Department Of Biostatistics And Medical Informatics, Ahi Evran University, Kirsehir, Turkey, 2Department Of Animal Science, Ahi Evran University, Kirsehir, Turkey.

The usage of mathematical models is very common in making forward looking estimates on high economic value characters in dairy cattle. In particularly, 305-day milk yield estimation is quite important problem structure. In this study, it was aimed to use the three different modelling techniques to estimate the 305-day milk yield estimate. For this purpose, multiple linear regression, artificial neural networks (ANN) and least squares support vector machines (LSSVM) are discussed. 10 different back propagation algorithms were used in ANN analyses and three different kernel functions were used in LSSVM analyses. Calving age, number of lactation, days in milk, calving season, and the first four-milk test days were used as independent variables in the study. The results of the analysis show that ANN and LSSVM can be an alternative to multiple linear regression and perform successful estimations with very low error rates. The results of the research are aimed to provide contributions to the researchers and decision-makers working in the field of animal husbandry in terms of time and economic aspects.

KEYWORDS: Artificial neural network, dairy cattle, least square support vector machine, 305 day milk yield.
The generalization of the well-known standard probability distributions are one of the most common process in statistical methodologies. In this paper, we introduce the generalization process proposed by Shaw et al (2009) which is called transmutation map approach. We apply this method to frequently used statistical distributions. The properties of these distribution are also examined. As an example we use the finance data which are obtained from Istanbul Stock Exchange (BIST) for illustration. At the end of the study, we estimate the Value at Risk (VaR) by modelling the data via transmuted distributions.

**KEYWORDS:** Generalization Procedure, Statistical Distributions, Transmuted Distributions, Value at Risk.
S26- RETROSPECTIVE RECONSTRUCTION OF ECONOMIC TIME SERIES

Margarita ROHR\textsuperscript{1}

\textsuperscript{1}University Of Valencia, Department Of Applied Economics, Valencia, Spain.

The reconstruction and modelling of economic time series is a difficult task due to the significant fluctuation levels, as well as to the presence of structural changes that occur under the influence of external/internal shocks. This paper presents an attempt of reconstruction of historical macroeconomic data, focusing on the estimation of export and import demand equations, in order to study the behaviour of the Russian foreign sector for the period 1960-2016. The simplicity of the well-known theoretical framework of the imperfect substitutes model allows performing the analysis of international trade flows with a reduced number of variables (income and competitiveness, essentially) and, in general, data sources are very abundant and reliable, which in this case has turned into one of the main difficulties. Therefore, the construction, from the original statistical sources of the USSR, Russian Federation and international organisations, of a database of macroeconomic time series related to foreign trade, including exports and imports prices indexes, makes possible not only the application of the mentioned theoretical model but also the use of other approaches to provide better understanding of different economic and social phenomena.

KEYWORDS: Time series reconstruction, Cointegration, Demand functions, Foreign trade, Russia
S27- RIDIT AND EXPONENTIAL WEIGHTS FOR MULTIRATER AGREEMENT STUDIES IN LATENT CLASS MODELS

Serpil AKTAŞ ALTUNAY¹, Ayfer Ezgi YILMAZ¹

¹Hacettepe University.

The weights exhibit the degree of discrepancy between the two categories of square cross-classifications in ordinal agreement studies. The choice of the weights affects the value of kappa. The existence of latent variable requires the latent class modeling when the raters not independent even if they rated independently. In such cases, the coefficient of agreement cannot be calculated in classical way. In this study, ridit and exponential weights are adopted to the Light’s Kappa, Hubert’s Kappa, and Mielke’s Kappa in latent class models. These weights are compared to the linear and quadratic weights on various numerical examples. The results show that all agreement coefficients indicate different level of agreement under the latent class models.

KEYWORDS: agreement, weighted Kappa, latent class models, weights
S28- THE LIFE SATISFACTION OF TURKEY: A COMPARISON OF MULTIPLE CORRESPONDENCE ANALYSIS AND JOINT CORRESPONDENCE ANALYSIS

Elif ÇOKER

1Department Of Statistics, Mimar Sinan Fine Arts University, Istanbul, Turkey.

Categorical variables are the types of variables which are observed very often especially in big surveys. Simple correspondence analysis investigates the relationship between two variables in a contingency table. In real life problems, we are generally faced with more than two variables. At this point where the interest is on the relationships between these variables, multiple correspondence analysis steps in. Joint correspondence analysis is developed as an alternative method to multiple correspondence analysis. The data set used in this study is taken from the latest available survey of Life Satisfaction of Turkey which was done by The Turkish Statistical Institute in 2016. The main aim is about analyzing the general profile of the life satisfaction of Turkey through the comparison of multiple correspondence analysis and joint correspondence analysis.

**KEYWORDS:** Multiple Correspondence Analysis, Joint Correspondence Analysis, Life Satisfaction
S29- POWER ANALYSIS OF MANTEL-HAENSZEL CHI-SQUARE STATISTIC BY USING SEVERAL SCORES FOR ORDERED CONTINGENCY TABLES

Hülya OLMUŞ¹, Ezgi NAZMAN², Semra ERBAŞ¹

¹Gazi University Faculty Of Science, Statistics, Ankara, Turkey, ²Gazi University Graduate School Of Natural And Applied Sciences, Ankara, Turkey.

The analysis of association between ordinal categorical data has received a considerable attention in practice, especially in social and medical studies since ordinal instead of nominal categorical data has several advantages. We consider contingency table having ordinal variables. Pearson and Likelihood chi-square statistics do not consider any ordering of the rows or columns in the contingency table. More powerful tests and more information can be obtained from the ordering data structure using some statistics such as Mantel-Haenszel chi-square statistic which considers the linear trend for ordinal data in order to test independence. Mantel-Haenszel Chi-square statistic is also called as linear association or linear by linear association chi-square. Correlation coefficient of this statistic is considered as the appropriate statistic to detect the association between ordinal variables. Integer scoring is generally used among several scoring systems and it can be considered while using the statistic. In this study, we used integer, midrank, ridit, exponential, Van der Waerden (Normal) and joint scores. A simulation study was performed to compare these scores in order to evaluate the difference related with the power of statistics. In addition, the differences in each score was examined with a real data set.

KEYWORDS: ordinal categorical data, scores, Mantel-Haenszel chi-square statistic, power
Credit risk modeling including the measurement of credit quality has been intensively investigated by academics and practitioners over the past decades. The aim is to contribute to this field of research by developing a framework for jointly modeling firm failures (e.g., bankruptcies) and ordinal credit ratings as outcomes. This model, unlike prior work, simultaneously incorporates failures and credit ratings and allows inference about the quantitative relationships between these outcomes by simultaneously making use of both sources of information. A dynamic latent trait model is employed for the latent creditworthiness process underlying the categorical rating and failure observations. Time dependent processes are defined for the systematic and idiosyncratic development of credit quality. Failure, firm-level and stock price data for publicly traded North American companies as well as issuer credit ratings from the big three rating agencies (S&P, Moody's and Fitch) are collected and analyzed to illustrate the proposed framework. Full Bayesian inference is performed using MCMC methods.

**KEYWORDS:** bankruptcy, credit ratings, credit risk, latent variable models, ordinal data analysis
Selecting which courses to enroll among many electives is one of the most influential decisions a student has to make in his education life. The courses a student takes might open up a career opportunity as well as affecting his GPA. Although this selection mostly thought to be trivial, the ambiguity of the factors to be considered leads the students to wrong decisions or missed chances. There have been studies on computer-aided recommendation systems for course selection to help students in making better decisions. In general, these systems depend on students’ performance in previously taken courses. However, taking only the past performances of a student into account may be misleading as there are other crucial factors that makes a course selection good or bad for him. In this study, we determined five factors that we thought to be directly affecting a course selection and investigated the level of importance of these factors according to both undergraduate and graduate students, and experts (i.e. academicians). We conducted two separate questionnaires for students and experts to this end. Then, by applying Analytic Hierarchy Process (AHP) we obtained three separate sets of weights for the determined factors. Each set of weights represents the preferences of a group of participants i.e. undergraduate students, graduate students or experts. This study helps understanding the needs and opinions of the user groups of course recommender systems. In future, we intend to develop a course recommender system using the determined factors and their respective weights.

**KEYWORDS**: Course Selection, AHP, Course Recommender Systems
S32- USING PARTICLE SWARM OPTIMIZATION FOR ESTIMATING THE PARAMETERS OF KAPPA DISTRIBUTION VIA MAXIMUM LIKELIHOOD METHOD: AN APPLICATION TO HYDROLOGY

Sukru ACITAS¹, Cagdas Hakan ALADAG², Birdal ŞENOĞLU³

¹Department Of Statistics, Anadolu University, Eskisehir, Turkey, ²Department Of Statistics, Hacettepe University, Ankara, Turkey, ³Department Of Statistics, Ankara University, Ankara, Turkey.

Two-parameter kappa distribution is commonly used in hydrology (Mielke, 1973). The popularity of it comes from the fact that its cumulative distribution and quantile function have closed forms. The quantiles have important role in hydrology, i.e. they are called as "design events" and the risk analysis is made on them (Ashkar and El Adlouni, 2015). The estimation of the quantiles depends on estimation of the parameters of the kappa distribution. Maximum likelihood (ML) is mostly utilized for estimating the unknown parameters of the kappa distribution in the related literature. Many statistical or mathematical softwares have some functions in optimization toolbox to obtain the ML estimates. Most of these functions have algorithms such as Quasi-Newton (QN) and Nelder-Mead (NM). In this study, to obtain the corresponding ML estimates, we propose to use particle swarm optimization (PSO) which is a population based heuristic optimization method and does not require any derivative (Kennedy and Eberhart, 1995). Furthermore, PSO searches different points in different regions of the search space and thus it avoids local optimum traps, see i.e. Aladag et al. (2012). We carry out an extensive Monte-Carlo simulation study to compare the performances of the ML estimators based on PSO, QN and NM algorithms. In the comparisons, we use the criteria known as deficiency and number of convergences for different sample sizes and shape parameters. Also, a data set taken from the hydrology literature is analyzed to show the applicability of ML estimation based on PSO.

KEYWORDS: Kappa distribution, estimation, maximum likelihood, particle swarm optimization, efficiency.
S33- TEXT MINING AND APPLICATIONS WITH TWITTER DATA IN R

Feten Başak DEMİR¹, Ayten YİĞİTER ²

¹Hacettepe University, Science Faculty, Statistics Department, Ankara, Turkey, ²Hacettepe University, Ankara, Turkey.

Social media tools that emerged with the recent developments in technology have begun to change the way people communicate with each other. A lot of data are available thanks to the social media sites where people share their feelings and thoughts. In other words, people whether they are aware of it or not, constantly produce data while using the social media sites. One of those sites, Twitter, just like any other social network, is a place for people to express their feeling, emotions and thoughts. Expressions of people using Twitter (tweets) are called as unstructured data in the data mining world. These data can be made sense of by analyzing them with text mining applications. In this study, under the text mining scope, the applications of Twitter data are analyzed by using R program.

KEYWORDS: Data mining, Text mining, Social network, Twitter
Anomaly detection (AD) is an important task in data analysis. It is used in different domains such as Intrusion Detection, Fraud Detection, and Health Care Systems. Similar to the outlier definition, an anomaly may be a data point or a pattern that is a collection of data points that do not conform with rest of the data. But there are conceptual differences between an anomaly pattern and an outlier. AD methods are evaluated in three categories; Supervised AD, Semi-Supervised AD, and Unsupervised AD. In unsupervised AD methods, the training dataset is not provided. Only raw data is used, that’s why unsupervised AD methods are more appropriate for real-life systems. In this study, we aim to evaluate the performances of four unsupervised AD algorithms (k-Nearest Neighbor, Local Outlier Factor, Local Density Cluster-based Outlier Factor and Histogram-Based Outlier Score) from the main three approaches of unsupervised AD methods (Nearest-Neighbors, Cluster-based Methods, and Statistical Methods) using data sets with varying characteristics. In application, KDD-99, Wisconsin Breast Cancer, Yahoo Anomaly Detection and Arrhythmia data sets are used. We use RapidMiner Software for anomaly detection tasks which has a special extension for the unsupervised AD. Also, R statistical programming language is used for constructing graphics. The results indicate that the nearest-neighbor based algorithms are better than cluster-based and histogram-based algorithms to detect anomalies.

**KEYWORDS:** anomaly detection, local outlier factor, local density cluster based outlier factor, histogram-based outlier score
A situational analysis was planned by General Directorate of Fighting with Desertification and Erosion in Ministry of Forestry and Water Affairs in Murat River Basin. The aim of this study is to analyze the survey data by data mining techniques. After the power analysis to determine a sample size, a survey was conducted by 1800 household responsible. The main theme was the benefit from support by government. The training set have contained 160 attributes which are demographical information, structure of house, income, costs, agricultural facilities, fields, productions etc. In this study, predictive analyses as binary logistic regression and two-step clustering were applied to predict the groups. In logistic regression, 30 attributes were included to the analysis and it was found that 14 attributes were significant after stepwise method. The predicted values were compared with actual values by chi-square test and diagnostic values were obtained as Sensitivity=89.26%, Specificity=41.37% and Accuracy=71.30%. In two-step cluster, silhouette measure of cluster quality was 0.36, and the diagnostic values were obtained as Sensitivity=76.90%, Specificity=8.44% and Accuracy=52.99%. According to the results, binary logistic regression gave us more accurate predicted values. This survey was the beginning of a Project which will be continued in later years. Therefore, these results should be examined by project planners carefully to decide whether the households in other villages benefit from governmental support for regional development or not. Predictive analytics by data mining algorithms is very useful method that analyze current or historical facts about future or unknown events.

**KEYWORDS:** Regional development, river basin, forestry villages, logistic regression, two-step cluster
S36- NEWS ANALYZE USING TEXT MINING METHODS

Kaan TOPRAK¹, Ugur SEVİK¹, Tolga BERBER¹

¹Statistics And Computer Sciences, Karadeniz Technical University, Trabzon, Turkey.

In text mining, classification, clustering, or relationships of data are analyzed using statistical data mining methods. Accordingly, datasets that contain text, such as documents, articles, e-mails can be used for text mining. The most common problems in datasets taken from the web sources for text mining are punctuation marks, html tags and stopwords. Stopwords are basically a set of commonly used words in any language which filtered out before or after processing of text analyzing. The most important problem in Turkish text analysis is due to character coding. There are some libraries and methods in R for encoding of any languages texts. The aim of this study is to analyze the last news in Trabzon and Sivas. Our datasets consist of the last 3000 news from www.hurriyet.com news site for both provinces. The data was analyzed using term frequency-inverse document frequency (TF-IDF) approach using R programming. As a result of the analysis, the words in the word cloud related to Trabzon are “Trabzon”, “Trabzonspor”, “Istanbul”, “Gol”, “Mac”, “Milli”, “Bulutlu”, “Teknik”, “Baskani” and those related to Sivas are “Sivas”, “Cumhuriyet”, “Türkiye”, “Belediye”, “Demir”, “Grup”, “İstanbul”, “Eğitim”, “Ankara”, “Belediyespor”.

KEYWORDS: Term Frequency, Inverse Document Frequency, Text Mining, Data Mining, Web Mining.
S37- A HYBRID CLASSIFICATION METHOD ON THE PREDICTION OF FOOTBALL MATCHES RESULT

Ismail Hakki KINALIOGLU¹, Coşkun KUŞ, Ismail KINACI²

¹Department Of Computer Programming, Distance Education Vocational School, Usak University, Usak, Turkey, ²Department Of Statistics, Faculty Of Science, Selcuk University, Konya, Turkey.

One of the biggest requirements of our time is to make sense of the data. The amount of accessible data is increasing rapidly day by day. This data is processed for many purposes such as classification, clustering, prediction, decision making, etc. There are many ways to do these and new methods are constantly being developed. In this study, classification algorithms and factor analysis have been used together to develop a hybrid model to derive meaning from the data. These are the most commonly used methods for the purpose of extracting information from data. The data set of the study consists of match statistics in UEFA organizations between 2010-2018. The required data was obtained from the site named WhoScored.com in which sharing the football statistics. In the first stage of the developed model, the most commonly used classification algorithms were applied on the data set and a new data set was created with the results. The columns in this new data set consist of the results of the classification algorithms used. In the second phase, factor analysis are applied on the new data set. The results of the Hybrid model, which was created to achieve a more successful prediction, are shared in this study.

KEYWORDS: Classification, Hybrid Classification, Factor Analysis, Match Result Predict
Document clustering (text clustering) is a research field in text mining, which aims to apply cluster analysis on text data to organize, navigate, summarize or retrieve text data in an efficient way. The immense quantity of data available on the web involves effective processing of text data. Cluster analysis can be utilized to extract useful information from text documents. With the use of cluster analysis, document organization, corpus summarization and document classification can be handled in an efficient way. Cluster analysis has been utilized in a wide range of applications, including text mining. In response, there are many clustering methods and algorithms developed in the literature. Conventional clustering methods (such as, K-means, K-means++ and expectation maximization) have been already utilized in clustering text documents. However, conventional clustering algorithms suffer from instability. Their performances are strongly influenced by the parameters, characteristics of data, etc. In order to build more robust clustering schemes, cluster ensembles (also known as consensus clustering) can be utilized. This paper presents the empirical results of two homogeneous and heterogeneous clustering ensemble frameworks for document clustering. In the empirical results, the results on 15 text classification benchmarks are evaluated. The empirical results indicate that cluster ensembles can enhance the clustering quality for document clustering in terms of F-measure.

KEYWORDS: cluster ensembles, text mining, document clustering, ensemble learning, machine learning
S39- SENTIMENT ANALYSIS ON USER COMMENTS IN VIDEO SHARING NETWORKS

İsmail Hakkı KINALIOĞLU¹, Adem KARATAŞ²

¹Department Of Computer Programming, Usak Univerity, Usak, Turkey, ²Social Studies Education, Faculty Of Education, Usak University, Usak, Turkey.

Social media platforms that vary depending on audience, content and intended use are based on the interactions between users. Since shared content reaches very high dimensions, the number of researches based on social media data is increasing continuously. Sentiment analysis and opinion mining are the most popular methods used in these researches. The behaviours (comment, share, like, dislike, etc.) of users on social media platforms constitute the data sets of these researches. This study used data obtained from Youtube.Com, one of the most popular video sharing sites. Comments of educational videos shared for children were analyzed with sentiment analysis and opinion mining methods. Sentiment analysis was considered as a text classification problem and classification was performed using the K-Nearest neighbor algorithm. Comments of randomly selected videos were reviewed for the classification process. Positive and negative dictionaries were formed by determining the most frequent words from selected comments. The words are simplified. At the next stage, ten thousand random comments were selected for use in training the model. Two thousand comments were randomly chosen to test the generated model and these comments were marked as positive or negative. The results of the classification process are shared in the last section.

KEYWORDS: Sentiment Analysis, Opinion Mining, Text Mining, K-Means Cluster
S40- DISEASE DIAGNOSIS with SUPPORT VECTOR MACHINES: AN APPLICATION for DIAGNOSIS of DIABETES

Melda KOKOÇ¹, Güvenç ARSLAN²

¹Gazi University Rectorate, Ankara, Turkey, ²Dept. Of Statistics, Kırıkkale University, Kırıkkale, Turkey.

With the increase of bad nutritional behavior and consumption of fast-foods nowadays, balance in the human body is gradually deteriorating and diseases are becoming more common. Diabetes is one of these common diseases that can occur independently of age. Because early diagnosis of diabetes is important for the successful treatment of the disease and for the quality of life of the patient, this study aims to investigate diagnosis of diabetes with a high accuracy rate. In this study we considered different ratios of training and test data in order to decide for an optimal partition of the data set for analysis. After determining test and training ratios, different trials were carried out under different approaches by using Support Vector Machines (SVM) for diagnosis. Finally we applied the k-means clustering algorithm before applying the support vector machines (SVM) to increase the performance of the standard SVM approach. The results show that this combined approach improves accuracy of diabetes diagnosis and gives much better results than similar studies in the literature.

KEYWORDS: Clustering, Data mining, Diagnosis of diabetes, K-Means, Support vector machines
In a randomized block design, testing ordered alternatives can be carried out by various nonparametric tests such as Page, Jonckheere, Hollander. Page and Jonckheere tests are based on rank correlations between the within-block rankings and the criterion alternative ranking. The key assumption of these tests is that the within-block observations must be independent. In medical studies, the blocks may include repeated measures and this assumption may be unreasonable. Therefore, making the use of Page and Jonckheere tests would be inappropriate. Zhang and Cabilio (2013) generalized Jonckheere test and Akdur et al. (2016) generalized weighted Jonckheere test based on circular bootstrap method for repeated measures in randomized block with ordered alternative hypothesis. In this study, we attempt to generalize the Page test to the case where the within-block measurements are correlated. In the simulation study, the generalized Jonckheere and the generalized page tests are compared empirically based on nonparametric bootstrap and circular bootstrap methods in terms of type I error and power values. References Zhang, Y., & Cabilio, P. (2013). A generalized Jonckheere test against ordered alternatives for repeated measures in randomized blocks. Statistics in medicine, 32(10), 1635-1645. Akdur, H. T. K., Gokpinar, F., Bayrak, H., & Gokpinar, E. (2016). A Modified Jonckheere Test Statistic for Ordered Alternatives in Repeated Measures Design. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 20(3), 391-398.

KEYWORDS: repeated measures, nonparametric tests, ordered alternatives.
S42- LINEAR CONTRASTS IN ONE-WAY ANOVA UNDER TYPE II CENSORED SAMPLES WHEN THE DISTRIBUTION OF THE ERROR TERMS IS NON-IDENTICAL JONES AND FADDY'S SKEW T

Talha ARSLAN¹, Sukru ACITAS², Birdal ŞENOĞLU³

¹Van Yüzüncü Yıl University, Department Of Econometrics, Van, Turkey, ²Anadolu University, Department Of Statistics, Eskisehir, Turkey, ³Ankara University, Department Of Statistics, Ankara, Turkey.

Arslan and Senoglu (2017) obtained the estimators of the model parameters in one-way ANOVA and proposed a new test for testing the equality of the treatment means when the samples are type II censored assuming the distribution of the error terms is Jones and Faddy's skew t (JFST). It is known that type II censored samples are encountered frequently in various areas of science, such as agriculture, engineering, medicine and so on. Arslan et al. (2017) also considered the same problem for the linear contrasts in one-way ANOVA for the complete samples and non-identical JFST error terms; see also Şenoğlu and Tiku (2002) in the context of non-identical generalized logistic (GL) error distribution. In this study, we assume that the samples are type II censored rather than complete and the distribution of the error terms is non-identical JFST. Similar to the earlier studies, we use the modified maximum likelihood (MML) methodology in the estimation procedure, see Tiku (1967). We propose a new test for testing the linear contrasts. An extended Monte-Carlo simulation study is conducted to compare the efficiencies of the MML estimators and the tests based on them with those based on maximum likelihood (ML) estimators. At the end of the study, an example is given to show the implementation of the proposed methodology.

KEYWORDS: Linear contrasts, Type II censoring, Non-identical treatments, Jones and Faddy's skew t distribution, Modified likelihood.
Seals used to provide insulation and sealing can be manufactured from a wide variety of materials depending on the usage area. In this study, it was aimed to maximize the tensile strength of rubber-based channel cover seals. It is desired that the tensile strength of the channel cover seal should be at least 9 MPa and as high as possible according to customer requirements. For this purpose, the parameters of the vulcanization process, which is the most important process during the production of the elastomeric-based channel cover seal, are optimized using Response Surface Methodology. Time, temperature and pressure were determined as control variables of the vulcanization process. The experiment intervals of these control variables were determined as 3-5 min. for time, 160-200 °C for temperature and 120-170 bar for pressure. Experiment combinations were determined by using Box-Behnken design and two repetitions were made in each combination in order to minimize errors that may be caused by randomness and to obtain more robust results. The channel cover seals’ tensile strength were measured using tensile testing device and the data were collected. As a result of the statistical analysis, linear, quadratic, and interaction effects of control variables on the tensile strength are obtained. Statistical significance of these effects was also tested with ANOVA. Optimum factor levels were obtained as 4, 15 min. for time, 185,05 °C for temperature, and 150,32 bar for pressure. Finally, the optimum conditions have been verified by confirmation experiments.

**KEYWORDS:** Channel Cover Seal, Response Surface Methodology, Parameter Optimization
S44- OPTIMIZATION OF MULTIPLE RESPONSES USING PRINCIPAL COMPONENT ANALYSIS AND RESPONSE SURFACE METHOD FOR AN EXPERIMENTAL DESIGN

Duygu KILIÇ¹, Deniz OZONUR², Hatice Tul Kübra AKDUR², Hülya BAYRAK²

¹Gazi University, Institute Of Sciences, Department Of Statistics, Ankara, Turkey, ²Gazi University, Faculty Of Sciences, Department Of Statistics, Ankara, Turkey.

Principal component analysis (PCA) is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set. Applications of experimental design may include correlated response variables. PCA method can be used to obtain one overall response variable that contains most of the information in the set of response variables. After obtaining one overall response variable of the experimental design model, response surface method can be used to optimize the response (output) variable which is influenced by several independent (input) variables. In this study, a real dataset which contains correlated response variables of an experimental design model was analyzed by PCA to obtain one output response. By using the output response response surface method was applied to independent variables of the model in order to identify the reasons for changes in the output response.

KEYWORDS: multiple responses, response surface method, principal component analysis
Control charts such as Cumulative Sum (CUSUM) and Exponential Weighted Moving Average (EWMA) have been developed as an alternative to the Shewhart control charts in the literature for small but continuous steep shift. Any specific signal such as spike, step, bump and rump that may occur in the process can be detected earlier by Cumulative Score (CuScore) control charts, which are a generalized version of the Shewhart, CUSUM and EWMA. In this study, the process for some particular signal type (bump) is modelled for ARMA (1,1) Computer Stock Data to find the optimum values of CuScore control chart parameters. Then the probabilities of identifying (detection rate) and misidentifying the signals (false alarm) were calculated by simulation codes in SAS. In the literature, an intuitive approach is used to select the levels of graphical parameters. However, no method has been found that can find compromised solutions while minimizing the probability of false signal and maximizing the probability of signal detection. In this study, Design of Experiment (DOE) and Response Surface Methodology (RSM) were used to create experimental plans considering the values that parameters can take. The relevant experiments were conducted taking into account the levels in the experimental plans and the results were analyzed statistically. As a result of the analysis, compromised solutions were found for the objective functions formed by the obtained linear/nonlinear regression models. Desirability functions and GAMS nonlinear solvers were used to obtain optimum parameter values of the CuScore control charts.

**KEYWORDS:** CuScore Control Chart, ARMA (1,1) Time Series, Design of Experiment, Response Surface Methodology
S46- A CRITERION TO EVALUATE THE POWER OF THE TEST

Mustafa CAVUS1, Berna YAZICI1, Ahmet SEZER1

1Department Of Statistics, Anadolu University, Eskisehir, Turkey.

There are many tests can be used to test same hypotheses but it may lead to a problem for the researchers to decide the most appropriate test for their researches. Monte-Carlo simulation studies are used to obtain the performance of the tests under the specific condition. Power of the test and type I error rate are used in Monte-Carlo simulation studies to compare the performance of the tests. Not being able to compare the power of the tests with different type I error rates cause a problem. Although, Zhang and Boos (1994) and Lloyd (2005) proposed solutions to overcome this problem, their proposals have some deficiencies; since the power is wrongly adjusted when the type I error rate is lower than the nominal level. Moreover, Zhang and Boos’s adjusted power could not be used for the tests based on same test statistics. In this study, we introduce a new criterion is called penalized power of the test and it can be used to compare the tests when the type I error rates are different. A Monte-Carlo simulation study is conducted to show the efficiency of our proposal over alternatives.

KEYWORDS: penalized power, Monte-Carlo, adjusted power, unequal type I error, size corrected power
S47- CLASSIFICATION OF AIR QUALITY MONITORING STATIONS VIA STATISTICAL METHODS: A CASE STUDY FOR KOCAELI

Murat ÇOLAK¹, Gülşen AYDIN KESKİN¹, Şenay ÇETİN DOĞRUPARMAK², Hatice ESEN¹

¹Industrial Engineering Department, Kocaeli University, Kocaeli, Turkey, ²Environmental Engineering Department, Kocaeli University, Kocaeli, Turkey.

Kocaeli is a migration-receiving city because of its geographic position and it has become one of the cities which the pollution is felt due to pollutants rooted from increasing vehicles, houses and industrial activities. Therefore, 13 measurement stations (Kocaeli, Kocaeli-Dilovası(3), Kocaeli-OSB, Kocaeli-Alikahya, Kocaeli-Gebze(2), Kocaeli-Gölcük, Kocaeli-İzmit, Kocaeli-Kandıra, Kocaeli-Körfez, Kocaeli-Yeniköy) have been constructed by the ministry of environment and urbanisation to monitor the pollutants and to develop early warning system. Nitrogen dioxide (NO2), nitrogen monoxide (NO) and nitrogen oxides (NOx) are monitored in all of the 12 stations except Kocaeli station. In this study, it is aimed to determine the most appropriate cluster memberships by classifying 12 measurement stations which have similar air pollution characteristics in terms of NO, NO2 and NOX pollutants through both principle components analysis (PCA) and k-means clustering analysis methods. The station data used in this study is obtained for 1.1.2017 - 1.1.2018 date interval and includes 24 hours data for one year. The study is important from two perspectives. Firstly, devices in one of the stations that have similar air pollution characteristics could be transferred another region where a new station is planned to be constructed in the Air Quality Monitoring Network in case of need, so it is possible to decrease equipment cost. Secondly, when measurement couldn’t be made in any station owing to technical reasons, the results of this study enable to estimate pollution values of the region using the data of station that have similar characteristics.

KEYWORDS: principle components analysis, k-means clustering analysis, air pollutants
Waste electrical and electronic equipment (WEEE), or e-waste, is a major waste type and it seems that managing WEEE will be at the center of focus of governments and companies. Due to environmental, social, health and legal dimensions of e-waste, it has become crucial to estimate e-waste quantity in order to establish an efficient e-waste management chain from product design to material recovery. There are various e-waste estimation methods in the literature and the most of which are based on the input-output analysis. Each method needs different types of data and assumptions to be used. In spite of the e-waste estimation literature is growing all over the world, there are a small number of e-waste estimation studies are available for Turkey since there are several challenges to overcome. This study aims to review and compare these methods in terms of strengths, weaknesses, applicability and lack of data for the case of Turkey. Alternative methods and new trends for e-waste estimation under the constraints of Turkey are also introduced. Findings of the study may contribute to the better understanding of e-waste potential of Turkey and sheds lights onto the future studies related to WEEE estimation.

KEYWORDS: E-waste, WEEE, Estimation methods, Turkey
S49- STATISTICAL ASSESSMENT OF AIR POLLUTANT CONCENTRATION IN ALIKAHYA REGION OF KOCAELI PROVINCE

Gülşen AYDIN KESKİN1, Murat ÇOLAK2, Şenay ÇETİN DOĞRUPARMAK2, Hatice ESEN1

1Industrial Engineering Department, Kocaeli University, Kocaeli, Turkey, 2Environmental Engineering Department, Kocaeli University, Kocaeli, Turkey.

Air quality measurement stations have been constructed by ministry of environment and urbanisation for determination of air pollution and policy making for this problem. Thus, it is aimed to improve air quality through air pollution policies year by year. Various air pollutants (PM10, PM2.5 etc.) and meteorological factors (temperature, humidity, pressure etc.) are measured in these stations. When the measurement results are analyzed, it is observed that each pollutant is not measured in every station and there are missing data for other measurements depend on some reasons. Therefore, it is aimed to determine relation between measured parameters by means of correlation analysis and principles components analysis and to compare these methods each other. The results obtained in this study enable to estimate concentration of relevant pollutant when the measurement for the pollutant type can not be made for any reason at the station. The station data used in this study is obtained for 1.1.2017 - 1.1.2018 date interval and includes 24 hours data for one year. Alikahya one of the most polluted region of Kocaeli province is selected as working area. While there are hills at the northern part of this region, settlement area and coach station are located at the western part. At the southern, there are E-5 highway and settlement area and at the eastern part it is bordered by villages and Yirim Brook. In addition, various factories and TEM highway positioned near the region. This region is a worth-analyzing town due to all of these reasons.

KEYWORDS: principle components analysis, correlation analysis, statistics, air pollution
In this study, generalized asymmetric student-t (GAST) distribution is applied to GARCH models. The performance of GARCH model with GAST innovation distribution is discussed in terms of accuracy of VaR forecasts. S&P-500 index is used to compare the out-of-sample performance of GARCH-GAST model with GARCH models specified under normal, Student-t, skew-normal and skew-T innovation distributions by means of unconditional and conditional likelihood ratio tests and loss functions. Based on the empirical findings, we conclude that GARCH models with normal, skew-normal and Student-t innovation distributions produce underestimated VaR forecasts. GAST distribution provides superior fits to standardized residuals of GARCH models and produces more accurate VaR forecasts than other competitive models based on the backtesting and loss functions results.

**KEYWORDS:** Generalized Asymmetric Student-t distribution, Value-at-Risk, Backtesting, GARCH model.
Forecasting the stock market movement has continuously been an appealing endeavor to many researchers. Different methods have been tested with the sole objective of finding the best forecasting tool. Being one of the most important markets in the MENA region, The Moroccan stock market has received little share of attention in terms of research. We shed light on how portfolio managers can use one of the most famous forecasting tools in order to make decisions in asset selection. We use ARMA-GARCH and Kalman-filtered ARMA-GARCH to forecast the rates of return of 36 companies from the Casablanca Stock Exchange over a period of 1003 trading days. Our goal is not to test the accuracy of the forecasts for individual stocks but rather to investigate how Kalman Filter can help portfolio managers make better decisions when forecasting the rates of return with ARMA-GARCH. Kalman filter has shown that it greatly helps make better decisions in stock selection. The performance of the portfolio can be from 4 to 9 times better than when relying only on ARMA-GARCH. Portfolios where we long and short positions must be taken are the ones with the highest performance over the 1003 trading day period.

**KEYWORDS:** Time Series, Forecasting, ARMA-GARCH, Kalman Filter, Portfolio Management
S52- ON BOOTSTRAP PREDICTION INTERVALS FOR GARCH MODELS

Beste H. BEYAZTAS

'Dept of Statistics, Bartin University, Bartin, Turkey.

In this paper, we propose a new resampling algorithm based on block bootstrap to obtain prediction intervals for future returns and volatilities of GARCH processes. The finite sample properties of the proposed methods are illustrated by an extensive simulation study and they are applied to Japan Yen (JPY) / U.S. dollar (USD) daily exchange rate data. Our results indicate: (i) that the proposed algorithm is capable of providing valid prediction intervals for future returns and volatilities and: (ii) it is computationally more efficient than traditional method(s).

KEYWORDS: Financial time series, Prediction, Resampling methods, Exchange rate
S53- INVESTIGATING DETERMINANTS OF FOREIGN DIRECT INVESTMENT WITH BAYESIAN MODEL AVERAGING

Erhan ÇENE¹, Filiz KARAMAN¹

¹Department Of Statistics, Yildiz Technical University, Istanbul, Turkey.

In this study, Bayesian Model Averaging (BMA) is employed to reveal the determinants of Foreign Direct Investment (FDI) in 54 countries from Europe, South America, Africa and Asia for the period of 1995-2016 within a panel data context. Countries are categorized according to their development level and continent which provides a basis for applying BMA. BMA is a method that can be used for model selection especially when the number of explanatory variables is high. BMA facilitates all the possible regression models and allows considering all significant models. In that way it provides more efficient and more convenient result rather than relying on only single model. Very few studies used BMA for exploring determinants of FDI and they used cross-sectional data. This study differs from them by applying BMA in a panel data context. Initial BMA results suggest that KOF globalization index and GDP are the only two factors that are common in all of the developing, developed and transition economies which have positive effect on FDI. Population and corruption are only effective on developing countries while exchange rate have impact on developed countries. Developing countries still depends on the market size and institutional quality as well as integration to the global world. Looking at the BMA analysis results among continents, labor productivity is effective at South America while human capital has influence at Africa and Asia. Different aspects of countries allow to create a separate policy both for the firms and governments in terms of FDI.

KEYWORDS: Bayesian Model Averaging, Foreign Direct Investment, Bayesian Modeling, Panel Data Analysis
S54- DO THE COUNTRIES’ MONETARY POLICIES HAVE SPATIAL IMPACT?

Cengiz ARIKAN1, Yeliz YALÇIN1

1 Econometrics, Gazi University, Ankara, TURKEY.

Nowadays, not land border but economic cooperation and borders determine the neighborhood and closeness by globalization. No doubt, any economic event happens in any country affects other partners more and less according to economic relationship in globalization process. The desire of measuring of this interaction make occur spatial econometrics. Initially, in spatial models take into account land borders. Subsequently, studies about spatial econometric models allow economic interactions and relationships. After the global economic crises in 2008 Central Banks have started to vary monetary policy tool to ensure economic and financial stability. It is estimated that which tool will be implemented by following the policies of the central banks in which they are closely related. The spatial effect of monetary policy can be not only geographical but also economic or social. Different spatial models have set up to examine whether any spatial effect on monetary policy. Unlike other studies in this study not only geographic weight matrix but also economic weight matrix have been used in the spatial models. Different weight matrix models results have been compared and construed. Our preliminary findings reveal that there is a spatial effect on monetary policy between OECD, EU and G-20 countries. And also, economic weight matrix

KEYWORDS: Monetary Policy, Spatial Model, Spatial Impact, Econometrics
S55- ARBITRAGE STRATEGY IN ELECTRICITY MARKETS- A CASE STUDY

Betül Zehra KARAGÜL

Actuarial Sciences Department, Hacettepe University, Ankara, Turkey.

Electricity is an asset that cannot be stored. The demand for electricity at a given time and in a given location will be satisfied only if the corresponding quantity is produced and transmitted from the production node to the demand node at the same time. In deregulated markets, the electricity delivery price is determined through a matching between the supply and demand curves, provided every 5 minutes by the various market players in the network. Every 5 minutes each player provides the independent with a curve expressing the relationship between the price and amount of electricity. When the price selected for a given time results from the curves provided one hour beforehand, the price is called real-time price (RT price). The aim of the study is to design an algorithm that decides which hours on the next day warrant a short, long, or neutral position. This algorithm should result in profits on an annual, trimester, and monthly basis, while satisfying criteria on the maximum daily loss. The data provided by a private energy company contains loads (the actual value, day ahead value and two day ahead value), production comes from wind and solar and price (DA price and RT price and the RT-DA spread). First, we examine the pattern in different seasons. Secondly, we select and calculate certain variables to build the model. And then, trying to reduce random noise by clustering similar observation. Finally, we did a basic analysis of bidding optimization problem.

KEYWORDS: arbitrage strategy, optimization, clustering, electricity price.
Solar energy is presented as the most important renewable energy source as an alternative to fossil fuels. Establishing solar power plants in the right place is important for ensuring the highest performance from systems and minimizing system costs. Therefore, simple and complex solar radiation prediction models have been developed that take into account the extraterrestrial, atmospheric, and terrestrial properties of solar irradiance. The clear sky models developed to predict terrestrial solar radiation are based on the position of the sun in the sky and atmospheric conditions. The simple clear sky models are defined as a function of the solar zenith angle only, while the complex clear sky models use many atmospheric parameters such as aerosols, gases and water vapor to more accurately model the irradiance intensity reaching the atmosphere and the ground. Although simple sky models depend only on the zenith angle, the open sky models contain uncertainty because the zenith angle reflects the air mass with variable and complex features. Estimation methods need to be recalculated on the basis of fuzzy logic in order to incorporate these uncertainties into open sky models. In this study, Daneshyar–Paltridge–Proctor (DPP), Kasten–Czeplak (KC), Haurwitz, Berger–Duffie (BD), Adnot–Bourges–Campana–Gicquel (ABCG), Robledo-Soler (RS), Meinel ve Laue simple clear sky models are redefined on the basis of fuzzy logic and compared with the sample application. The application results show that clear sky models developed based on the fuzzy logic consistently reflect terrestrial measurement values.

**KEYWORDS:** fuzzy logic, solar radiation, estimation methods, clear sky models
Fossil fuels are used as the primary source of energy in meeting the rising energy demand with increasing population. The harmful gases and substances that occur during the conversion of fossil fuels to energy cause environmental and social problems. Alternative energy sources are proposed against fossil-based fuels containing environmental, social and economic problems. Renewable energy sources emerge as the most important alternative energy source with its environmental sensitivity, inexhaustible and renewable characteristics. Besides, the sun is the most important renewable energy source with its direct and indirect use characteristics as a source of all the energies on earth (except geothermal, tidal, nuclear). The economic analysis of solar power plants with high initial cost should be done with high accuracy in order to compete with other energy sources in the free market. The results of the economic analysis provide a decision on whether or not to invest in solar energy systems. Life Cycle Cost (LCC) and Levelized Cost of Energy (LCOE) methods are widely used economic analysis methods in the economic evaluation and the comparison of the large-scale solar energy system. Because the LCC and LCOE methods involve time-dependent variability and uncertainty, the calculation methods need to be redefined on the basis of fuzzy logic. In this study, LCC and LCOE methods are redefined by using the hesitant fuzzy set and energy prices are evaluated and compared more accurately and realistically involving uncertainties with a sample application.

**KEYWORDS**: hesitant fuzzy sets, solar power plant, economic analysis, life cycle cost, levelized cost of energy
S58- OUTLIER DETECTION METHODS FOR TIME SERIES DATASETS

Ekin Can ERKUŞ¹, Vilda PURUTÇUOĞLU²

¹Department Of Biomedical Engineering, Middle East Technical University, Ankara, Turkey, ²Department Of Statistics, Middle East Technical University, Ankara, Turkey.

In statistical analyses, the detection of outliers is one of the fundamental steps to clean the batch effects and to prepare the data to further analyses such as the construction of different hypothesis testing or modeling. Due to its importance, there are many outlier detection methods suggested in the literature. Some of them is particularly designed for univariate datasets such as box-plot and z-score. Some of them is applicable for high dimensional data for instance pcout or sign methods, and some of them is used for multivariate datasets like chi-square plot, adjusted quantile method and distance-distance plot. Hereby, we have recently developed a novel outlier detection method, specifically, for time series datasets. Our new nonparametric approach, called “Fourier transform detection” (FOD), is based on the fourier transformation of original measurements. From our preliminary analyses under normal density with different numbers of outliers, we have observed that FOD is a promising method to find the noisy observations. Thereby, in this study, we extend our analyses under distinct distributions, percentages of outliers and locations in the complete measurements such as the allocation of outliers in the right or left tail of the data via various Monte Carlo scenarios. We also perform it in real datasets. Then, we compare our findings with well-known outlier detection approaches and evaluate their accuracies via certain measures and computational demand. From the outputs, it is found that FOD has a better accuracy in majority of Monte Carlo cases and improves the computational cost in the

KEYWORDS: Fourier transform, outlier detection, time-series data, accuracy measures, data analyses
The entropy is a tool to get an information from a data set. The maximum entropy distributions are derived by means of the entropy functions. The entropy type distributions from Tsallis statistic which is based on Jackson q-derivative have been derived. In this paper, we derived new entropies via using the newly defined fractional derivatives in two senses. The maximum entropy principle is used to derive new probability density functions for this new entropy as well. The applications in an image analysis are performed to detect the disorderly pattern in the image. We make a comparison with the classical calculus and the entropies based on the classical calculus. It is observed that the parameters from fractional calculus have a role to overcome the non-identicality problem in data set from the image. Thus, we can have a robust version of entropy and maximum entropy distributions.

**KEYWORDS:** entropy, fractional calculus, distribution, image analysis
S60- DYNAMIC LOT-SIZING DECISIONS UNDER FUZZY DEMAND AND THE COMPARISON WITH HEURISTIC METHODS

Metin ÖNER, Yağmur USTA

1Manisa Celal Bayar Üniversitesi Uygulamalı Bilimler Yüksekokulu Bankacılık Ve Finans Bölümü Manisa Turkey, 2Manisa Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü Uluslararası Ticaret Ve Finans Manisa Turkey.

Purpose: To introduce the decision support system developed by MATLAB GUI programming for the order decisions based on popular heuristic algorithms with methods providing optimal and fuzzy solutions in dynamic inventory control models. Method: For the cases that the crisp demand exists, the optimal lot-sizing decisions were obtained by using a mix integer programming model equivalent to the Wagner-Whitin algorithm based on dynamic programming. For the cases that the fuzzy demand exists, the mix integer model was solved by the Zimmerman method. Furthermore, the order quantities have been determined by using all the popular heuristic algorithms (Silver-Meal, Least Unit Cost, Part Period Balancing, Fixed Period Demand, Lot-for-Lot, Least Total Cost, Period Order Quantity, Groff's Method, and McLaren's Method) developed for the cases that the crisp demand is known. The decision support system has been developed in which the results of the above methods can be obtained with MATLAB GUI. Findings: The operation of the methods for specific stock items of an industrial business in the automotive sector as a subsidiary manufacturing firm has been demonstrated. For crisp and fuzzy demand, the results of the optimum and fuzzy solutions and the results of the heuristic algorithms were compared respectively. Conclusion: In industrial field, heuristic algorithms are mostly preferred at the applications such as MRP, MRPII, and ERP. The solution support is provided for not only the heuristic algorithms but also for the cases where the demand is known precisely or can be expressed in a fuzzy way, with the decision support system developed

S61- A LEAN PRODUCTION STUDY IN AN AUTOMOTIVE COMPANY

Hatice ESEN¹, Murat ÇOLAK¹, Gülsen AYDIN KESKİN¹

¹Industrial Engineering Department, Kocaeli University, Kocaeli, Turkey.

In today’s competition conditions, companies have to meet customer demands by accelerating product development via customer-oriented approaches and delivering products at desired time and quality. The companies adopt lean production philosophy to reach this production performance. Lean production aims to provide resource usage for value added activities by classifying activities as value added and non-value added. In this manufacturing strategy, the value of product is determined according to expectation of the customer. One-piece flow, one of the lean manufacturing techniques used to minimize time losses in material movements, is defined as the placement of all the machines required for the production based on the workflow of the part. In this study, a one-piece flow application which aims to reduce the production time and stockless production was realized in a company operating in automotive subsidiary industry. In this company, the results of statistical process control and ABC analysis were used to determine the line in which the one-piece flow is to be applied. Therefore, material layout and production area were arranged as U line in order to make production easier. Standing works have been adapted to production process instead of sit-down works to provide effective communication and mobility. As a result, an improvement of 17% in operator number and 24% in production process was achieved. There was also a 70% reduction in inventory level. Besides, ergonomic conditions have become better by improvements made in the scope of this study.

KEYWORDS: Inventory management, lean production, one-piece flow, automotive industry.
Nonlinear regression models are more common as compared to linear ones for real life cases e.g., modeling the growth of a plant in biology, attenuation relationships in earthquake engineering, weather forecasts in climatology, inflation and economic growth in economics, etc. Besides, nonlinear regression models are much more complex to fit and to interpret. Classical parameter estimation methods, such as least squares and maximum likelihood, can be adopted to fit the model in nonlinear regression as well nevertheless explicit solutions are not obtainable unlike linear models. Although many iterative algorithms are developed to solve the problem numerically, there is no extensive study which compiles, classifies and compares these algorithms. In this study, we aim to compile these recursive methods which are used for parameter estimation in nonlinear regression and compare them with respect to several criteria such as execution time, mean squared error of the residuals, number of iterations, variances of the estimated parameters. The comparison will be conducted based on both real and simulated data under different scenarios such as small vs. large sample sizes, normal vs. non-normal error terms, simple vs complex models.

**KEYWORDS:** nonlinear regression, nonlinear least squares, iterative algorithms
S63- EXAMINATION ON THE PERFORMANCE OF SOME BIASING PARAMETERS FOR SHILLER’S ESTIMATOR IN DISTRIBUTED LAG MODEL

Nimet ÖZBAY¹, Selma TOKER¹

¹Department Of Statistics, Çukurova University, Adana, Turkey.

In the distributed lag model, biased estimators are required in the presence of multicollinearity. As a result of the use of biased estimators, the problem of selection of biasing parameter arises. This problem is also troublesome in the linear regression model, hence there are many studies in the literature to get rid of the foregoing issue. In the present work, we examine several methods on the selection of biasing parameter for the distributed lag model. A numerical example and a Monte Carlo simulation study are conducted to determine which of these selection methods perform better. As a result of these applications, best estimators of the biasing parameter in terms of mean square error are revealed and recommended. It can be inferred that the performance of the estimator of the biasing parameter depends on the degree of the multicollinearity.

KEYWORDS: Biasing parameter, Distributed lag model, Multicollinearity, Shiller estimator
The problem of classifying events to binary classes has been popularly addressed by Logistic Regression Analysis. However, there may be situations where the most interested class of event is rare such as an infectious disease, earthquake, financial crisis etc. The model of such events tends to focus on the majority class, resulting in the underestimation of probabilities for the rare class. Additionally, the model may incorporate sampling bias if the rare class of the sample is not representative of its population. It is therefore important to investigate whether such rareness is genuine or caused by an improperly drawn sample. We here conducted a simulation study by creating three populations with different rareness levels and drawing samples from each of those which are either compatible or incompatible with the actual rare classes of the population. Then, the effect of sampling bias is discussed under the two correction methods of bias due to rareness as suggested by King and Zeng.

**KEYWORDS:** Logistic Regression, Rare Event, Sampling Bias, Bias Correction
Joint location and scale models (JLSM) are valuable tools for modeling symmetrical heteroscedastic data and have received a lot of attention in recent years. However, it is common for the observables to contain outliers and involve asymmetric outcomes. In such situations, using the skew-t distribution instead of normal distribution to model the observables can be a useful tool for robustifying an analysis. Thus, in addition to modeling location and scale parameters, it may also be necessary to model the skewness parameter. To handle this, joint location scale and skewness model (JLSSM) has been proposed in the literature. Because it includes three model in itself, JLSSM doesn’t deal well with a large number of irrelevant variables. So far variable selection methods are mainly concerned with location and scale models not with skewness models. Therefore, we consider the subset selection of variables in JLSSM of the skew-t distribution. In this study, we use information complexity (ICOMP) criterion proposed by Bozdoğan (1990,1998, 2000,2004) to select the best subset of variables for all location, scale and skewness models. We demonstrate our results in a Monte Carlo simulation study and a real dataset. We also compare the performance of information complexity (ICOMP) and Akaike’s information criterion (AIC) in choosing the best model.

**KEYWORDS**: joint location scale and skewness models, skew-t distribution, variable selection
S69- ASSESSING THE PERFORMANCE OF MODIFIED RIDGE ESTIMATOR IN
SIMULTANEOUS SYSTEMS

Selma TOKER, Nimet ÖZBAY

1Department Of Statistics, Çukurova University, Adana, Turkey.

Two stage least squares estimation produces unstable estimates of the coefficients for a simultaneous equations models when the problem of multicollinearity is confronted. Using biased estimators with some prior information represents an appealing option in this context. We recommend two stage modified ridge estimator employing two stage ridge estimator as a prior information. Our new estimator will be preferable to the two stage least squares and two stage ridge estimators with regard to mean square error criterion. The numerical results are computed by an example using an aggregate econometric model of the U.S. economy. Graphical representations are also illustrated to distinguish the estimated mean square error performances. The outcomes are in favor of the two stage modified ridge estimator while estimating the coefficients of simultaneous equations model suffering from the problem of multicollinearity.

KEYWORDS: Mean square error, Modified ridge estimator, Multicollinearity, Simultaneous equations model
Nonlinear regression analysis is a statistical technique which is often preferred for solving real world problems. That analysis is frequently used in the field of applied science such as medicine, engineering, space sciences, chemistry etc. One of the biggest problems encountered in making this analysis is the problem of multicollinearity. The problem of multicollinearity which is also arising from the structure of the nonlinear models, affects the parameter estimation process negatively. Consistent and low error estimation of parameters is important in regression analysis. When the problem of multicollinearity is ignored in the parameter estimation process, estimated statistics of the model are obtained with high error. This leads to misinterpretation of the results of the analysis outputs. In this study, a new estimation method is proposed, which is developed by using the priori knowledge about the parameters in the parameter estimating process of multicollinear nonlinear regression model.

**KEYWORDS:** Nonlinear regression, multicollinearity, prior information, Mean Square Error
S71- AN INTERCITY BUS SCHEDULING PROBLEM WITH CENTRAL CITY LOCATION

Hüseyin GÜDEN¹, Barış Keçeci², Ertan Yakıcı³, Mumtaz Karatas⁴

¹Industrial Engineering Department, Eastern Mediterranean University, Gazimağusa, KKTC, ²Industrial Engineering Department, Başkent University, Ankara, Turkey, ³Industrial Engineering Department, Milli Savunma Üniversitesi Deniz Harp Okulu, İstanbul, Turkey.

This study is motivated by a real-life case of a company which was about to invest in intercity bus service business in Turkey. Companies in this business generally have central offices in the cities that they have more customers. The companies may keep a reserve fleet of buses in these central cities for unexpected events such as breakdowns and accidents. The maintenance and repair operations of the buses are performed in these cities, as well. Since the considered company wants to gain an insight into their investment and future profit, we have prepared several possible scenarios and analyzed those scenarios using optimization techniques. The suggested problem attempts to maximize the total profit by determining (i) locations of the central cities, (ii) the trips that will be performed, and (iii) the bus routes-schedules. A mathematical model and a decomposition based heuristic method are developed to solve the problem. Performance of the proposed solution method is evaluated on the prepared scenarios and we have found that it performs well.

KEYWORDS: Intercity bus scheduling, Passenger transportation, Vehicle routing, Location-routing
Nonlinear programming (NLP) is a mathematical programming model in which the objective function is not linear or one or more of the constraints do not have a linear structure. It is paid attention in past years as an important operational research area, NLP has a wide range of applications in the areas such as military, economic, engineering optimization and management. In the literature, there are many methods based on optimization theory and algorithms proposed to solve NLP problems. However, a general and efficient method for solving NLP problems has not been developed. According to the structure of objective function and constraints, suitable solution methods should be used for special models. In this work, various NLP problems are solved by using particle swarm optimization (PSO) which is one of the effective heuristic algorithms. As a result of extensive analysis, it is aimed to show the performance of the PSO method in the solution of NLP problems.

**KEYWORDS:** Optimization, Nonlinear Programming, Particle Swarm Optimization, Heuristic Algorithms
S73- A COMBINED CUSTOMER AND SUPPLIER LOGISTICS PROBLEM FOR A COMPANY: A MATHEMATICAL MODEL

Mujgan SAĞIR ÖZDEMİR

This paper presents a real case for a production company who delivers its products among different customers and simultaneously visits the suppliers on the same route to optimize the vehicles’ capacity usage. Whenever a vehicle visits either a customer to deliver required products or a supplier to get some input materials, each vehicle is required to be fully emptied or fully uploaded. A mathematical model is proposed to find the vehicle routes. The problem is different than the well-known vehicle routing problems.

KEYWORDS: Customer and supplier logistics, mathematical model, vehicle routing
S74- OPTIMIZATION OF MULTI-ITEM VENDOR MACHINE OPERATION

Ertan YAKICI1, Hakan MERDANOĞLU2, Osman Tufan DOĞAN2, Serhan DURAN3

1Industrial Engineering Department, National Defense University (Turkish Naval Academy), İstanbul, Turkey, 2Innova IT Solutions, Ankara, Turkey, 3Industrial Engineering Department, Middle East Technical University, Ankara, Turkey.

In this study, we model and solve the operation policy problem for managing vendor machines at different locations. Each vendor machine has towers which can hold different items; therefore the decision of which items should be kept at which location and with how many towers becomes a vital decision. Moreover, dynamic pricing of the items and inventory routing is also handled within this problem. According to the uncertainties faced in demand, solution to the problem determines the allocation of towers to the items throughout a time horizon, along with the item prices at each time slot and transfer requirements of items from one location to another. Experimental results support usage of the proposed model.

KEYWORDS: Linear Programming, Dynamic Pricing, Inventory Routing
S75- PARTICLE SWARM OPTIMIZATION FOR MULTI-OBJECTIVE OPTIMIZATION PROBLEMS

Bülent ALPTEKİN¹, Deniz ALPTEKİN¹, Çağdaş Hakan ALADAĞ²

¹Department Of Statistics, Middle East Technical University, Ankara, Turkey, ²Department Of Statistics, Hacettepe University, Ankara, Turkey.

A large amount of real-life and engineering problems belong to multi-objective optimization problems (MOPs), which have multiple conflicting performance indexes or objectives to be optimized simultaneously to achieve a tradeoff, such as aerospace systems, electrical systems, biological sciences and data mining. The presence of constraints brings difficulties in optimization since the search space has to be restricted to a feasible region. Population-based optimization techniques such as Genetic Algorithms (GAs), Particle Swarm Optimization (PSO), Differential Evolution (DE) and Ant Colony Optimization (ACO) have been popular choices for multi-objective optimization problems. The main reason is that these algorithms are capable of finding a set of Pareto optimal solutions in a single run. With the success of the PSO in single objective optimization, researchers are motivated to extend the use of PSO in multi-objective optimization problems. The approach has been tested by the constrained multi-objective optimization problem and the results are presented.

KEYWORDS: Optimization, Particle Swarm Optimization, Multi-Objective Optimization
S76- MULTI-OBJECTIVE MATHEMATICAL MODELS FOR SENSOR NETWORK COVERAGE PROBLEMS

Mumtaz KARATAS, Tuğçe HATİPOĞLU

1National Defense University, Turkish Naval Academy, 2Kocaeli University.

Coverage is a fundamental issue in sensor networks and it reflects how well an area of interest or critical infrastructures are monitored. Coverage problems can be categorized as area coverage, barrier coverage, and point coverage, each having a different objective and structure. In this study, we consider comprehensive multi-objective coverage problems in which decision makers aim to achieve multiple types of coverage simultaneously while satisfying various constraints, e.g. budget, topology, organizational, resource. For this purpose, we develop non-linear and linear mathematical models and discuss the performance of those approaches in terms of solution quality and computation effort.

KEYWORDS: Coverage, Mathematical Modeling, Multi-Objective Programming
S77- BAGGED PI-SIGMA ARTIFICIAL NEURAL NETWORKS WITH ARTIFICIAL BEE COLONY ALGORITHM FOR FORECASTING

Erol EĞRİOĞLU¹, Ufuk YOLCU², Eren BAŞ³, Ali Zafer DALAR¹

¹Department Of Statistics, Giresun University, Giresun, Turkey, ²Department Of Econometrics, Giresun University, Giresun, Turkey.

Artificial neural networks have been used for forecasting of time series in the literature. In the literature, there are many types of artificial neural networks for obtaining forecasts of time series. Pi-Sigma artificial neural networks are high order networks and they can produce accurate forecasts for time series. Although point estimations can be obtained from pi-sigma artificial neural networks, variance of random sample is ignored and confidence intervals cannot be obtained. In this study, “independent identical distributed residual bootstrap with rejection sampling” bootstrap method is applied for Pi-Sigma artificial neural networks and forecasts are obtained from sampling distributions as a suitable central tendency measure. Artificial bee colony algorithm is used to train Pi-Sigma artificial neural network. The proposed method is compared with some common methods in the literature by using real world time series data sets.

**KEYWORDS**: Pi-Sigma artificial neural network, Bootstrap methods, Artificial Bee colony algorithm
S78- COMPARISON OF CLASSIFICATION ACHIEVEMENTS OF LOGISTIC REGRESSION AND DECISION TREES ALGORITHMS ON TIMSS 2015 DATA

Enes FILIZ', Tugay KARADAG', Oykum Esra ASKİN', Ersoy OZ'

1Yıldız Teknik Üniversitesi; Fen Edebiyat Fakültesi; İstatistik Bölümü, 2Yıldız Teknik Üniversitesi Fen Edebiyat Fakültesi İstatistik Bölümü ; Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü.

Trends in International Mathematics and Science Study (TIMSS), is held every four years with mathematics and science students in the fourth and eighth grades. It allows researchers to make comparisons between results in terms of student achievement at international level. Such studies at international level play an important role in predicting the success of students. Machine learning algorithms are becoming popular in recent years in order to guide the researchers in classification, prediction and identification of achievement. The purpose of this study was to classify students' mathematical achievements with using 37 independent variables. In this respect, the results of mathematical achievement for 8th grade students in Turkey was classified with using decision trees (C4.5, RepTree, Random Forest) and logistic regression methods. In addition, correlation based, ReliefF, Info Gain and One R feature selection algorithms were used in order to determine the variables affecting the classification. According to the analyzes performed with the Weka program, the results were given comparatively. Logistic regression was found the most successful classification algorithm with 80.2% when using all independent variables. Also, Random Forest algorithm gave the most successful result with 80.1% within all decision tree algorithms. According to the results of analysis made with 6 variables by using feature selection method, Logistic Regression was determined the most successful method with %79 success rate.

KEYWORDS: Machine Learning, Decision Tree Algorithms, Feature Selection, Classification
S80- RANDOM FOREST AND BOOSTING TREE ALGORITHMS IN THE PREDICTION AND DIMENSIONALITY REDUCTION MODELING OF BIOMASS ENERGY SYSTEM

Halil AKBAŞ¹, Gültekin ÖZDEMİR²

¹Graduate School Of Natural And Applied Sciences, Department Of Industrial Engineering, Süleyman Demirel University, Isparta, Turkey, ²Faculty Of Engineering, Department Of Industrial Engineering, Süleyman Demirel University, Isparta, Turkey.

A prediction model for renewable biomass energy system in industry is established. The model is built by random forest based on industrial data collected on a daily basis. The high dimensionality of system variables are reduced by using boosting tree algorithm. Prediction accuracy for both random forest prediction models is estimated by using statistical indicators. The results of the study show that random forest prediction model with dimensionality reduction provides better results than the prediction model without importance analysis. Prediction results of biomass energy system are presented with analyses in this study.

KEYWORDS: Prediction, dimensionality reduction, random forest, boosting tree, biomass
Any ranking problem that minimizes a pairwise ranking error can be represented by a system of linear equations. A fast version of gradient descent algorithm with near optimal step size and momentum factor is proposed to solve this system of linear equations iteratively. Tikhonov regularization is also integrated in this framework to avoid overfitting problems where we have very large and high dimensional, but sparse data. Resulting algorithm outperforms standard gradient descent with respect to convergence times regardless of the difficulty of the problem and the level of regularization.

**KEYWORDS:** Bipartite ranking, pairwise ranking, regularized least squares, gradient descent
ATTITUDES TOWARDS IMMIGRANTS IN GERMANY

Malika IDIZOVA

"Nazarbayev University, School Of Humanities And Social Sciences, Astana, Kazakhstan.

This research analyzes the attitudes towards immigrants in Germany from 2004 to 2016 and tries to uncover the determinants behind, using an ordered logit regression model. My findings suggest that high income, university education, adherence to Islam, high religiosity, being employed, being from the West region of Germany, being widowed, and having never been married positively affect the view about immigrants. In contrast, those respondents whose parents were born in Germany have a less favorable attitude towards immigrants.

KEYWORDS: Germany, immigration, attitudes towards immigrants
Customer satisfaction is one of the most important issues in today’s globalized world. Product designs considering customer’s requirements for higher customer satisfaction make companies superior than others and help companies to protect and increase market shares. It is known that number of substitutes is quite high as a result of rapidly changing technology. Hence, protecting customer loyalty has become a challenging issue for companies. Predicting what customer wants and measuring customer perceptions and preferences on the related product are important in order to provide customer loyalty. These efforts can be overcome through conjoint analysis. Conjoint analysis is a statistical method that assists to measure preferences in terms of a newly developed product. This method is generally used in marketing area in order to predict customer perceptions and preferences. In this study, which car specifications become prominent by academicians is presented by using conjoint analysis. The aim of this study is finding out the prominent car specifications that academicians give importance to. Therefore, hypothetic car types which are generated by using predetermined factors and their levels are presented to target group in order to rate these cars. This target group is composed of academicians who continue their M.Sc. or Ph. D. education or are graduated from these degrees in Engineering Faculty of Dokuz Eylül University. As a result of this study, potential customers’ perceptions for buying a car are specified by this way. Additionally to this, preferences that belong to different groups of the potential customers are also compared each other.

**KEYWORDS:** Conjoint analysis, product design, customer satisfaction, marketing, car selection
In civil and military aviation, a great majority of accidents are caused by human error. Therefore, the selection of qualified air traffic controller (ATCO)s is especially important in order to improve safety in flight operations. One measure that can be taken to achieve this is to develop and implement an appropriate selection process or to overcome deficiencies in the existing system. Interviews are frequently used as methods for assessing the suitability of candidates for a particular profession by questionnaire and giving points to evaluate their discernment, comprehension ability, overall appearance, appropriateness, competence, and suitability of behavior and reactions. Therefore, conducting the interview examination in a more systematic way can avoid possible inconveniences and lead to a more effective selection procedure. Within this context, the student selection process in an air traffic control department is investigated, with the aim of supporting the interview examination through the Analytic Network Process. Thirty-nine candidates are ranked according to the ANP, with the current and proposed rankings being compared.

**KEYWORDS:** Air Traffic Control, Interview, Analytic Network Process
S85- INVESTIGATING CORPORATE INTEREST IN SOCIAL INNOVATION

Sercan MADANLAR¹, Burcu FELEKOĞLU²

¹The Graduate School Of Natural And Applied Sciences, Dokuz Eylul University, Izmir, Turkey; ²Department Of Industrial Engineering, Dokuz Eylul University, Izmir, Turkey.

Social innovation is finding innovative solutions that create value by solving social problems or fulfilling social needs. In this study, twenty-five high-end corporations of America are chosen and investigated in terms of how much they are interested in social innovation, what kind of applications they made and what kind of firms desire more to take part in social innovation projects. From a review of the literature on social innovation, ten criteria are identified which affect interest of corporations to pursue social innovation. Optimization direction of three criteria were determined as maximum and seven criteria were determined as minimum in this project. For each corporation, criteria values are obtained by detailed secondary data investigations. Also criteria weights were determined by experts. Obtained data is analyzed by COPRAS method. COPRAS method gives us the rank of the companies from 1 to 25 and also this method uses a point scale that is between 100 to 0. Corporations are compared by their interests in social innovation and types of firms which desire more for social innovation are identified and implications are discussed.

KEYWORDS: social innovation, multi criteria decision making, COPRAS
This study aims to develop a new approach to multiple criteria decision problems. For this purpose, we use simplified total difference formula to measure the distance of triangular fuzzy numbers to the origin for improving the Fuzzy Analytic Hierarchy Process (FAHP) in solving complex multiple criteria decision-making problems. Also, we conduct an experiment. The experiment applies different decision-making techniques (AHP, Fuzzy LLSM and proposed approach) to benchmark. Then, we discuss the different techniques based on ranking results.

**KEYWORDS:** Decision making, Analytic hierarchy process, Fuzzy sets
Measurement error is a common problem in statistical analysis that affects the results of any statistical test and that causes biased estimations. When there is a measurement error in any of the variables in the model it is important to consider the effect of this error and chose the most suitable statistical method to obtain correct results. On the other hand, when the variables include some measurement error the most common statistical approaches to analyze the data set can be listed as Type II Regression Analysis, Bland Altman method (Especially in clinical chemistry) and Structural Equation Modeling (SEM). In this study Type II regression analysis and SEM are considered as two of the measurement error models and the performances of these two techniques are compared. While the data set and the results of Type II Regression is obtained via MATLAB Software, the results of SEM is obtained via LISREL software. Mean Square Error, Akaike Information Criteria and Bayes Information Criteria are used to compare these two techniques. The results are given in related tables and figures.

**KEYWORDS**: Type II Regression Analysis, Structural Equation Modeling, Measurement Error Models, Statistical Modeling
S88- COMPARISON of LINK FUNCTIONS in BINARY RESPONSE MODEL by USING FIRTH’S PENALIZED MAXIMUM LIKELIHOOD ESTIMATION METHOD

Ezgi NAZMAN¹, Semra ERBAŞ², Hülya OLMUŞ²

¹Gazi University Graduate School Of Natural And Applied Sciences, Ankara, Turkey, ²Gazi University Faculty Of Science, Statistics Ankara, Turkey.

Binary logistic regression is one of the most widely used method in which response variable has two possible outcome providing relation with one or more than explanatory variables. Maximum Likelihood Method is frequently applied in estimating parameters of the binary logistic regression model. When sample size is small, however, maximum likelihood estimation of parameters can be biased. The lower event rate in the binary response, the more severe is the bias in the estimated parameters. On the other hand, Firth’s penalized maximum likelihood estimation reduces first-order bias in maximum likelihood estimations of parameters. Link functions of binary logistic regression are used to estimate linear relationship between binary response variable and explanatory variables. The usual link functions in binary regression model are logit, probit and clog-log which are based on cumulative distribution function. In this study, these link functions in Firth’s logistic regression model were compared in terms of Root-Mean-Squared Error (RMSE) and Akaike Information Criterion (AIC) by using various event rate, sample size and number of explanatory variables with a detailed simulation study. As a result, Probit link function in binary response model using Firth’s penalized logistic regression model has the lowest RMSE and AIC in comparison with logit and clog-log link functions.

KEYWORDS: Binary response model, Firth’s penalized maximum likelihood method, link functions, bias
INDEPENDENT COMPONENT ANALYSIS: AN ALTERNATIVE APPROACH FOR MULTICOLLINEARITY PROBLEM

Nurbanu BURSA, Hüseyin TATLIDİL

Department Of Statistics, Hacettepe University, Ankara, Turkey.

Independent component analysis (ICA) is a recently developed method for finding underlying factors or components from multivariate statistical data. What distinguishes ICA from other methods is that it looks for components that are both statistically independent, and non-Gaussian. ICA can be seen as an extension of principal component analysis (PCA). However, ICA is the more powerful technique to find the sources. Because it considers high order statistics for independence. Therefore, the ICA in analyzing multivariate data sets can reflect the intrinsic properties of the original data in a better way. ICA is also very closely related to the method called blind source separation or blind signal separation. The data analyzed by ICA could originate from many different kinds of application fields, including brain waves, telecommunication signals, digital images, psychometric measurements etc. This study aims to introduce ICA and to solve multicollinearity problem in multiple linear regression analysis with this technique. As is known, multicollinearity is an important problem and has several effects on the ordinary least squares estimates of regression coefficients. ICA decomposes the observed multivariate data into statistically independent components. For this reason, it can be used before applying multiple linear regression to eradicate multicollinearity, if data is strongly dependent. To show the performance of this technique in multicollinear data, a simulation study and a real application on Eurobond price data were conducted. Prediction ability of proposed model is found better when compared with multiple linear regression and principal component regression.

KEYWORDS: Independent component analysis, multicollinearity, regression, Eurobond
S90- ON ESTIMATION OF KENDALL'S DISTRIBUTION FUNCTION USING BERNSTEIN POLYNOMIAL APPROACH FOR SOME ARCHIMEDEAN COPULA FUNCTIONS

Selim Orhun SUSAM, Burcu HUDAVERDİ UCER

1Munzur University, Faculty Of Economics And Administrative Sciences, Department Of Econometry, 2Dokuz Eylul University, Faculty of Science, Department of Statistics.

In this study, we estimate the Kendall’s distribution function (K(t)) for Archimedean copula family using Bernstein polynomial approximation and we investigate its performance by Monte Carlo simulation. Also, Cramér-von-Mises statistic which is based on the new estimate of Kendall’s distribution function is provided for some well-known Archimedean copula functions. We investigate the power of the test for each of these copula functions and we compare the results with the conventional approach which is based on the empirical Kendall’s distribution function.

KEYWORDS: Kendall's distribution function; Bernstein polynomial; Cramer-von-Mises statistic; Archimedean copula
In 2007, Statistics Lithuania introduced the quality management system conforming to the ISO 9001 standard. The essence of the ISO based quality management is meeting customers’ requirements and striving to exceed their expectations. Long-term success is achieved when organization attracts and retains the confidence of its customer, in our case, user of statistics. Regular monitoring and reporting of quality in statistics is not only a good tool for quality improvement, it also helps to enhance the transparency, comprehensibility and usability of official statistics, what leads to the greater confidence in it. With this in view, a system for monitoring and reporting of quality and performance indicators was introduced at Statistics Lithuania. Publicly available quality reports of Statistics Lithuania are based on the ESS Single Integrated Metadata Structure (SIMS) and provide information not only on statistics quality, but also describe statistical processing, presentation and other important aspects. Such kind of information is especially relevant to the researchers and academia, as revealed by their opinion surveys, regularly performed by Statistics Lithuania. The paper introduces the main principles of the quality management system of Statistics Lithuania and explains how the system promotes user confidence in statistics through the regular quality monitoring and reporting exercise. It is explained how quality of statistics is measured and presented to the public; the usefulness of quality reports to the users, especially to researchers and academia, is discussed. Moreover, some results from the monitoring exercise and examples of quality reports are presented.

**KEYWORDS:** ISO 9001, quality monitoring, quality reporting, user-oriented quality reports, ESS SIMS.
**S92- STATISTICAL DISCLOSURE CONTROL – BENEFITS AND CHALLENGES**

**Arijana Amina RAMIC**

†(Scientific Staff At Federal Statistical Office Of Germany), Bonn, Germany.

Statistical disclosure control (SDC) is a technique used in data-driven research to ensure protection of the confidentiality of the respondents and subjects of the research. Researchers are not usually interested in information about one single person or business; they are looking for trends among larger groups of people. However, the data they use is, in the first place, linked to individual people and businesses, and SDC ensures that these cannot be identified from published data, no matter how detailed or broad. SDC identifies disclosure risk and ensure the results of the analysis are altered to protect confidentiality. It requires a balance between protecting confidentiality and ensuring the results of the data analysis are still useful for statistical research. Statistical confidentiality is a fundamental principle of official statistics enshrined in the Treaty and in the European statistics Code of Practice. Harmonisation of principles and guidelines as regards protection of confidential data is the obligation of Eurostat and national statistical authorities in the European Statistical System. This paper presents some details about some SDC methods for microdata and for the tables, and presents some benefits and challenges of their application from the point of view of the statistics producers, which include official statistics producers as well as universities, and from the point of view of the statistics users.

**KEYWORDS**: Statistical confidentiality, Statistical disclosure control, SDC methods
S93- GOOD QUALITY OF STATISTICAL TRAINING - GOOD QUALITY OF STATISTICAL INFORMATION

Vadym PISHCHEYKO¹, Arijana Amina RAMIC²

¹State Statistics Service Of Ukraine, Shota Rustavelly Str. 3, Kyiv, Ukraine, ²Germany.

The State Statistics Service of Ukraine has its own Academy of Statistics, which does the training of students and re-training of experts in official statistics. We are connecting academic work and official statistics production through systematic education and preparation for statistics profession. The National Academy of Statistics, Accounting and Auditing is a higher educational institution with a history of almost 30 years. Licensing and accreditation by the Ministry of Education and Science of Ukraine of the educational activities of the Academy in the following specializations: “Accounting and Auditing”, “Economic Statistics”, “Finance”, “Economic Cybernetics”, “Banking”, “Management of Foreign Economic Activities”. The Academy uses various forms of education and trains students for professions that meet the needs of the economy and the challenges of today; Academy graduates receive a diploma of the European sample. Academy is a basic institution of higher education for the preparation and training of employees of State Statistics of Ukraine. There are 700 Academy graduates working at the State Statistics of Ukraine and its territorial bodies, including 250 in management positions. There are 110 employees of State Statistics bodies enrolled to training programs at the Academy. Adopted Global Assessment that was conducted in Ukraine by Eurostat and UNECE in 2016 has recommended us to establish a regional training center for the countries of our region. This paper provides some more details about the mentioned collaboration, as well as about the future work of the regional training center.

KEYWORDS: Academy of Statistics, Official statistics, Training Center
The purpose of this study is to determine an alternative approach to estimate the human development index (HDI) by using economical, demographical and educational variables that are thought to be influential on the country’s human development indexes with the path analysis method and artificial neural networks. In the study, the variables analyzed is determined, variables which have significant effect on HDI, are expressed separately as direct and indirect effects, and the factorial structure between the variables is described. After these analyses, the HDI indexes are seperated into 3 classes as low-developed, medium-developed,high-developed. By evaluating these classes, classification on countries by using artificial neural networks is executed. We utilize educational, demographical and economical variables of 2015 for 100 countries on World Bank metadata.

**KEYWORDS:** Human development index, path analysis, artificial neural networks
S95- SURVEY QUALITY MEASUREMENTS AT UNIVERSITIES AND STATISTICAL OFFICE IN MONTENEGRO

Gordana RADOJEVIĆ1, Vuk CADJENOVIC1

1Statistical Office Of Montenegro, Podgorica, Montenegro.

In the era of surveys expansion, more and more attention is paid to the quality of the data obtained. This paper will provide information on methods for measuring quality and on the quality reporting manners, as well as on the institutional framework, statistical process of data processing and output in the Statistical Office of Montenegro and at universities in Montenegro. The work of the Statistical Office is based on the European Statistics Code of Practice and the TQM Quality Measurement Model. In essence of surveys carried out by universities are the epistemological principles of scientific work/research (generality, objectivity, reliability, precision, verifiability, publicity). When we look at these principles, we see that the European Code of Practice also contains scientific principles. However, a key difference lies in the ways the principles are applied and the methods of measuring quality. The Statistical Office of Montenegro, as a producer of official statistics, has a legal basis for accessing individual data, which leads to a better quality of surveys, as well as to a developed metadata database, which among others serves to produce quality reports. Nevertheless, these advantages have also another side, which is reflected in the need for more financial resources for carrying out surveys and a longer period of time for the data publication. Universities as producers of statistical data that are most often used for verification, proofing and creation of theories, as well as determination of relations between phenomena, have greater flexibility in the realization of the surveys and the ability to adapt faster to the needs of data users. This paper will present the main similarities and differences in the manners of carrying out the surveys at the universities and at the Statistical Office of Montenegro, with a special emphasis on measuring the quality of the research and the production of quality reports.

KEYWORDS: European Statistics Code of Practice, Quality measurement, Quality reports, Metadata
In last few decades, the role and popularity of statistics is constantly increasing in the society. Development of new techniques and technologies in official statistics is the consequence of continuous improving the scientific (theoretical) background in statistics followed by significant progress in development of IT solutions. Academic and official statistics play a key role in this development process. Ideally, this means that academic and official statistics representatives should closely cooperate and contribute to the overall progress of statistics. However, practical experiences show that the nature and characteristics of these relations, to a certain extent, differ and experiences vary from country to country. This paper will describe the main differences between academic statistics and official statistics as well as links between academic and official statistics productions. By using practical experiences from Bosnia and Herzegovina, this paper will provide information about three main aspects of cooperation between academic and official statistics: needs for statistician as a profession – employment opportunities, work in academic and official statistics production in parallel and cooperation between academic and official statistics producers in Bosnia and Herzegovina, through examples of good practices. At the end of this paper, future aspects and challenges in the cooperation between academic and official statistics producers in Bosnia and Herzegovina will be shortly described.

**KEYWORDS:** Academic statistics, Official statistics, Information technology, Training
S97- OPTIMIZATION OF MULTISTATIC SENSOR LOCATIONS FOR POINT COVERAGE PURPOSES

Mumtaz KARATAS1, Emily CRAPARO2, Ertan YAKICI3

1Industrial Engineering Department, Turkish Naval Academy, National Defense University, Istanbul, Turkey, 2Department Of Operations Research, Naval Postgraduate School, Monterey, CA, USA.

Multistatic Sonar Networks (MSNs) consisting of non-collocated sources and receivers are a promising development in sonar systems, but they present distinct mathematical challenges compared to the monostatic case in which each source is collocated with a receiver. In this survey, we present a portfolio of mathematical models and solution algorithms, developed with the help of students and collaborators, designed to tackle the MSN point coverage problem. The MSN point coverage problem involves locating sources and receivers to monitor a given set of critical points in space. We first discuss the optimal placement of one type of sensor, given a fixed placement of the other type. We formulate a nonlinear program and an integer linear program for solving this problem. We also develop the Divide Best Sector algorithm, which quickly provides an epsilon-optimal source position assuming fixed receivers. Next, we study the problem of optimally placing both sources and receivers. For this purpose, we describe two integer linear programs and two heuristics. We finally discuss adaptations of the point coverage problem for both definite range and probabilistic sensing models.

KEYWORDS: Sensor Networks, Optimization, Sonar
S98- THRESHOLD ACCEPTING ALGORITHM FOR THE CAPACITATED P-MEDIAN PROBLEM

İslam ALTIN¹, Aydın SİPAHİOĞLU²

¹Osmangazi University, ²Osmangazi University.

The Capacitated P-Median Problem (CPMP) is the variation of p-median which is one of the best known facility location problem. CPMP finds the optimal location of p facilities among n candidate nodes and allocates every node to a facility by taking into account capacity and demand constraints. Besides, it aims at minimizing the total distance between demand nodes and its associated facility. Since CPMP is an NP-Hard problem, many heuristic and approximation algorithms have been proposed. In this study, Threshold Accepting Algorithm (TAA) which is a deterministic variant of Simulated Annealing is applied to solve the CPMP. TAA is faster than Simulated Annealing regarding computation time, due to the fact that TAA acceptance function doesn’t contain generation of random number. Furthermore, TAA escapes from local optimal by accepting not only improving solution but also nonimproving solution. In order to test performance of the proposed algorithm, medium and large scale problems which is taken from the literature have been solved. The results have shown that the proposed algorithm can obtain quite good solutions in a reasonable computation time.

KEYWORDS: Capacitated P-Median Problem, Threshold Accepting Algorithm, Simulated Annealing.
Parameter estimation is one of the primary steps in the modelling process. Maximum likelihood (ML) is the most widely used method among various parameter estimation methods such as the least squares, the maximum likelihood, the moments method and etc. This method based on the maximization of the likelihood function with regard to unknown parameters. However, likelihood equations are nonlinear for some distributions and solving these equations by mathematical operations is a quite difficult task. Therefore, it can be advisable to use metaheuristic methods for maximizing likelihood function. In this study, four different metaheuristic algorithms, which are simulated annealing (SA), particle swarm optimization (PSO), differential evolution (DE), and artificial bee colony (ABC) are employed for the parameters estimation of considered distribution. Furthermore, the performance of these algorithms under the same conditions is evaluated in terms of time and quality of solution and also the best efficient algorithm is selected for ML parameter estimation.

**KEYWORDS:** Maximum Likelihood, Metaheuristic Methods, Parameter Estimation
Empirical likelihood estimation method can be used as an alternative to the ordinary least square (OLS) and likelihood function based methods to estimate parameters of a linear regression model when there is no distribution assumption on the error terms. The empirical likelihood method aims estimating parameters by finding probabilistic weight for each observation. To achieve this an empirical likelihood function, which contains these probabilistic weights, is maximized under some constraints. One of these constraints is a general form of the estimating equations in the OLS with probabilistic weights, since the empirical likelihood method does not assume that all observations have the same weight unlike the OLS. It is known that the OLS is easily affected by outliers. Therefore, robust alternatives such as MM estimation method can be used to deal with these outliers. In this study, we consider to obtain robust empirical likelihood estimation of parameters in a linear regression model by using the MM estimation method. Some simulation studies are performed and observed that this robust empirical likelihood estimation is less affected by the outliers compare to the classical empirical likelihood estimation method according to their mean square values.

**KEYWORDS:** Empirical likelihood, Linear models, MM estimation
S101- A NEW GOODNESS OF FIT TEST FOR NORMALITY

Deniz ALPTEKİN, Süleyman GÜNAY

'Department Of Statistics, Hacettepe University, Ankara, Turkey.

Goodness of fit tests are used for checking the validity of the distributional assumptions. The most known goodness of fit test is Pearson Chi-Square which is suggested by Karl Pearson in 1900. The usage of Pearson Chi-Square is easy, however, the power of the test is lower than the other goodness of fit tests. In this study, a new goodness of fit test based on empirical distribution function is suggested for normality. By using Monte Carlo simulation technique, critical values are calculated. Type I error rate and powers of some goodness of fit tests, such as Kolmogorov-Smirnov, Cramer von-Mises, Anderson-Darling and the suggested test, are calculated for different sample sizes. Results show that the suggested test has more power than other tests, especially in small samples.

KEYWORDS: Anderson-Darling, Cramer von-Mises, Kolmogorov-Smirnov, Monte Carlo, power of test.
S102- STATISTICAL INFERENCE FOR THE MULTIPLE LINEAR REGRESSION MODEL: SKEW-T ERROR TERMS

İklim GEDİK BALAY¹, Birdal ŞENOĞLU²

¹Department Of Banking&Finance, Ankara Yıldırım Beyazıt University, Ankara, Turkey, ²Department Of Statistics, Ankara University, Ankara, Turkey.

In this study, we extend Acıtaş et al. (2013a) to Azzalini-Capitanio's skew-t (ST) distribution and to multiple linear regression, Gedik-Balay (2014) and Islam et al. (2001a, 2001b). See also Acıtaş et al. (2013b) in the context of simple linear regression when the distribution of the error terms is Jones and Faddy's skewed-t. Likelihood equations are nonlinear functions, we therefore use maximum likelihood (ML) methodology based on iteratively reweighting algorithm (IRA) to solve them. Alternatively, we use modified maximum likelihood (MML) methodology which is non-iterative; see Tiku (1967). MML estimators have closed form, therefore they don’t have any computational difficulties. An extensive Monte Carlo simulation study is carried out to compare the performances of the ML, the MML and the traditional least squares (LS) estimators. Simulation results show that the ML estimators have the best performance among these estimators in terms of efficiency as expected. However, the MML estimators have high efficiencies besides having low computational time. We analyze a data set taken from the literature to show the usefulness of the proposed estimators.

KEYWORDS: Multiple linear regression model, skew-t distribution, modified maximum likelihood, iteratively reweighting algorithm, non-normality
S103- ESTIMATION OF PARAMETERS OF SKEW-T DISTRIBUTION UNDER PROGRESSIVE CENSORING VIA MAXIMUM LIKELIHOOD USING GENETIC ALGORITHM

Abdullah YALÇINKAYA¹, Ufuk YOLCU², Birdal ŞENOĞLU³

¹Department Of Statistics, Ankara University, Ankara, Turkey, ²Department Of Econometrics, Giresun University, Giresun, Turkey.

Skew-t (St) nests a variety of distributions having a wide range of skewness and kurtosis, see Azzalini (1985). It reduces to the well-known skew normal distribution when the degrees of freedom of St distribution tends to infinity. Therefore, it provides great flexibility in modelling not only symmetric heavy-tailed distributions but also positively skewed and negatively skewed distributions. In most of the real life problems, censored samples are more frequently encountered than the complete samples. There exist various types of censoring, among these progressive censoring is one of the most popular and widely used in statistics literature, see Balakrishnan and Aggarwala (2000) for details. In this study, we estimate the parameters of the St distribution under progressive censoring via Maximum Likelihood (ML) using the Genetic Algorithm (GA) which is a population based optimization method, see Holland (1975). In GA, we use the robust confidence intervals based on the Modified Maximum Likelihood (MML) estimators to identify the search space, see Tiku (1967) for details of MML. We use an extensive Monte Carlo simulation study for comparing the efficiencies of the ML estimators using GA with the ML estimators using classical optimization methods. At the end of the study, an example is analysed to show the implementation of the proposed methodology.

KEYWORDS: Skew-t distribution, progressive censoring, genetic algorithm, modified maximum likelihood, Monte Carlo simulation
S104 - A DISTRIBUTION FREE TEST FOR SYMMETRY COMBINING RUNS AND SIGNS

Ayman BAKLIZI

Department Of Mathematics, Statistics And Physics, Qatar University, 2713 Doha, Qatar.

A distribution free test for the symmetry of a continuous distribution about a specified median is proposed. The test is based on combing runs and signs in the sequence of ordered observations in magnitude. The performance of the proposed test is investigated and compared with other well-known tests. Simulations show that the proposed test has a superior power performance under a variety of asymmetric alternatives. A table of critical values for some small sample sizes is included.

KEYWORDS: Distribution free, Runs test, Sign test, Symmetry
In this study, we obtain the estimators of the parameters in the multiple linear regression model when the distribution of the error terms is Azzalini’s skew Normal (SN); see Gedik-Balay (2014) and Islam et al (2001a, 2001b). In estimating the model parameters, we use maximum likelihood (ML) and modified maximum likelihood (MML) methodologies; see Tiku (1967). We use iteratively reweighting algorithm (IRA) to compute the ML estimates of the parameters. Unlike the ML estimators, MML estimators are the functions of the sample observations therefore we don’t need any iterative methods to obtain them. It should be noted that MML estimators are asymptotically equivalent to the ML estimators. Therefore, they have all the asymptotic properties of the ML estimators such as unbiasedness, efficiency, normality, etc. We use an extensive Monte Carlo simulation study to compare the efficiencies of the ML and the MML estimators of the model parameters with the corresponding least squares (LS) estimators. At the end of the study, we analyze a data set taken from the literature to demonstrate the performances of the proposed estimators.

**KEYWORDS:** Multiple linear regression model, skew-normal distribution, modified maximum likelihood, iteratively reweighting algorithm, Monte Carlo simulation
Parametric methods assume that the data is normally distributed. Therefore, normality tests have an important place in statistics and many methods are found. Considering normality tests, the most effective method must be used to test whether the pre-condition of normality in the data is provided. In this study, it is aimed to compare the normality tests (Kolmogorov-Smirnov, Lilliefors, Anderson-Darling, Cramer-Von Mises, Shapiro-Wilk, Jarque-Bera) used commonly for whether data is suitable for normal distribution. In these comparisons, type-I errors and powers of the normality tests were examined considering various sample sizes, normal and non-normal distribution patterns.

**KEYWORDS:** Type-I Error, Power of Test, Kolmogorov Smirnov, Anderson Darling
S108- SCORE TEST FOR HOMOGENEITY OF INVERSE GAUSSIAN SCALE PARAMETERS

Esra GÖKPINAR¹, Gamze GÜVEN², Fikri GÖKPINAR³

¹Department Of Statistics, Gazi University, Ankara, Turkey, ²Department Of Statistics, Eskisehir Osmangazi University, Eskisehir, Turkey.

The Inverse Gaussian (IG) has many applications in different areas ranging from reliability, life testing to remote sensing and it has been widely used in modelling right-skewed data. In this study, a new test is proposed for testing the homogeneity of IG scale parameters. The proposed test based on score statistic is easy and fast to implement. We compare the proposed test and existing tests via the Monte Carlo simulation. The simulation results show that the actual size of the tests are close to the nominal level even in small sample sizes. In addition, the proposed test is more powerful than the other tests for almost all of the considered cases.

KEYWORDS: Score statistic, Inverse Gaussian distribution, Restricted Maximum Likelihood Estimator
S109- FIDUCIAL BASED APPROACH FOR TESTING THE EQUALITY OF THE TREATMENT MEANS IN ONE-WAY ANOVA WHEN THE USUAL NORMALITY AND HOMOGENEITY OF VARIANCES ASSUMPTIONS ARE NOT MET

Gamze GÜVEN¹, Özge GÜRER², Hatice ŞAMKAR¹, Birdal ŞENOĞLU²

¹Department Of Statistics, Eskisehir Osmangazi University, Eskisehir, Turkey, ²Department Of Statistics, Ankara University, Ankara, Turkey.

In this study, a new test is proposed for testing the equality of the treatment means in one-way ANOVA when the distribution of the error terms is long-tailed symmetric (Tiku and Kumra, 1985) and the variances are heterogeneous. In the estimation part of the study, we use the Tiku’s modified maximum likelihood (MML) methodology, see Tiku (1967). MML estimators are asymptotically equivalent to the maximum likelihood (ML) estimators but they do not have any computational difficulties since they are explicit functions of the sample observations. In the testing part of the study, we use the fiducial approach based on the idea of Fisher’s fiducial inference, see Fisher (1935). By using an extensive Monte Carlo simulation study, we compare the size, power and robustness property of the proposed test with the corresponding test based on normality assumption, see Li et al (2011). Simulation results show that our results are more efficient than the normal theory results especially for the small values of the shape parameter. Two tests have more or less the same efficiencies when the shape parameter increases as expected since long-tailed symmetric distribution approaches to the normal distribution while the shape parameter goes to infinity. At the end of the study, we give an example to demonstrate the usefulness of the proposed test.

KEYWORDS: One-way ANOVA, fiducial approach, nonnormality, heterogeneous variances, modified likelihood.
S110- STATIONARY BOOTSTRAP BASED MULTI-STEP FORECASTS FOR UNRESTRICTED VAR MODELS

Ufuk BEYAZTAŞ¹

¹Department Of Statistics, Bartin University, Bartin, Turkey.

This study proposes a new asymptotically valid stationary bootstrap procedure to obtain multivariate forecast densities in unrestricted vector autoregressive models. The proposed method is not based on either backward or forward representations, so it can be used for both Gaussian and non-Gaussian models. Also, it is computationally more efficient compared to the available resampling methods. The finite sample properties of the proposed method are illustrated by extensive Monte Carlo studies as well as a real-world example. Our records reveal that the proposed method is a good competitor or even better than the existing methods based on backward and/or forward representations.

KEYWORDS: Forecast density, Multivariate forecast, Vector autoregression, Resampling methods
S111- A SIMULATION STUDY ON THE PERFORMANCE OF WEIGHTED BOOTSTRAP IN THE PRESENCE of Outliers

Uğur BİNZAT¹, Engin YILDIZTEPE¹

¹Dokuz Eylül University Statistics Department İzmir Turkey.

Bootstrap is one of the most prominent resampling methods which is also commonly used in the regression analysis. But, bootstrap resamples may contain more outliers than the original sample and even a single outlier may affect the model badly. Using robust methods with bootstrap is a sensible approach in this case. But, this could not be applied easily due to some limitations such as mathematical difficulties and high computation time. Alternatively, giving less probability to outlying observations rather than equal probability in the bootstrap procedure helps to protect the model against outliers. The weighted bootstrap technique (WBT) simply consists of two phases, the probabilities are determined in the first phase, followed by the bootstrap procedure. In this study, a simulation of WBT to estimate regression coefficients is conducted and results are discussed.

**KEYWORDS:** resampling, weighted bootstrap, outlier, regression analysis
Problem of determining a risk measure for an economy with a condition of another one in distress is naturally led to the notion of systemic risk. Famous risk indicator value-at-risk (VaR) is very useful to describe the condition of an institution or a financial market. But it covers the measurement without considering dynamics of outside parties. Conditional Value-at-Risk which is called as CoVaR is used to measure a financial institution contribution to systemic risk. In this study, we analyse the systemic risk contribution of fluctuations of the brent crude oil to stock markets by using CoVaR. Together with other risk indexes, CoVaR helps us to understand better of the risks threatening the stability of the system. Also, classic VaR method is applied to realize how systemic risk computation differs from monitoring risk in isolation.

**KEYWORDS**: conditional value-at-risk, systemic risk, copula
Regression analysis is a statistical method that models the relationship between variables. The conventional ordinary least squares technique which is used for estimating the regression parameters has some assumptions such as normality and homoscedasticity. These assumptions are met uncommonly and under small departures from the underlying assumptions the conventional method can lead highly inaccurate results. In this study, after explaining the problems with ordinary least squares, five robust regression methods Coakley – Hettmansperger estimator, OP estimator, deepest regression line, SNM method and TSTS method are introduced. By conducting a simulation study, robust regression methods and ordinary least squares are compared in terms of actual significance levels. Moreover, relative efficiencies over the least squares are also investigated. Results and concluding remarks are discussed, some recommendations are given.

**KEYWORDS:** Coakley – Hettmansperger estimator, OP estimator, deepest regression line, SNM method, TSTS method
S114- ROBUST CONFIDENCE INTERVALS FOR THE DIFFERENCE OF TWO POPULATION VARIANCES

Hayriye Esra AKYÜZ\textsuperscript{1}, Hamza GAMGAM\textsuperscript{2}

\textsuperscript{1}Department Of Statistics, Bitlis Eren University, Bitlis, Turkey, \textsuperscript{2}Department Of Statistics, Gazi University, Ankara, Turkey.

The confidence intervals for the ratio of the two population variances is well known especially when the observations come from Normal distribution. On the contrary, although the interpretation of variance differences in randomized and clinical trials is very important, there is little statistical methods based on the difference of variances. In this study, confidence intervals based on robust estimators were obtained for the difference of variances of two nonnormal populations. They are estimators based on binary distances, the median of binary distances, and comedian. The performance of these confidence intervals was compared with a Monte Carlo simulation. Simulation studies were conducted based on 10000 replications for nominal alpha 0.05 with different sample sizes and distributions. The coverage probabilities and the average widths of confidence intervals based on robust estimators for the difference of the variances of two nonnormal populations were summarized. We used sample sizes as $n = 10, 20$ and $50$. Random samples are generated from Normal, Gamma, Beta, Chi-square, and Uniform distributions. According to the simulation results; the coverage probabilities of the confidence intervals for all estimators are very close to each other and the confidence intervals obtained with estimator based on binary distances have the narrowest average widths. As a result, it is recommended to use the confidence interval based on binary distances if it is required to establish a narrower confidence interval for the difference of the variances of two nonnormal populations.

**KEYWORDS:** Average width, coverage probability, robust confidence interval, robust estimator
In regression analysis, two important issues are estimation and variable selection. Popular robust regression estimation methods are combined with variable selection methods simultaneously to obtain robust estimation and variable selection in the presence of outliers. Unfortunately, some recent works show that the robust estimation methods used in those estimation and variable selection procedures are only resistant to the casewise (rowwise) outliers in the data. Also, cellwise outlier will be problem in high dimension data even though the ratio of cellwise outlier is so low. Therefore, robustness of estimation should carefully examined when both cellwise and casewise outliers are present. In this study, we proposed a robust estimation and variable selection method to deal with both cellwise and casewise outliers in the data. The proposed method has three steps. In the first step, cellwise outliers are identified, deleted and marked with NA sign in each explanatory variable. In the second step, the cells with NA signs are imputed using a robust imputation method. In the last step, robust regression estimation methods are combined with the variable selection method LASSO (Least Angle Solution and Selection Operator) to estimate the regression parameters and to select the significant explanatory variables. The simulation results reveal that the proposed estimation and variable selection procedure performs well in the presence of cellwise and casewise outliers.

**KEYWORDS:** robust variable selection, LASSO, outlier, cellwise outlier, regression
S116- RSSAMPLING: A PIONEER PACKAGE FOR RANKED SET SAMPLING WITH R

Bekir CETİNTAV1, Büşra SEVİNC2, Melek ESEMEN2, Selma GÜRLER3

1Statistics, Mehmet Akif Ersoy University, Burdur, Turkey, 2Institute Of Natural And Applied Sciences, Dokuz Eylül University, Izmir, Turkey, 3Department Of Statistics, Dokuz Eylul University, Izmir, Turkey.

Ranked set sampling (RSS) is an advanced method for data collection which is substantial for the statistical and methodological analysis in scientific studies. Especially when the exact measurement of an observation is difficult and/or expensive and ordering visually or with an auxiliary variable can be done easily, RSS is an efficient and advantageous sampling method for estimation. The basic RSS method and its modified versions come into prominence recently due to the efficiency and cost effectiveness. This paper introduces the first R package that implements the RSS and its modified versions for sampling. With RSSampling package, the researchers can sample with basic RSS and the modified versions Median RSS, Extreme RSS, Percentile RSS, Balanced-group RSS, Double RSS, L-RSS, Truncated-base RSS, Robust-Extreme RSS. The RSSampling package also allows imperfect ranking using an auxiliary variable (concomitant) which is widely used in the real life applications. Applicants can also use this new package for parametric and nonparametric inference such as mean, median, variance and correlation estimation, regression analysis, sign test and etc. where the samples are obtained via basic RSS.

KEYWORDS: ranked set sampling, modified ranked set sampling, R package, simulation
S117- A NEW FAMILY OF EXPONENTIAL TYPE ESTIMATORS FOR THE POPULATION MEAN IN STRATIFIED RANDOM SAMPLING

Ceren ÜNAL; Cem KADIŁAR

Department Of Statistics, Hacettepe University, Ankara, Turkey.

In sampling theory, information of auxiliary variable is commonly used in order to increase efficiency of estimators. Many authors have used the auxiliary information in their proposed ratio, product and regression type estimators. Alternatively, exponential type estimators have also been proposed in recent years. The aim of these studies is to propose more efficient estimator than existing estimators. For this purpose, we propose family of estimators based on the exponential function for the population mean in stratified random sampling. The expressions for the Bias and Mean Square Error (MSE) of the proposed family are derived to the first order of approximation. The properties of the exponential family of estimators are obtained and comparisons are made with some of the existing estimators. We prove that the proposed estimators are more efficient than the ratio estimator, Singh R., Kumar M., Singh R. D., Chaudhary M. K. (2008) exponential ratio estimator and classical regression type estimator under the obtained conditions. In addition, these theoretical results are supported by a numerical illustration.

KEYWORDS: Stratified Random Sampling, Exponential Type Estimators, Efficiency
S118- A NEW CLASS OF EXPONENTIAL TYPE OF MEDIAN ESTIMATOR USING AUXILIARY INFORMATION

Sibel AL¹, Hulya CINGI¹

¹Statistics, Hacettepe University, Ankara, Turkey.

In this study, we have suggested a new class of exponential type of median estimator using auxiliary information in simple random sampling. Auxiliary information is rarely used in estimating population median, although there have been many studies to estimate population mean using auxiliary information. We have found minimum mean square error (MSE) expression to compare the efficiency over other median estimators in the sampling literature. It has been shown that the proposed class of estimator is always more efficient than estimators considered by Gross (1980), Kuk and Mak (1989), Singh et al. (2003b), Bahl and Tuteja (1991) and Singh et al. (2009). In addition theoretical findings are supported by an empirical study based on three populations to show the superiority of the proposed estimators over others.

KEYWORDS: Auxiliary information, exponential estimator, mean square error, median estimation, simple random sampling
S119- ESTIMATION OF POPULATION PROPORTION FOR BINARY VARIABLE IN MODIFIED RANKED SET SAMPLING METHODS

Aylin GÖÇOĞLU¹, Neslihan DEMİREL ²

¹The Graduate School Of Natural And Applied Sciences, Department Of Statistics, Dokuz Eylül University, Izmir, TURKEY, ²Faculty Of Science, Department Of Statistics, Dokuz Eylül University, Izmir, TURKEY.

Ranked set sampling (RSS) is an alternative to simple random sampling that has become popular in recent years. RSS is used in cases for which sample units can be easily ranked, but where the exact measurement of sample units is time consuming, difficult or expensive. In this study, proportion estimators and associated variance estimators are obtained under modified ranked set sampling methods for a binary variable ranked via a concomitant variable. A Monte Carlo simulation study is undertaken to evaluate the relative efficiency of estimators of modified ranked set sampling designs for different sets and cycle sizes for various distributions. The results indicate that the proportion estimator under modified RSS methods performs better than the proportion estimator under simple random sampling (SRS).

KEYWORDS: Ranked set sampling, Modified ranked set sampling, Concomitant variable, Proportion estimator, Relative efficiency
As a modified version of ranked set sampling (RSS), percentile RSS (PRSS) is suggested by Muttlak (2003) to estimate the population mean. This sampling method is general form of quartile RSS (QRSS), extreme RSS (ERSS) and median RSS (MRSS). That means, QRSS, ERSS and MRSS can be obtained by using PRSS. In PRSS, pth and qth percentile of the sample are selected for full measurement, 0 < p < 1 and q = 1 - p. To collect sample data in PRSS, k2 units are selected without replacement from the population. Then, these units are divided randomly into k sets, each of size k. In each set, units are ranked without actual measurement. If set size k is odd, (i) From the first (k-1)/2 sets, the (p(k+1))th smallest unit is measured. (ii) The median ranked unit is measured from the (k+1)/2 set. (iii) Then, the (q(k+1))th smallest unit is measured from the remaining (k-1)/2 sets. If set size k is even, (i) From the first k/2 sets, the (p(k+1))th smallest unit is measured. (ii) Then, the (q(k+1))th smallest unit is measured from the remaining k/2 sets. To obtain n=kl sample observations, this process is repeated l cycles. In this study, we proposed the empirical distribution function (EDF) for PRSS. It is proved that this estimator is unbiased for F(x) = {0.5, 1, 0}, theoretically. And this property of the estimator is supported with numerical results. Also, the EDF when using PRSS is compared with EDFs when using ERSS, MRSS and RSS, numerically.

**KEYWORDS:** Empirical distribution function, percentile ranked set sampling, extreme ranked set sampling, median ranked set sampling
MODEL BASED INFERENCE USING JUDGMENT POST STRATIFIED SAMPLES

Konul BAYRAMOGLU KAVLAK¹, Omer OZTURK²

¹Hacettepe University, Department Of Actuarial Sciences, ²The Ohio State University, Department Of Statistics.

In survey sampling studies, statistical inference can be constructed either using design based randomization or super population model. This study develops statistical inference based on super population model in a finite population setting using judgment post stratified (JPS) samples. For a JPS sample, first a simple random sample (SRS) is constructed without replacement. The sample units in this SRS are then stratified based on judgment ranking in a small comparison set to induce a data structure in the sample. The study shows that the mean of a JPS sample is model unbiased and has smaller mean square prediction error (MSPE) than the MSPE of a simple random sample mean. Using unbiased estimators of MSPE, the paper also constructs prediction confidence intervals for the population mean. A small scale empirical study shows that the estimator performs better than SRS estimators when the quality of ranking information in JPS sampling is not poor. The paper also shows that the coverage probabilities of prediction intervals are very close to the nominal coverage probability. Proposed inferential procedure is applied to a real data set obtained from an agricultural research farm.

KEYWORDS: super population model; judgment post stratified sample, finite population
A principal components analysis of local alternatives of certain one-sided tests for the signal detection problem is presented. These tests reject a null hypothesis if a certain diffusion process crosses a specified boundary $b$. Identifying local alternatives with high power is thus equivalent to identify directions from $b$ with maximal increasing boundary crossing probabilities (BCP). In cases, where the BCP is given in closed form, the alternatives can be computed analytically. In all other cases only numerical solutions are possible. We discuss how local alternatives with high power depend on the choice of $b$.

**KEYWORDS:** Local Alternatives, Credit Risk Modelling, Boundary Crossing Probabilities, Monte Carlo Simulation
The double auction is a complex market mechanism which allows buyer and seller traders to send price offers to the market. In this market mechanism, transactions and prices occur by the interaction of buyers’ and sellers’ price offers. Buyers and sellers form their price offers by using their bidding strategies. In this study, we introduce six new bidding strategies for traders in double auction. And, these strategies are simulated through multi-agent simulations. Four of these bidding strategies are called price targeting strategies. Working principle of them is to determine a target price and to approach to this target price with an amount of learning rate. Among six strategies two of them are based on fuzzy logic. First fuzzy logic strategy is more sophisticated than the second one in terms of number of fuzzy rules. We also use zero intelligence strategy for the intention of comparison. In this work, we simulated first fuzzy logic strategy versus other strategies in buyer and seller positions. In these simulations, buyers (sellers) utilize first fuzzy logic strategy and corresponding sellers (buyers) utilize one of the other six strategies. As a result of the simulation study, it is found that first fuzzy logic strategy has higher performance than the other strategies in terms of trader profits.

**KEYWORDS:** Double Auction Market, Market Micro Simulation, Trader Behaviors, Multi-Agent Systems, Fuzzy Logic
This study focuses on the problem of traffic jam at the crossroads and attempts to address it using simulation method. In this study, the goal is to offer an improvement which reduces the traffic jam at the crossroad during the rush hours in Gaziemir which is a district of Izmir in Turkey and one of the fastest growing district due to its location near airport and Aegean Free Zone. Traffic jam is seen because of the rapid increase in population while the traffic infrastructure remains the same. The intersection of Kemal Reis and Önder streets one of the busiest junction has been determined as problematic. Two simulation models are constructed. The first model represents the current status and the aim is analyzing the traffic situation during the rush hours. Analysis of the crossroad is done using Arena Simulation Software. The inputs; vehicle arrival distributions and possibilities of car types are determined by Input Analyzer using recorded videos of the region as data sources. Performance criteria; the number of waiting vehicles at queues and the times that vehicles spent in the system, are obtained via simulation model. In the second proposed system, the traffic lights are added in order to make what-if analysis. It is seen that the queue on one of the main roads is substantially and the queue on one of the side roads is slightly reduced while the queues on the other main road and side road are slightly increased. Results of the two simulation models are compared.

**KEYWORDS:** Traffic simulation, Crossroad, Junction, Traffic jam
Tough competition in the market has led many organizations to deal with their wastes and become lean in their operations. To achieve this goal, different methods and techniques have been applied so far. The most widely used one is the Value Stream Mapping (VSM) which shows the current state of the organization. Based on these wastes, a future state map is drawn with proposed improvements. However, VSM fails to assess these improvements, therefore another technique which can be used with VSM is needed. To evaluate the proposed changes, simulation has been commonly preferred. In this paper, an real application of combining VSM and simulation in a furniture company has been shown. Furniture industry is a typical example in which a huge amount of wastes occur. VSM has been used in forming both the current and future state maps for a particular product family and then suggestions for improvement have been made. Along with VSM, a simulation model has been developed to show the current and future states in detail and a statistical comparison has been carried out.

**KEYWORDS:** Lean manufacturing, value stream mapping, simulation, furniture industry
S127- A SIMULATION-BASED RESPONSE SURFACE STUDY FOR AN EMERGENCY DEPARTMENT OF A HOSPITAL

Ceren GEGE1, Özgür YALÇINKAYA1

1Industrial Engineering Department, Dokuz Eylül University, Izmir, Turkey.

The long waiting times of patients and congestion are the most important problems of Emergency Departments (EDs) in the hospitals. In this study, for an ED of a hospital in Izmir, the aim is to find the optimal allocation of resources while minimizing the overall flow time of patients and keeping the utilization ratios of doctors and nurses at the required levels. First of all, the simulation model of the ED is built to evaluate the current system performance. Next, the data obtained from the simulation model is used for building metamodels in where response surface design is used. Lastly, metamodels are used for optimization. This study will provide help for healthcare decision makers to find the optimal configuration of resources in EDs via response surface methodology.

KEYWORDS: Emergency Departments, Simulation, Optimization, Response surface methodology, Healthcare systems
In this study, the effect of quality dimensions such as tangibles, reliability, responsiveness, assurance, and empathy, and accordingly the perceived overall quality level of the patients on their satisfaction is analyzed. For this aim, the Servqual questionnaire is applied in a teaching and research hospital in Turkey. After demonstrating the validity and reliability of the questionnaire, using the correlation analysis, the associations between the variables of the quality and satisfaction are assessed. For this, Pearson's, Spearman's and Kendall's correlation coefficients are used. Then, the surveyed relation is explained more detailed through a regression model.

KEYWORDS: Servqual, Quality, Satisfaction, Correlation Analysis, Regression
GENERALIZED ESTIMATING EQUATIONS APPROACH FOR THE RISK ASSESSMENT OF FRAGILE STATE INDEX

Harun YONAR1, Neslihan IYİT1

1Department Of Statistics, Selcuk University, Konya, Turkey.

Generalized estimating equations (GEE) approach is an extension of generalized linear model approach used in modelling longitudinal data including repeated and correlated measurements taken on the same subject. Additionally, GEE approach is suitable for modelling categorical, count and continuous outcomes in a broad perspective. Parameter estimation method in GEE approach is based on quasi-likelihood theory. The most important feature of the GEE approach is to provide modelling the correlation structure between the subjects included into the model by using different working correlation structures. The repeated measurements within subjects are not independent from each other, from this perspective GEE approach also provides opportunity of being able to model these within-subject correlations effectively. The aim of this study is to model the risk factors effective in the “fragile states index” which examines the fragility of 178 countries with 12 indicators between 2006 and 2016, by using the GEE approach. Furthermore, risk assessment of these countries will be tried to be done for each risk class in this index by GEE approach.

KEYWORDS: fragile states index, generalized estimating equations, risk assessment
S130- THE RELATION BETWEEN THE LEVEL OF SATISFACTION OF PATIENTS, AND THEIR PERCEIVED QUALITY AND LOYALTY IN A HOSPITAL IN TURKEY

Aydin TEYMOURIFAR1, Menderes TARCAN2, Onur KAYA1, Gurkan OZTURK1

1Anadolu University, 2Osmangazi University.

Generally, there is a bilateral relationship between the quality of service that patients perceive and their satisfaction level. Satisfied patients usually perceive a higher level of the quality and also a good level of quality increases the satisfaction of the patient. Although with a constant quality level, the satisfaction of patients may be reduced if their expectation changes over time and all of these affect the loyalty of the patients. In this study, a model is proposed to describe this relationship. To test the model, a modified Servqual questionnaire is applied in a training and research hospital. Based on the obtained data, the constructed model is tested using the structural equation modeling technique (SME). The confirmed model and the achieved results show the effect of the perceived quality level of the patients and their satisfaction on their loyalty.

KEYWORDS: Quality, Satisfaction, Loyalty, Servqual, Structural Equation Modeling Technique
ABSTRACT: Depression, anxiety and stress are among the most common mental disorders in human society. Therefore psychological studies on Depression Anxiety and Stress scales (DASS) have been performed frequently in recent years. It is very important to be able to analyze such psychometric disorders correctly with effective measurement techniques. One approach to address the need for an accurate measurement is to use a Rasch measurement model. This model is used to facilitate the computations of the data analysis and related interpretations with greater confidence. The aim of this study is to examine the psychometric properties of the Depression Anxiety and Stress Scales (DASS-42) using Rasch analysis model. For this purpose participants (n=316) are included to complete a pen and paper version of DASS-42 living in Istanbul and the Rasch model of measurement instruments is discussed. Ministep version of Winsteps is used to analyze the Rasch topics: item and person measures, maps and fit statistics.

KEYWORDS: Rasch, DASS, item, person
S132- GENERALIZED LINEAR MODEL BASED CONTROL CHARTS WITH POISSON RESPONSE

Ulduz MAMMADOVA, Mahmude Revan ÖZKALE ATICIOĞLU

Statistics Department, Çukurova University, Adana, Turkey.

Statistical Process Control (SPC) is a set of statistical methods that are widely used to monitor, control and improve processes. One of the primary techniques of SPC is the control chart, which is employed to study the process variation that may affect the quality of the process outcome. The control chart which detects false alarm quickly has a good performance and it is important in industrial processes. Generalized Linear Model (GLM) based control chart uses when response data are exponentially distributed. In this study, the GLM based control charts are studied when the response variable is followed a Poisson distribution. The deviance residuals obtained from Poisson regression model are used as a control chart statistics. ARL has been used as a metric for evaluating the performance of Poisson Shewhart, Poisson EWMA, and Poisson CUSUM control chart. a control chart. It is determined through simulation results that the performance of CUSUM control chart is the best. Shewhart control chart shows a good performance when the mean is shifted. EWMA control chart has the worst performance. The performance analysis of control charts constructed based on the plastic plywood data showed the same result.

KEYWORDS: Generalized Linear Models, Poisson distribution, Control charts, ARL
In this study, a Cuscroe Control Chart was designed for a particular signal type (spike) in IMA (1,1) viscosity data of a chemical process. The simulation code of this chart has been in SAS environment. Then, Design of Experiment (DOE) and Response Surface Methodology (RSM) were used to find the optimum values of CuScore control chart parameters by the objectives of maximising the detection rate and minimising the probability of false signal. In the literature, no method has been found that can find compromised solutions while minimizing the probability of false signal and maximizing the probability of signal detection. Experimental plan was made by considering the values that parameters can take. The relevant experiments were conducted taking into account the levels in the experimental plans and the results are analyzed statistically by ANOVA. As a result of the ANOVA, two non linear regression models were obtained. After validation of the results, these models were considered as objectives that conflicting each other. To optimize these conflicting objectives simultaneously, compromised local solutions were found by Desirability Functions. At the same time, global optimum solution was found by the nonlinear solvers of GAMS optimization package. In the GAMS code minimisation of the gap between detection rate and false signal was used as the single objective function. The results were consistent with the results of Desirability Function. Finally, the parameters of Cuscore control charts are determined that maximize the Desirability level and/or minimise the gap between detection rate and false signal.

KEYWORDS: Cuscroe Control Chart, Time Series, Response Surface Methodology, Desirability Function, Optimization
S134- ROBUST ALTERNATIVES TO THE TUKEY’S CONTROL CHART

Hayriye Esra AKYÜZ1, Moustafa Omar Ahmed Abu-SHAWIESH2

1 Department Of Statistics, Bitlis Eren University, Bitlis, Turkey, 2Department Of Mathematics, The Hashemite University, Al-Zarqa, Jordan.

The median absolute deviation from the sample median (MAD), Sn, and Interquartile range (IQR) are some robust estimators for population standard deviation. In this study, two robust estimators of scale are used to construct alternatives to the Tukey's Control Chart (TCC) by substitution MAD and Sn estimators, instead of the Interquartile range (IQR). To monitor the control chart’s performance, many symmetric and skewed distributions were selected to examine the Average Run Length (ARL) with various observations generated by a Monte Carlo simulation. The results found that in-control ARL values of the modified Tukey’s control chart are higher than ARL values of the traditional Tukey’s control chart (TCC) for every distribution. We conclude that the modified Tukey’s control charts have more efficient to detect the process mean when the process is in control. The out of control ARL values are worse when the distribution is nonnormal. The TCC based on estimator MAD is better than the others for the process mean monitoring performance. The TCC based on estimator Sn has approximately the same process monitoring performance as the TCC. According to these results, the TCC based on estimator MAD is recommended for practitioners over the other two estimators.

KEYWORDS: Statistical Process Control, Tukey's Control Chart, Robust Scale Estimator, Average Run Length
Economic literacy is defined as the ability to follow developments in the economy, to examine and interpret the effects of these developments, to lead individuals to be more rational and participative, and to understand the existence of economic policies. Due to today's rapid economic changes, making right decisions is becoming difficult for individuals economically illiterate and not capable of interpreting economic events. Now, being a good economy literate become one of the basic necessities of being a good citizen. This study aims to determine the economic literacy levels of university students and to make related suggestions. For this, it is planned to determine the economic literacy levels of the students studying at the faculty of economics and administrative sciences of a state university and to contribute to the related field. To achieve this, a Literacy Test and a survey on the demographic characteristics of junior and senior students at the departments of Econometrics, Economics, Business, Finance, Political Science and International Relations, which provide a 4-year education at the faculty of economics and administrative sciences of a state university, will be carried out. The tests will help to evaluate how often and how deep the students are interested in economy, whether their education is satisfying, how they respond to various economic problems and whether there is difference in their economic literacy levels from department to department.

**KEYWORDS:** Economy, economics literacy, economic consciousness.
S136- BUFFER DESIGN FOR FAILURE PRONE PRODUCTION LINES UNDER FINANCIAL CONSIDERATIONS

Tuğçe DABANLI, Seren Özmehtem TAŞAN, Derya Eren AKYOL

1Econometrics, Manisa Celal Bayar University, Turkey, 2Industrial Engineering, Dokuz Eylül University, Turkey.

In terms of cost and storage, work in process inventories (WIPs) are great burdens for the companies. The buffer capacity assigned to WIPs has a certain limit and usually cannot be unbounded. Even if there is enough space in the storage area for WIPs, for each additional buffer, unnecessary cost should be handled. Additionally, machines in real life may fail, and the time between failures and the time for repairs are random and independent of each other for each machine. The stochastic structure of the production lines allows machines to be treated as states of a Markov Chain. Within this perspective, the performance analysis of a production line becomes a vital step while considering the productivity of a system. For the performance analysis, product amount-based performance measures are usually preferred. However, since financial data may be more informative for measuring the performance of a system, the usage of cost-based performance measures should become more significant. In this study, a failure prone system is investigated with decomposition method and an approach is proposed for performance analysis of a failure prone system under financial considerations. Moreover, in order to show the applicability of the proposed approach, an illustrative example is presented. The results indicate that the optimum buffer allocations and the maximum financial measures for the production lines are reached with the proposed approach. We conclude that as the production rate increases, the annual cash flow also increases accordingly.

KEYWORDS: Buffer Allocation, Markov Chains, Decomposition Method
S137- SHORT-TERM PROBABILITY FORECASTING BY HIDDEN MARKOV MODELS:
AN APPLICATION TO A TEAM PERFORMANCE IN SOCCER

Özgür DANİŞMAN¹, Umay UZUNOĞLU KOÇER¹

¹Deparment Of Statistics, Dokuz Eylül University, İzmir, Turkey.

Hidden Markov models (HMM) are composed of two stochastic processes which can be applied in many areas such as earthquake analysis, signal processing, speech processing, communications or text recognition. The process makes transitions between unobservable (hidden) states in accordance with the Markovian property while emitting another observable stochastic sequence of random variables. The first stochastic process is a finite-state Markov chain. The second stochastic process is related to the observations in any state of the Markov chain. Any state of the hidden Markov chain depends only on the last state before it and the associated observation. The observation sequence at time t depends only on the hidden state at time t. In this study, hidden Markov model is used for modeling the structure of match results and inter-event times between successive matches. The Baum-Welch algorithm which is an iterative expectation-maximization algorithm is used for estimation of hidden state transition probability distribution, observation probability distribution and initial probability distribution. The Viterbi algorithm is used to obtain the most likely hidden state sequence, given the observation sequence. As a result, the proposed model is capable of forecasting the short-term winning, drawing or losing probabilities of next match depending on times between two successive matches.

KEYWORDS: Hidden Markov model, Baum-Welch algorithm, Viterbi algorithm, Forecasting, Parameter estimation
S138- THE EXACT DISTRIBUTION AND GEOMETRIC FUNCTION OF A GEOMETRIC PROCESS WITH ERLANG INTERARRIVAL TIMES

Mustafa Hilmi PEKALP, Halil AYDOĞDU

Department Of Statistics, Ankara University.

Geometric process (GP) arises in a wide variety of applications such as system reliability, software engineering, maintenance, queueing systems, risk and warranty analysis. The probability distribution of N(t) and its mean, also called geometric function, M(t) are basic tools in many of these applications involving GP. However, the distribution of N(t) and its mean M(t) can not be obtained in closed forms. In this study, the exact one-dimensional probability distribution of GP is derived as a power series expansion of the convolution functions by assuming that the first interarrival time follows an Erlang (gamma with integer shape parameter) distribution. Furthermore, the geometric function M(t) is also expanded into a power series by using an integral equation.

KEYWORDS: geometric process, mean value function, exact distribution, Erlang distribution
S139- THE EFFECTS OF CONTRACT MECHANISMS BETWEEN PRIVATE HOSPITALS AND THE GOVERNMENT ON THE PUBLIC BENEFIT

Aydin TEYMOURIFAR¹, Onur KAYA¹, Gurkan OZTURK¹

¹Anadolu University.

Nowadays, the health sector is as competitive as others and the price of health services seriously affects the choice of patients for the hospital choice. Generally, the examination fee in private hospitals is defined based on the contract mechanisms with the government. In this study, the effects of different contract mechanisms on the public benefit, considering the interaction between the public and private hospitals in a region is investigated. These hospitals have different characteristics such that although the payments are low in the public hospitals, the level of perceived quality and accordingly the level of satisfaction of patients are low due to crowding. Therefore, some of the patients of these hospitals will prefer the private hospitals if the prices were lower. To find a more balanced system, policies like increasing the subsidies given by the government to the patients of the private hospitals, discounting the prices in return for a fixed payment, and differentiated payment strategies according to the patient's income level are examined. In this system, while the government is trying to increase the public benefit, the private hospital wants to maximize its own profit. The comparison between the policies is done based on the public benefit concept defined as a multi-objective function, which consists of patient's average waiting time, their perceived quality level and public expenses. The results show that the well-defined contract mechanisms according to the features of the system increase satisfaction of the patients and also the public benefit. Especially, the policy of differentiated pricing according

S140- ESTIMATION OF DURATION OF MAXIMUM DRAWDOWN IN OIL PRICES

Merve Sale Salıcı, Ceren VARDAR ACAR, Fatma Pınar ERDEM KÜÇÜKBIÇAKÇI, İbrahim ÜNALMIŞ


The purpose of this study is to estimate the duration of maximum drawdown of highest possible drop in the oil prices. Maximum drawdown can be defined as the indication of the highest possible market risk in finance. In order to detect the maximum drawdown, super cycles which are specified as total of an upward and a downward movement are determined in the dataset. This is done by using filtering methods such as Hodrick-Prescott and Band Pass filter. Fractional Brownian motion model is used for modelling oil prices. We have conducted simulation studies for calculating the expected value of maximum drawdown and duration of maximum drawdown for fractional Brownian motion. In this study, comparison of simulation results to the real life data set is provided in order to make prediction and give discussion on the duration of maximum drawdown of the oil prices.

KEYWORDS: Fractional Brownian motion, stochastic, oil prices, maximum drawdown
The particles are assumed to arrive at a counter that registers the particles according to a Poisson process \( \{N_1(t), t \geq 0\} \) with rate \( \lambda \). When a particle is registered, it locks the counter for a constant locking time \( L \). Particles arrived at the counter during the locking period are lost. Only the particle registered can be observed. If we define the random variables \( Y_1, Y_2, \ldots \) as the consecutive times between two particles registered, then it can be constituted a registration process \( \{N_2(t), t \geq 0\} \) based on these random variables where \( N_2(t) \) is number of the particles registered up to time \( t \). It is obvious that \( Y_1, Y_2, \ldots \) are independent. While the random variable \( Y_1 \) has exponential distribution with mean \( 1/\lambda \), \( Y_i \)'s, \( i=2,3,\ldots \) have same distribution but different from \( Y_1 \). Hence, the counting process \( \{N_2(t), t \geq 0\} \) is delayed renewal process. In the literature, this process is called a type II counter process. In this study, we deal with the problem of estimating the mean value function of Type II counter process. A parametric plug-in estimator is proposed and a simulation study is implemented to evaluate the performance of the estimator.

**KEYWORDS:** delayed renewal process, mean value function, Type II counter, estimation.
In this study, we mainly propose an algorithm to generate correlated random walk converging to fractional Brownian motion, with Hurst parameter, $H \in [1/2, 1]$. The increments of this random walk are simulated from Bernoulli distribution with proportion $p$, whose density is constructed using the link between correlation of multivariate Gaussian random variables and correlation of their dichotomized binary variables. We prove that the normalized sum of trajectories of this proposed random walk yields a Gaussian process whose scaling limit is the fractional Brownian motion.

**KEYWORDS:** Random walks, Fractional Brownian motion, discretization, simulation
S145- A NEW DISTRIBUTION WITH FOUR PARAMETERS: PROPERTIES AND APPLICATIONS

Kadir KARAKAYA¹, İsmail KINACI¹, Coşkun KUŞ¹, Yunus AKDOĞAN¹

¹Statistics, Selçuk University, KONYA, TURKEY.

In this paper, a new distribution called compounded Geometric-mixed Exponential distribution is proposed by compounding the mixed exponential and geometric distributions. Some properties of the new distribution such as survival function, hazard function, moments, Lorenz and Bonferroni curves etc. are obtained. Parameters of the proposed distribution have been estimated by the maximum likelihood estimation (MLE) method. Performances of MLE estimators have been studied via a detailed simulation study. Four real data applications are also provided to illustrate the practical importance of the proposed distribution.

KEYWORDS: Compound distributions; Estimation; Geometric distribution; Mixed exponential distribution
S146- A NEW GENERALIZATION OF THE LOMAX DISTRIBUTION: TWO-SIDED LOMAX DISTRIBUTION

Fatih ŞAHİN1, Yunus AKDOĞAN1, Kadir KARAKAYA1, Mustafa Çağatay KORKMAZ2

1Statistics, Selçuk University, Konya, Turkey., 2Artvin Çoruh University, Artvin, Turkey.

Lomax (Lo) was originally pioneered for modelling business failure data by Lomax. It is known as second type of the beta distribution or Pareto II model. The Lo distribution has found a wide application in many fields such as size of cities, income and wealth inequality, actuarial science, engineering, medical and biological sciences, reliability and lifetime modelling. In this paper, we study the properties of the TSG-Lo distribution such as, hazard rate function, series expansion and moments. We derive its parameters estimation via maximum likelihood estimation (MLE) method. A simulation study is given to see the performance of MLEs. Finally, we fit the TSG-Lo distribution to real data set to compare it with some models in the literature.

The bimodality and asymmetry in distributions have been studied via using the different generators. The flexible distributions are important to fit data efficiently. The number of parameters and the role of parameters which are completely specified play a crucial role in modeling to increase the quality of modeling capability. For these reasons, deriving new distributions is important to overcome these problems such as the number of parameters and the role of parameters. In this study, cumulative distribution function and Kumaraswamy type distributions are used to generate a class of distributions. Asymmetry parameter is added to model the asymmetry in a data set. Thus, this class of distributions on the real line is proposed. The observed and exact Fisher information matrix for some distributions in this class are provided. The real data sets are applied to illustrate the modeling capability of distributions.

**KEYWORDS:** bimodality, asymmetry, modeling, Kumaraswamy
S149- POWER DAGUM DISTRIBUTION: PROPERTIES AND APPLICATIONS

Merve ÇOKBARLI¹, Caner TANIŞ², Buğra SARAÇOĞLU³

¹Department Of Statistics, Graduate School Of Natural Sciences (master Student), Selcuk University, Konya, Turkey, ²Department Of Statistics, Science Faculty, Çankırı Karatekin University, Çankırı, Turkey, ³Department Of Statistics, Science Faculty, Selcuk University, Konya, Turkey.

In this study, it has been introduced a new distribution with four parameters called as Power Dagum distribution using a power transformation on an random variable having to dagum distribution suggested by Dagum, C. (1977). The some statistical properties of this new distribution such as quantile function, moments, moment generating function, variation coefficient, skewness and kurtosis coefficients are derived. Maximum likelihood estimators (MLEs) and least square estimators (LSEs) of the unknown parameters of this distribution are obtained. Also, a Monte Carlo simulation study is performed in order to compare the performances of MLEs and LSEs for the parameters of this new distribution in terms of MSE and bias. Finally, a real data application for this new distribution is presented.

KEYWORDS: Power dagum distribution, maximum likelihood estimators (MLEs) and least square estimators (LSEs), hazard function, moment generating function.
S150- ON FUNCTIONS BOUNDED BY KARAMATA FUNCTIONS

Edward OMEY

KULeuven.

The class $M(\rho)$ is the class of functions satisfying $\log U(x)/\log x \to \rho$ as $x \to \infty$. This class of functions extends the class of regularly varying functions which were introduced by Karamata. It can be proved that $U \in M(\rho)$ implies that $U$ is bounded by regularly varying functions with the same index $\rho$. We extend this class and define a new class of positive and measurable functions that are bounded by general regularly varying functions. We study integrals and Laplace transforms of these functions. We use the results to study the tail of convolutions of distribution functions. The results are extended to functions and the difference of functions that are bounded by regularly varying functions or $O$-regularly varying functions. [1] Cadena, M. (2016). Contributions to the study of extreme behavior and applications. Ph.D., University Pierre et Marie Curie, Paris, France. [2] Cadena, M., Kratz, M. and Omey, E. (2017). On the order of functions at infinity. J. Math. Anal. Appl. 452 (1), 109 - 125. [3] Cadena, M., Kratz, M. and Omey, E. (2017). New results on the order of functions at infinity. Submitted.

KEYWORDS: Regular Variation, Extensions, Order of functions at infinity, Convolution product
S151- EXPONENTIED-MIXED EXPONENTIAL DISTRIBUTION

Kadir KARAKAYA¹, İsmail KINACI¹, Yunus AKDOĞAN¹, Coşkun KUŞ¹

¹Statistics, Selçuk University, Konya, Turkey.

In this study, a new continuous distribution called exponentied-mixed exponential distribution with four parameters is introduced. Some properties of the new distribution (survival function, hazard function, moments, Lorenz and Bonferroni curves etc.) are obtained. The estimates of the parameters are obtained by the Maximum likelihood, Maximum Lq-likelihood and least square methods. In addition, the MSE and Bias values of the estimates are tabulated by the simulation study. Four numerical examples are also provided.

KEYWORDS: Exponential Distribution, Maximum likelihood, Maximum Lq-likelihood, Mixed Distribution, Least square.
S152- A NEW FAMILY OF DISTRIBUTIONS: EXPONENTIAL POWER-X FAMILY OF DISTRIBUTIONS AND ITS SOME PROPERTIES

Noorsl Ahmed ZEENALABIDEN1, Buğra SARAÇOĞLU2

1Department Of Statistics, Selcuk University, Konya, Turkey., 2Department Of Statistics, Science Faculty, Selcuk University, Konya, Turkey.

In this study, it has been aimed to introduce a new family called as Exponential Power - X family of distributions by using the method suggested by Alzaatreh et al. (2013). Exponential Power Weibull (EP-W) distribution as a special sub model of this new family is discussed and its some statistical properties has been obtained. Moreover, the maximum likelihood estimators (MLEs) for unknown parameters of EP-W distribution have been derived and a simulation study based on biases and Mean square errors (MSEs) of this estimator for various sample sizes has been performed. Finally, an application with real data set has been presented.

KEYWORDS: Exponential Power - X family of distributions, Exponential Power Weibull distribution, maximum likelihood estimation (MLE).
S153- TOPP-LEONE NORMAL DISTRIBUTION WITH APPLICATION TO INCREASING FAILURE RATE DATA

Vikas Kumar SHARMA

1Institute Of Infrastructure Technology Research And Management (Iitram), Ahmedabad, India.

In this article, we proposed a new three-parameter probability distribution, called Topp-Leone normal, for modelling increasing failure rate data. The distribution is obtained by using Topp-Leone-X family of distributions with normal as a baseline model. The basic properties including moments, quantile function, stochastic ordering and order statistics are derived here. The estimation of unknown parameters is approached by the method of maximum likelihood, least squares, weighted least squares and maximum product spacings. An extensive simulation study is carried out to compare the long-run performance of the estimators. Applicability of the distribution is illustrated by means of three real data analyses over existing distributions.

KEYWORDS: Normal Distribution, Topp-Leone Generated Family, Hazard Rate, Moments\Sep Quantile Function, Orderings, Order Statistics, Maximum Likelihood Estimator, Maximum Product Spacing Estimator, Least Squares Estimator, Weighted Least Squares Estimator
Censored data are frequently encountered in many areas of science such as engineering, medicine, finance, etc. There exists many censoring schemes including Type I and Type II. Another widely used censoring scheme is called as Hybrid which is the mixture of Type I and Type II, see Epstein (1954). In this study, we estimate the parameters of the Azzalini’s skew normal (SN) distribution under hybrid censoring, see Celik (2012) in the context of Type II censoring. The parameter estimates are obtained by using maximum likelihood (ML) methodology via iteratively reweighting algorithm. We also use Tiku’s modified maximum likelihood (MML) methodology which is a noniterative alternative of ML methodology. We compare the performances of the estimators of the parameters by using an extensive Monte Carlo simulation study. At the end of the study, a data set is analyzed for illustrative purposes.

**KEYWORDS:** Hybrid censoring, Skew normal, Maximum likelihood, Iteratively reweighting algorithm, Modified maximum likelihood.
S155- TESTING FOR EQUALITY OF TWO COPULAS BASED ON EXCEEDANCE STATISTICS

Ayşegül EREM

Cyprus International University, Faculty Of Arts And Science, Via Mersin 10, TURKEY.

In two-sample problems, most of the studies are based on univariate random samples. However, in real life application bivariate two sample problems are quite important. In this study, a non parametric test is proposed for testing equality of two copulas. The new test statistic relies on bivariate exceedance statistics. The critical values are provided for some well-known copulas. Then, the power of the test is studied under different alternative hypothesis. At the end, a power study is performed under some hypotheses in R programme.

KEYWORDS: Bivariate exceedance statistics, Copula, Hypothesis tests, Nonparametric tests
In literature, many new families of distributions have been proposed in order to model real data set having to different hazard rates. As it known, real data sets can show fitting to many statistical distributions. It is important to obtain new flexible statistical distributions that can be used for data having to various hazard rates. In this study, a new family of statistical distributions called as “Generalized Cubic rank transmuted Family of Distributions” is introduced. This new family of distributions is a generalization of Cubic rank transmuted distributions. Generalized Cubic rank transmuted Family of Distributions includes transmuted G family of distributions, generalized transmuted G family and Cubic rank transmuted family of distributions. A special model of this new family is Generalized Cubic Rank Transmuted Weibull (GCTW) distribution. Some statistical properties of this distribution are examined in this study. Also the maximum likelihood, least squares and maximum product of spacing estimators of the unknown parameters of this new distribution are obtained. Moreover, a Monte-Carlo simulation study is performed to see the performances of these estimators in terms of mean square errors. One of the advantages of GCTW distribution is that this distribution can be used to model various real data sets having to increasing, decreasing and bathtub shape hazard rates. Kolmogorov Smirnov test to compare GCTW and other selected distributions that fit to a real data set has been used. Consequently, GCTW distribution has to the best fit according to other selected distributions.

**KEYWORDS:** Generalized cubic transmuted family of distributions, Generalized Cubic Transmuted Weibull distribution, Maximum likelihood estimation, Monte-Carlo simulation study
S157- A MODIFIED LINDLEY DISTRIBUTION

Coşkun KUŞ, Mustafa Çağatay KORKMAZ,İsmail KINACI

1Statistics, Selsuk University, KONYA, TURKEY, 2Artvin Çoruh University, Artvin, TURKEY.

In this paper, a new extended version of the ordinary Lindley distribution called modified Lindley distribution is proposed. The distributional properties are studied. Statistical inference is discussed under maximum likelihood framework. The performance of maximum likelihood estimates are observed by Monte Carlo simulation for difference cases. The usefulness of the proposed model is illustrated by using real data set.

KEYWORDS: Modified Lindley distribution; Mixture distribution; Lindley distribution; Power Lindley distribution.
S158- COMPARISON OF THE POWER OF SOME NORMALITY TESTS: A DECISION SUPPORT SYSTEM USING SIMULATION METHOD

Metin ÖNER¹, Yağmur USTA²

¹Manisa Celal Bayar Üniversitesi Uygulamalı Bilimler Yüksekokulu Bankacılık Ve Finans Bölümü Manisa Turkey, ²Manisa Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü Uluslararası Ticaret Ve Finans Manisa Turkey.

Purpose: To introduce the decision support system developed by the MATLAB GUI for getting the power of normality tests in the case of that some popular normality tests used in testing the data set with goodness of fit tests for normal distribution, (Limiting form, Marsaglia et al. Method, Lilliefors Correction, Stephens Modification on Kolmogorov-Smirnov Normality Test, Anderson-Darling, Cramer-vonMises, Shapiro-Wilk, Shapiro Francia, Jarque-Bera, D’Agostino-Pearson), the population in normal, t, uniform, triangle, lognormal, exponential, Weibull, gamma, beta, logistic, Rayleigh, Cauchy and Extreme values of distribution. Method: In the developed decision support system, the user in himself can choose the population size, sample size and distribution pattern of the population (one of the normal, t, uniform, triangular, lognormal, exponential, Weibull, gamma, beta, logistic, Rayleigh, Cauchy and extreme values distributions) and the trial number of simulation. The developed decision support system based on these selections produces the simulation results of the test power of 10 popular normality tests. In addition to that, the graphical support is provided in order to visually evaluate the p-value values of normality tests and the distributions of test statistics. Findings: It has been observed that they produce stronger results than the Shapiro Wilk, Shapiro Francia, Anderson Darling and Cramer-vonMises normality tests with the variation of sorting in different distribution types, different population sizes and different sample sizes while on the other hand the Kolmogorov-Smirnov normality tests produce weaker results.

KEYWORDS: Normality Tests, Simulation, Decision Support Systems, Power of the Test, I. Type Error.
S159- CHANGE POINT DETECTION IN VEHICLE TIME HEADWAYS

Aylin ALIN¹, Almila HACIOĞLU¹, Serhan TANYEL², Ufuk BEYAZTAŞ³

¹Statistics, Dokuz Eylül University, Izmir, Turkey, ²Civil Engineering, Dokuz Eylül University, Izmir, Turkey, ³Statistics, Istanbul Medeniyet University, Izmir, Turkey.

Vehicle time headways are fundamental in many traffic engineering applications. The headway is a measure of the temporal space between two vehicles. Specifically, it indicates the time that elapses between the arrival of the leading vehicle and the following vehicle at the specified point. There may be some points where headway distribution may change. Detecting these points is crucial for determining the headway characteristics. One of the problems related to detecting changes in the data is the outliers. Those points may falsely be flagged as change points. In this study, changes in a headway data will be estimated using AR(p) time series model. Robust measures based on weighted likelihood approach will be proposed to determine if the point is a change point. The proposed method will be applied on the headway data collected from Ankara Street which is one of the most important arterials in Izmir which connects Bornova, Kemalpaşa to the city center (Konak). Observations were made in the morning peak hour with a video camera placed on a pedestrian overpass which was located close to the observation point when the weather conditions were favorable. Headway data was extracted from the recordings in bureau by using a software which was developed at DEU.

KEYWORDS: Change point, Headway, Outlier, Ar(p) model, Weighted Likelihood
Empirical Mode Decomposition (EMD) has been broadly used to analyse nonlinear and nonstationary time series data in a wide range of fields. Due to the adaptive nature of EMD and under the assumption that every data consist of oscillations of intrinsic mode functions; EMD decomposes nonlinear and nonstationary data into series of intrinsic mode functions (IMFs) via the shifting process in order to achieve reliable and credible analysis. Despite its sophistication, mode mixing problem occurs during the decomposition process. The issue of mode mixing is remedied by an adjusted noise assisted algorithm of the original EMD algorithm known as the Ensemble Empirical Mode Decomposition (EEMD). At the end of the process of EEMD, the added noise may not be entirely separated from the signal or data leading to spurious results of the analysis. So as to contain this issue, noise of certain magnitude and amplitude is added to the data to counteract the scales of the IMF of the less dominant component. Therefore, in this study, the main focus will be on the extraction of the less dominant components from the IMF as required. Moreover, the Complete Ensemble Empirical Mode Decomposition with Adaptive noise is applied so as to completely segregate the noise from the signal.

**KEYWORDS:** Empirical Mode Decomposition, intrinsic mode functions, Ensemble Empirical Mode Decomposition, Complete Ensemble Empirical Mode Decomposition.
In this study, it is aimed to compare the performances of Classification and Regression Trees (CART), CART with Bootstrap Aggregating (Bagging) and Random Forests (RF) on seasonal time series forecasting. Time series 28 days (640 hours) long containing daily and weekly periods, a logarithmic increasing trend and a random error with uniform distribution were generated. The models were trained by 21 days long training data and the remaining 7 days long test data was predicted. This procedure was repeated 1000 times. Performances of methods were measured by mean absolute error percentage (MAPE). As a result, the means of MAPE values were 2.66%, 2.58% and 2.32% for CART, CART with bagging and RF respectively. In conclusion, Bagging helps decrease forecasting error for CART. RF has best results among all tested methods, so it is a suitable approach for seasonal time series forecasting.

**KEYWORDS**: Seasonal Time Series, CART, Bagging, Random Forests, Simulation
S162- BOOTSTRAPED FUZZY TIME SERIES MODEL BASED ON PARTICLE SWARM OPTIMIZATION

Ozge CAGCAG YOLCU, Erol EGRİOGLU, Eren BAŞ, Ufuk YOLCU

1Department Of Industrial Engineering, Gireun University, Giresun, Turkey, 2Department Of Statistics, Gireun University, Giresun, Turkey, 3Department Of Econometrics, Gireun University, Giresun, Turkey.

One of the fuzzy logic based models which is taken advantage of by many researchers in time series prediction problems is the fuzzy time series model. Most of the fuzzy time series models available in the literature take the analysis process into consideration as three separate processes. Thus, the model error that will occur in each separate process will generate and increase the total model error. In this respect, the total model error can be reduced by simultaneously evaluating these three separate analysis process in a single process. Moreover, when compared to outputs of probabilistic models, the outputs of these models are more difficult to interpret because it only contains the point estimations. In this study, from this point of view, a fuzzy time series prediction model is proposed to overcome this kind of shortcomings. The proposed fuzzy time series prediction model realizes the whole analysis process in a single process by using particle swarm optimization to determine the fuzzy cluster centers and to train the artificial neural network with single multiplicative neuron which is used the identification of fuzzy relation. And also, it produces more interpretable outputs including confidence intervals can be achieved in just the same way as in the probabilistic inference systems under favour of IID bootstrap technique. The prediction performance of the proposed bootstrapped fuzzy time series prediction model is investigated by some real world time series implementations. And the results are evaluated together with the results of some other models.

KEYWORDS: Time series prediction, fuzzy time series, IID bootstrap, particle swarm optimization
S163- AN INTUITIONISTIC FUZZY TIME SERIES FORECASTING MODEL BASED ON BAT ALGORITHM

Ufuk YOLCU¹, Eren BAŞ², Erol EĞİOĞLU³, Ozge CAGCAG YOLCU³

¹Department Of Econometrics, Gireun University, Giresun, Turkey, ²Department Of Statistics, Gireun University, Giresun, Turkey, ³Department Of Industrial Engineering, Gireun University, Giresun, Turkey.

Time series forecasting (TSF) models aim to find out the time series behavior in future by taking advantage of the information of the current and the past behavior of them. For this purpose, a great variety of forecasting tools have been produced in TSF literature. Especially in recent years, fuzzy logic based approaches, working intensively on them including adaptive network fuzzy inference system, fuzzy functions approach, fuzzy regression and, fuzzy time series models, constitute a family of these forecasting tools. Fuzzy time series approaches introduced in recent studies use membership values as inputs of the forecasting model. Usage of non-membership values, however, can affect in a positive way to the forecasting performance of the model due to by making benefit of more information in time series. Considering this idea, in this study, an intuitionistic fuzzy time series forecasting model (IFTS-FM) is proposed. In the proposed model, membership and non-membership values are obtained by using intuitionistic fuzzy C-means clustering method. Obtained membership and non-membership values are used as inputs in separate two inference systems composed of two separate Pi-Sigma Neural Networks (PS-NN) to determine the fuzzy relations. For both inference systems, real observations of time series are used as target values. The outputs of each PS-NN are combined by utilizing non-membership weight. The training of both PS-NN is realized via Bat optimization algorithm as well as the identification of non-membership weight. The evaluation of the forecasting performance of the proposed IFTS-FM is carried out with various real-world time series implementations.

**KEYWORDS:** Fuzzy time series, intuitionistic fuzzy C-means, Pi-Sigma neural network, bat algorithm, time series forecasting
S164- INTUITIONISTIC HIGH ORDER FUZZY TIME SERIES METHOD BASED ON REGRESSION AND PRINCIPAL COMPONENT ANALYSIS

Cem KOÇAK¹, Erol EGRİOGLU², Eren BAŞ³, Ufuk YOLCU⁴

¹School Of Health, Hitit University, Çorum, Turkey, ²Department Of Statistics, Giresun University, Giresun, Turkey, ³Department Of Econometrics, Giresun University, Giresun, Turkey.

Intuitionistic fuzzy sets are general form of classical fuzzy sets. Hesitation degrees play important role in intuitionistic fuzzy sets. The summation of memberships and non-membership values can be less than one for an intuitionistic fuzzy set. In this study, a new intuitionistic high order fuzzy time series method is proposed based on regression and principal component analysis. The fuzzification is done by using intuitionistic fuzzy c-means algorithm, multiple regression analysis is used to define fuzzy relations. The inputs of the forecasting model are linear functions of lagged memberships, lagged non-memberships and lagged variables. The linear functions are obtained from principal component analysis. Energy consumption time series for some countries has been used to test performance of the proposed method. The performance of proposed method compared with some methods in the literature.

KEYWORDS: Intuitionistic fuzzy sets, forecasting, intuitionistic fuzzy c-means, regression, intuitionistic high order fuzzy time series method
In recent years, Bitcoin has been very popular in finance markets. It was developed anonymously and it’s maintained by a collective group of people who use technology. This is currently a new form of currency so-called cryptocurrency. Normally the currency has been based on gold but on the contrary, Bitcoin is not based on any precious stone like gold, silver etc. Furthermore, Bitcoin has many sub-currency such as Ethereum (ETH) and Ripple (XRP). It is hard to predict all of these cryptocurrency because they have some risks and higher volatility just like other currencies. In this study, ETH Return and USD Exchange Rate Return Series are used by taking logarithm of daily returns (5 day in a week) after 8.8.2015 in order to analyze them. Initially, the stationary of these two series are determined, and optimal autoregressive moving average models for both of them are estimated. Later, the best heteroscedastic variance models among different type of models for each two series are determined. In addition, both of two series are extended by 30 days with time series forecasting which enables a comparison between ETH Return and USD Exchange Rate Return. Comparison of these two series revealed that one of these two return series is more profitable than the other.

**KEYWORDS:** Bitcoin, Ethereum, Time Series Analysis, Heteroscedastic Variance Models, Forecasting
PARAMETER ESTIMATION OF THE AUTOREGRESSIVE MODELS WITH SKEW DISTRIBUTED INNOVATIONS USING EM ALGORITHM

Yetkin TUAC, Olay ARSLAN

Statistics, Ankara University, Ankara, Turkey.

In this study, we consider pth order autoregressive model (AR(p)) with asymmetric innovations. In particular, Azzalini type skew normal and skew t distributions (Azzalini, 1985, 1986) are used to model the innovations. We provide expectation maximization (EM) algorithm (Dempster et al., 1977) to compute the maximum likelihood estimates for the parameters. The performances of the purposed estimators are demonstrated with simulation study and a real data example. The simulation results reveal that the result obtained from skew distributions are superior to the symmetric ones when the skewness is present in the data.

KEYWORDS: Autoregressive stationary process, skew distributions, EM algorithm
S167- THE USE OF DATA MINING AND NEURAL NETWORKS FOR FORECASTING PATIENT VOLUME IN AN EMERGENCY DEPARTMENT

Ceren Öcal TASAR1, Görkem SARıYER2

1Computer Engineering, Yaşar University, İzmir, Turkey, 2business Administration, Yaşar University, İzmir, Turkey.

The ever-growing demand for emergency services has created circumstances such that overcrowding in emergency departments (EDs) is becoming a major threat for public health. Since the main purpose of EDs is to provide timely care to patients, improving the timeliness of emergency care in such an overcrowded environment should be a primary focus from an operational viewpoint. Fortunately, rapid advancement in science and technologies provide opportunities to generate forecasts for patient volume (PV) which enable better plans and consequent improved outcome. This study presents an ED PV forecasting model using data mining and neural networks. Data of this study is obtained from a large-scaled training hospital’s ED that serves on average 800 patients in one day. Since this is a time series data, instead of using demographical variables, we identified time-based input parameters (year, month, day of the week, and period of the day: on duty versus out of duty) for our model to generate numerical forecasts for PV in a determined time index. Based on the interviews accomplished by hospital management, we aim to focus on summer periods. Data for three months of summer 2016 is used as training set, and corresponding data of 2017 is employed as test set. We show that multilayer perceptron neural network model gives ideal prediction accuracy of more than 80% on average, whereas the accuracy of random forest as a data mining methodology is lower. By the use of these techniques, diagnosis based forecasts for the given time index can also be generated.

KEYWORDS: emergency department, patient volume, time series, random forest, multilayer perceptron
S168- BOOTSTRAPPED HOLT METHOD WITH AUTOREGRESSIVE COEFFICIENTS
BASED ON HARMONY SEARCH ALGORITHM

Eren BAŞ1, Erol EGRİOGLU1, Ufuk YOLCU2

1Department Of Statistics, Giresun University, Giresun, Turkey, 2Department Of Econometrics, Giresun University, Giresun, Turkey.

Exponential smoothing methods are classical time series forecasting methods. It is well known that exponential smoothing methods are powerful forecasting methods. In these methods, exponential smoothing parameters are fixed on time and they should be estimated with efficient optimization algorithms. According to time series component, suitable exponential smoothing method should be preferred. Holt method can produce successful forecasting results for time series have trend. Obtaining confidence intervals for forecasting of Holt method and obtaining sampling distribution are easy tasks. In this study, Holt method is modified by using time varying smoothing parameters instead of fixed on time. Smoothing parameters are obtained for each observations from first order autoregressive models. The parameters of the autoregressive models are estimated by using harmony search algorithm. The confidence intervals and sampling distributions for forecasts and model parameters are obtained by using subsampling bootstrap approach. The forecasts of the proposed method are obtained from sampling distributions of forecasts. The real world time series are used to show the forecasting performance of the proposed method.

KEYWORDS: Holt Method, Bootstrap, Harmony Search Algorithm
S169- Statistical Analysis in Assembly Line Balancing: Lean Manufacturing Application in Energy Sector

Burcu Ozcan, Kubra Mutlu, İpek Dolu

Industrial Engineering, Kocaeli University, Kocaeli, Turkey.

The elimination of non-value added activities, maximizing the use of resources, reducing costs, ensuring continuity of these efforts, increasing quality and productivity are essential for firms. The lean manufacturing technique developed for this purpose is based on statistics data. Statistical analysis provides an improvement by presenting the current situation and future situation. The lean manufacturing operation was carried out in the "Qatar Breaker" assembly line, which operates in the energy sector. Circuit breakers are switchgears which are used to cut off the load current and short-circuit currents occur in high voltage and high-current switches. In this study, The Lean Production Activity Plan is conducted. Throughout this plan, process data collection, process analysis, process fluctuation and then standardization with kaizen have been applied in this order. Due to the fact that the use of automation is not widespread in the production system of the factory and manual labour is of first priority, process analyses were made with video records. Then standard operation tables were prepared. New process design was created according to the combination of these tables and sales planning departments data. As a result, assembly line balancing was performed by creating work flow according to this new process design.

KEYWORDS: Lean Manufacturing, Statistical Analysis, Assembly Line Balancing, Improvement
S170- BAYESIAN ESTIMATION FOR INVERSE GAUSSIAN DISTRIBUTION USING DIFFERENT APPROXIMATION METHODS

Ilhan USTA¹, Merve AKDEDE²

¹Department Of Statistics, Anadolu University, Eskisehir, Turkey, ²Department Of Statistics, Usak University, Usak, Turkey.

The Inverse Gaussian distribution is commonly used for modeling lifetime, reliability and survival data, since it is suggested as an alternative to gamma, log-normal and Weibull distributions by many authors. Thus, estimating the unknown parameters of the Inverse Gaussian distribution has become the purpose of the many studies using different approaches. In this study, the Bayes estimators are obtained based on the symmetric and asymmetric (linear exponential and general entropy) loss functions under non-informative Jeffrey’s and informative gamma priors for unknown parameters of inverse Gaussian distribution. Since the Bayes estimators can not be obtained in closed-forms, Lindley and Laplace approximation methods are used to compute the Bayes estimates. The performances of the proposed approximate Bayes estimates are also compared with the corresponding maximum likelihood estimates by using an extensive simulation study.

KEYWORDS: Inverse Gaussian distribution, Lindley approximation, Laplace approximation, Linear exponential loss, General entropy loss
S171- Supervised Term Weighting Schemes for Opinion Spamming

Aytuğ Onan

'Department of Software Engineering, Celal Bayar University, Manisa, Turkey.

The immense quantity of user-generated text information become available on the web. User-generated content includes ideas, comments, reviews, sentiments and opinions about products and services. User generated reviews can be important source of information for individual decision makers and business organizations. Reviews can be utilized to identify problematic characteristics of services and products. In addition, reviews with positive sentiments towards a particular service can be beneficial to organizations in terms of competition in marketing. Due to the power of the reviews on business and marketing, there are fake, deceptive and untruthful reviews. Identifying potentially fake reviews can be modelled as a text classification problem. In order to obtain classification schemes with high predictive performance, text documents should be represented in an efficient way for supervised learning algorithms. Term weighting methods aim to assign appropriate weight values to different terms in terms of their importance on text documents to enhance the predictive performance of text classification. This paper presents the empirical analysis of two unsupervised (term frequency and term frequency–inverse document frequency) and five supervised term weighting schemes (tf-chi square, tf-information gain, tf-gain ratio, tf-odds ratio and tf-relevance frequency) on opinion spamming. In the empirical analysis, four supervised learning methods (namely, Naïve Bayes, k-nearest neighbour algorithm, support vector machines and logistic regression) are utilized. The empirical analysis indicate that supervised term weighting schemes enhance the predictive performance on text classification. The highest predictive performance among the compared schemes is achieved by tf-gain ratio in conjunction with Naïve Bayes.

KEYWORDS: Deceptive spam, machine learning, supervised learning, supervised term weighting, text classification.
S172- Estimation of Censored Regression Model with Maximum Entropy Distributions

Ilhan USTA¹, Yeliz MERT KANTAR¹, Ismail YENILMEZ¹

¹Statistics Department, Anadolu University, Eskisehir, Turkey.

In regression analysis, when the dependent variable is censored, the classical ordinary least squares (OLS) estimator yields biased and inconsistent estimates. In such case, the Tobit estimator is one of the alternative methods. However, since the Tobit estimator depends on the assumption of normality, it produces inefficient results under non-normal distribution cases. In this study, we consider the maximum entropy (MaxEnt) distributions as an error distribution in the censored regression model. It is well-known that the MaxEnt distributions have flexible functional forms and they nest most commonly used statistical distributions under special moment conditions. A simulation study is conducted to show the performance of the estimators based on MaxEnt distributions relative to the OLS and Tobit estimators in the case of non-normal errors.

KEYWORDS: Censored regression model, Maximum entropy distribution, Tobit estimator
S173- Confidence Interval for the Concentration Parameter of Von Mises Distribution

NOR HAFIZAH MOSLIM1, YONG ZULINA ZUBAIRI2, ABDUL GHAPOR HUSSIN3, SITI FATIMAH HASSAN2, NURKHAIRANY AMYRA MOKHTAR3

1INSTITUTE OF GRADUATE STUDIES, UNIVERSITY OF MALAYA, 50603 KUALA LUMPUR, 2CENTRE OF FOUNDATION STUDIES IN SCIENCE, UNIVERSITY OF MALAYA, 50603 KUALA LUMPUR, 3FACULTY OF DEFENCE SCIENCES AND TECHNOLOGY, NATIONAL DEFENCE UNIVERSITY OF MALAYSIA, KEM SG BESI, 57000 KUALA LUMPUR.

The von Mises distribution is the ‘natural’ analogue on the circle of the Normal distribution on the real line and is widely used to describe circular variables. The distribution has two parameters, namely mean direction, μ and concentration parameter, κ. This study considers the calibration bootstrap in constructing a confidence interval for the concentration parameter, κ. The confidence interval based on the calibration bootstrap method was compared with the existing method of obtaining confidence interval based on the asymptotic to the distribution of . Monte Carlo simulation studies were conducted to examine the empirical performance of the confidence interval based on these two methods. Numerical results from the study suggest that confidence intervals based on the calibration bootstrap outperform the confidence interval based on the asymptotic distribution of . As a consequence, confidence interval based on the asymptotic distribution of has a shorter expected length as compared to confidence interval based on calibration bootstrap. As an illustration, daily wind direction data recorded at maximum wind speed for seven stations in Malaysia were considered. From point estimates of the concentration parameter and the respective confidence interval, we note that the method works well for a wide range of κ values. Implication of this study suggests that the method of obtaining the confidence intervals can be applied with ease and provide good estimates.

KEYWORDS: calibration bootstrap, circular variable, concentration parameter, confidence interval, von Mises
POSTER PRESENTATIONS
Combining Logistic Regression and Random Forest Methods for the Assessment of Albuminuria in Diabetes Mellitus Patients

Aslı SUNER, Banu SARAR YÜREKLİ, Timur KÖSE, Mehmet Nurullah ORMAN

1Department Of Biostatistics And Medical Informatics, Faculty Of Medicine, Ege University, İzmir, Turkey, 2Department Of Endocrinology, Faculty Of Medicine, Ege University, İzmir, Turkey.

Albuminuria is the primary predictor for diabetic nephropathy and also associated with the cardiovascular mortality in the patients with diabetes mellitus. Therefore, identifying the parameters related to the presence of albuminuria is of great importance. Herein, we aimed to determine the factors associated with albuminuria in diabetes mellitus patients. We analyzed the data of 129 patients (aged ≥18 years) admitted to the endocrinology department of a public hospital in İzmir from June to September 2013. The univariate logistic regression method was utilized to assess the factors associated with albuminuria in single variables, and then multiple logistic regression method was performed. Twenty-seven variables determined by expert opinion, and 12 of them were significant (p<0.05). Additionally, random forest method was applied to classify the presence of albuminuria by using 12 variables. Classification performance was assessed using a 10-fold cross-validation and evaluated in terms of accuracy, sensitivity, specificity, precision, Kappa statistics, F-measure, area under the ROC curve and Youden index. All comparisons were two-sided, and a p-value<0.05 was considered statistically significant. Statistical analyses were performed with IBM SPSS Version 21.0 and WEKA 3.8.2. The following risk factors including; gender, age, bone morphogenetic protein-4, noggin, lipocalin, creatine, uric acid, total cholesterol, low-density lipoprotein, glomerular filtration rate, presence of microvascular complication, small-vessel occlusion and fibrinogen were statistically significant in logistic regression analysis. The random forest (0.91) achieved higher accuracy than multiple logistic regression (0.90). In conclusion; these risk factors may be used as predictive for albuminuria and nephropathy in diabetes mellitus patients in clinical practice.

KEYWORDS: Logistic regression method, random forest method, diabetes mellitus, albuminuria, endocrinology
Flexible and reliable profile gradients provide important information about a physical system. There are many techniques available to estimate the profile gradient of a noisy system based on finite differencing or smoothing the noisy spatial data with subsequent analytic differentiation, in which both cases claimed to result in huge uncertainties. A new approach in the recent literature to estimate the gradient was to represent the gradient profile directly using the exponential cubic spline to avoid huge uncertainties. In the latter case, the smoothness of the gradient is directly affected by the smoothness of the estimated exponential cubic spline which is determined by the value of tension of the spline. In the literature, there are different heuristic methods implemented to determine the value of the tension of the exponential cubic spline. A new Bayesian approach is presented to infer the tension value of the exponential cubic spline that represents the profile gradient of a noisy system. Different priors; uniform, Jeffreys’ and Gaussian are used for the tension parameter in the Bayesian recipe and tested for the accuracy of the gradient estimates. It is argued that the Jeffreys’ and uniform priors work well when the noise in the system is less, but the gradient estimates go haywire as the noise increases. However, the results show that the Gaussian prior works really well providing reliable estimates for the tension parameter as well as the profile gradient at both lower and higher noise levels.

**KEYWORDS:** Bayesian inference, profile gradient, exponential cubic spline, tension parameter
P3- INVESTIGATION OF SOME INDIVIDUAL FACTORS AFFECTING THE SUCCESS OF MATHEMATICS WITH MULTIPLE CORRESPONDENCE ANALYSIS

Mehmet Tahir HUYUT\textsuperscript{1}, Siddik KESKİN\textsuperscript{2}

\textsuperscript{1}Department Of Medical Services And Techniques, Vocational School, Bitlis Eren University, Bitlis, Turkey; \textsuperscript{2}Department Of Bioistatistic, Medical Faculty, Van Yüzüncü Yıl University, Van, Turkey.

Nowadays, the most important objective of the reform studies on education is to create a system that can help students learn understand mathematics better. In mathematics there are many factors that affect students' achievement averages. In this study, it was aimed to determine the environmental factors affecting the success of Mathematics in the ninth grade of secondary education and the relation between them by Multiple Correspondence Analysis. In the preliminary study, 35 different questionnaires were determined and these questions were applied to 471 students from 7 different school. Multiple correspondence analysis method was used to evaluate the obtained data. In the study, it was found that the students had target consciousness were more successful and who believing that they learned mathematics permanently want to study at university. Also it was emphasized that students who did not pass enough dialogue with their teacher had fear of the note and were more unsuccessful. In addition, it was seen that the students were more successful if the expressions of the teachers were clear. Students who see their teachers as friendly, helpful and humorous are found to be positively associated with the achievement category. The results of the study has shown that multiple correspondence analysis yields satisfactory results to examine of individual factors affecting mathematics success. Therefore, it is thought that the application of multiple correspondence analysis will be important for examine whether it is meaningful or not the changing test results depending on many categorical factors in clinical laboratories or other disciplines.

KEYWORDS: Multiple correspondence analysis, individual factors affecting mathematics success, categorical variables, dimension reduction
P4- AN EMPIRICAL STUDY ON CLASSIFICATION PERFORMANCES OF ARTIFICIAL NEURAL NETWORKS, LOGISTIC REGRESSION, ANFIS, K-NEAREST NEIGHBOR ALGORITHM AND BAGGING CART

İmran KURT ÖMÜRLÜ¹, Fulden CANTAŞ¹, Mevlüt TÜRE¹, Hakan ÖZTÜRK¹

¹Adnan Menderes University Medical Faculty, Biostatistics Department, Aydın/TURKEY.

The goal of this study is to empirically examine classification performances of artificial neural networks (ANNs), logistic regression (LR), adaptive neuro-fuzzy inference system (ANFIS), bagging classification and regression trees (CART) and k-nearest neighbors (KNN) algorithm methods on data sets generated according to varying correlation level and sample size. Two simulation scenarios were created by generating four independent variables which obey normal distribution. In the first simulation scenario, there is no correlation between independent variables whereas there are different levels of correlation, \(0 \leq r < 1\) between independent variables in the second simulation scenario. The sample size was set to \(n=250, 500, 750\) and \(1000\). Sensitivity, specificity, accuracy rates of the models were calculated and the best classifier was chosen according to these criteria. All models except ANFIS shown better performance as sample size increased. Descending order of model accuracy rates was found as ANFIS, ANN, LR, KNN, bagging CART in correlated data whereas it was found as ANFIS, LR, ANN, KNN, bagging CART in uncorrelated data. Besides, the order according to sensitivity and specificity of models was found to be similar to what was found according to accuracy rates.

**KEYWORDS:** Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System, Bagging, Classification and Regression Trees, k-Nearest Neighbors Algorithm
P5- POINT SYMMETRY MODELS IN SQUARE CONTINGENCY TABLES

Gökçen ALTUN¹, Serpil AKTAS ALTUNAY¹

¹Department Of Statistics, Hacettepe University.

Square contingency tables are the cross classification tables where the same variable is measured repeatedly. Point symmetry models in these kinds of tables represent the non-symmetric structure of the tables. In this study, the extensions of point symmetry models are shown for ordinal classification. Conditional point skew symmetry model is discussed for the non-symmetric structure of the table. These models are applied to the cross-classification of the high school final year student’s parents’ educational level in Turkey. We calculate the odds ratios which are the furthest cells to the middle point of the diagonal by using the parameter estimation.

KEYWORDS: square contingency tables, point symmetry models, skew symmetry models
P6- A SENTIMENT ANALYSIS ON THE SOCIAL MEDIA POSTS OF TURKISH FOOTBALL FANS.

Ismail Hakki KINALIOGLU¹, Coşkun KUŞ², Ismail KINACI³

¹Department Of Computer Programming, Distance Education Vocational School, Usak University, Usak, Turkey, ²Department Of Statistics, Faculty Of Science, Selcuk University, Konya, Turkey.

Today, social media platforms with hundreds of millions of users have become tools that enable easy access to sentiment of people. Apart from data leaked by social media bosses, all the information allowed by users is easily accessible. This data is processed for a variety of purposes. User opinions are obtained to make evaluation about elections, commercial product campaigns, or movie, song, book etc. The number of these examples can be increased for various purposes. In this study, the views of soccer fans who uses social media very intensively, were examined. The data are obtained from Twitter, one of the most popular social media platforms. Matches played until the 24th week of 2017-2018 season in Turkish Supper League have been examined. For a match between teams A and B, tweets about both teams were obtained starting from 72 hours before the match time to the match time. Advertising, spam, etc. cleaning have been made on the obtained data. Afterwards, sentiment analysis was performed on the data which organized according to Turkish linguistics structure. Sentiment analysis is considered as a text classification problem and K-Nearest neighbor algorithm is used. Finally, the relationship between the dominant result of sentiment analysis and the result of the match was investigated.

KEYWORDS: sentiment analysis, opinion mining, data mining, text mining, knn algorithm
P7- MORE THAN A DATA: MIGRATION PROBLEM AND ITS IMPACTS IN TURKEY

Ozan EVKAYA¹

¹Atılım University.

Over the last 10 years, Turkey has witnessed the arrival of increasing number of migrants and refugees. Especially, Syrian migrants have devastating impacts from various perspectives. This naive study is about to understand and highlight the migration statistics of Turkey, with some findings on workforce and education. For this reason, this universal problem has been discussed briefly with various indicators. Thereafter, the developing integration policies about the flood of migrants in Turkey has been considered. The simple motivation of this study is to enable an enriched discussion for migration problems of Turkey beyond the migration statistics.

KEYWORDS: migration, refugees, integration policy, economy
P8- THE CIRCULAR MARKET FLOW AS AN APPROACH TO EXPLAIN THE VALUE OF OFFICIAL STATISTICS TO USERS

Florabela CARAUSU¹, Margarita ROHR²

¹DevStat, Valencia, Spain, ²Department Of Applied Economics, University Of Valencia, Spain.

Starting from the idea that official statistics are a public good, the authors argue that ensuring the value of official statistics can tackle positively the challenges that official statistical producers should face. From the legal point of view a public good belongs to or is provided by the State at any level through all bodies being part of the public sector. From the economic standpoint, it is a good that is available to all population and the use of it by one person does not reduce its use by another. In this sense, official statistics are very important for the economic and social development of a country and its significance can be explained by the circular market flow. The main objective of this paper is to offer a new way of understanding the importance of the statistical system for the economic and social development of societies, but as well to promote the continuous and systematic dialogue between statistics producers and users, focused towards communicating the value of official statistics to the society in general. Statistical literacy, the communication of the value of officials statistics to stakeholders, and a closed feedback loop are capable to engage all actors with the process of producing official statistics, ensuring its sustainability.

KEYWORDS: Official Statistics, Value, Public good
P9- CALCULATION OF LOSS PROBABILITY IN THE GI/M/2/K QUEUEING MODEL WITH ORDERED ENTRY

Hanifi Okan IGUDER

1Statistics, Dokuz Eylul University, Izmir, Turkey.

This study is mainly concerned with the finite-capacity queueing system with recurrent input, and two heterogeneous servers represented by GI/M/2/K. Interarrival times are independent and have an arbitrary distribution whereas the service process is Markovian. The service discipline is addressed in a way that customers choose the server with the lowest index number among the empty servers with probability 1. An arriving customer joins the queue when both servers are busy. When the system capacity is full, customers cannot enter the system and leave without taking service. These customers are called 'lost customers' and the flows of lost customers are called 'stream of overflows'. The queueing model GI/M/2/K with ordered entry is analyzed using semi-Markov process. Loss probabilities are computed numerically for the queueing systems where the interarrival times are assumed as exponential, Erlang and deterministic distribution.

KEYWORDS: Loss probability, Heterogeneous servers, Semi-Markov process, Stream of overflow, Ordered entry
P10- ESTIMATION BASED ON PROGRESSIVE FIRST-FAILURE CENSORING FROM BURR XII DISTRIBUTION

İlhan USTA¹, Nihat ÜNAL¹

¹Anadolu University, Faculty Of Science, Department Of Statistics, Eskisehir, Turkey.

In this study, we study the estimation problems for the parameters, reliability and hazard functions of the Burr XII distribution under progressive first-failure censored data. Based on this censoring scheme, the maximum likelihood and bootstrap estimates for the Burr XII distribution are obtained. In addition, approximate confidence intervals and two bootstrap confidence intervals, namely percentile bootstrap and bootstrap-t, for the parameters are derived. Finally, a simulation study is performed to assess the performance of the proposed estimators.

KEYWORDS: Burr XII distribution, Progressive first-failure censored data, Bootstrap, Maximum likelihood, Confidence interval
P11- INVESTIGATION OF DATA FROM VARIOUS AIR QUALITY MONITORING STATIONS IN ANKARA USING QUANTILE REGRESSION ANALYSIS

Nur Efsan TİĞLİ, Şengül CANGÜR

1Department of Biostatistics and Medical Informatics, Faculty of Medicine, Duzce University, Turkey.

With growing population density increasing day by day, the increased use of energy and vehicles in traffic, industrialization and irregular urbanization are the main causes of air-pollution. Five stations (Keçiören, Sıhhiye, Cebeci, Dikmen, Sincan) belonging to different human and geographical characteristics in Ankara were chosen in this study. Between 1 January and 31 January 2017, the data with daily 6-hour intervals published by the Ministry of Environment and Urbanism were used. The temperature, wind speed, wind direction, relative humidity and pressure measurements from the climate elements were taken as independent variables when dependent variables each of PM10, SO2, NO2, CO, O3 was taken as dependent variables from the air-pollution parameters selected according to the air quality criteria proposed by the World Health Organization. Quantile Regression analysis was performed on these variables using the R program. This method is a very useful method for examining the values of lower quantile of air population parameter that have less effect on public health and the values showing excessive air-pollution in the upper quantile values. In this study, the quantile regression equations for QR=0.25, QR=0.50 and QR=0.75 of air-pollution parameters were generated. Among these equations, Akaike Information Criteria (AIC) was used for selecting the most appropriate model. As a result of analyses, it was found that the temperature has significant effect on SO2, NO2, O3, CO air-pollution parameters, and that the wind direction has significant effect on PM10, CO, O3 values and the relative humidity has a significant influence on O3, CO values.

KEYWORDS: Air-pollution, Quantile Regression, Air Quality, Ankara
Emergency departments of hospitals are the most dense units in the health care system. It is stated by the system administrators that this intensity in the emergency services has increased day by day. Simulation methods provide the convenience to identify bottlenecks in such dense units and to evaluate possible scenarios without experiencing them in real life. In this study, it is aimed to simulate the emergency department of a public hospital in Izmir according to patient triage tags based on the patient registry data of March 2017. Appropriate statistical distributions have been determined with the Promodel package program for the periods between visits to the relevant departments and the examination periods of the patients. For the alternative scenario in which the geriatric patient group (aged 65 years and over) was removed, new distributions were found in accordance with the data set and the system was restarted. The results obtained in the comparative studies have led to the conclusion that the establishment of separate emergency services for the geriatric patient group in hospitals and the increase in the use of Home Health Care Services will ease the current system. * Work done while the first author was a graduate student at the Izmir Katip Celebi University under supervision of Assoc. Prof. Dr. Femin Yalçın.

**KEYWORDS:** Geriatric patients; emergency department; triage; simulation
P13- ON A NEW GENERALIZATION FOR INVERTED EXPONENTIAL DISTRIBUTION

Gamze ÖZEL¹, Selen ÇAKMAKYAPAN², Ceren ÜNAL¹

¹Department Of Statistics, Hacettepe University, ²Department Of Statistics, Istanbul Medeniyet University.

In reliability studies commonly used models in life testing include the gamma, lognormal and inverse Gaussian distributions. Also, the inverted exponential distribution is widely used in the field of reliability and was studied by Keller and Kamath (1982) and Duran and Lewis (1989). However, inverted exponential and these distributions have not enough flexibility. In this paper, we proposed a flexible model which present a generalization of inverted exponential distribution using the Lindley distribution. We provide a comprehensive study of some statistical properties of the Lindley inverted exponential distribution. The possible shapes of the corresponding probability density function and hazard function are obtained and graphical demonstration is presented. Quantile function and various relevant measures are studied. The distribution is found to be unimodal. The proposed method of maximum likelihood is used for the estimation of model parameters. The usefulness of the proposed model is illustrated by mean of a real data set.

KEYWORDS: Maximum likelihood, estimation, inverted exponential distribution
Different functions have been used to assess pretest and posttest changes. The most widely used of these functions is the percent change measure. But there is no detailed information on which test statistic should be used for this measure. Although percent change is often said that the distribution is negatively skewed, it can be tested with non-parametric tests but there are also different critics in this regard. So in this study, it was aimed to examine the statistical properties of empirical distribution of four different test statistics for hypothesis testing when using percent change. Six different sample size (n=10 to n=10000) and 3 different correlations (low, medium and high) were drawn from bivariate normal distribution using the written simulation program with 60000 replication. In each sample, the arithmetic mean is divided by its standard error to obtain the first test statistic, trimmed mean is divided by standard error to obtain the second statistic, median is divided by interquartile range (IQR) to obtain the third statistic and trimmed mean divided by IQR for the fourth statistic. As a result of the simulations, it was seen that the 4 different test statistics used to test the percent change did not show t-distribution or even symmetric distribution. These results emphasize that it cannot be tested by the t-test. Percent change values could only be given as a descriptive value in the study, but statistical evaluation would not be appropriate. So, it has been shown that the proposed statistics in the literature on the percent change will not give accurate results because sampling distributions are not examined.

**KEYWORDS:** Pre-post designs, Simple difference, Change measures, Percent change, Simulation
Different functions have been used to assess pretest and posttest changes. One of these functions is the ratio measure. The suitable change measure or relative change functions is selected according to the ease of interpretation and ease of data analyses. It was aimed to examine the statistical properties of empirical distribution of four different test statistics for hypothesis testing when using ratio. Six different sample size (n=10 to n=10000) and 3 different correlations (low, medium and high) were drawn from bivariate normal distribution using the written simulation program with 60000 replication. In each sample, Ratio values were calculated by dividing Post measurements by pre-measures. In each sample, the arithmetic mean is divided by the its standard error to obtain the first test statistic, trimmed mean is divided by standard error to obtain the second statistic, median is divided by interquartile range (IQR) to obtain the third statistic and trimmed mean divided by IQR for the fourth statistic. It was observed that the sampling distribution of 4 test statistic analyzed in the study was positively skewed in small or medium sized samples (n <60). In large samples, it was determined that the distribution of the samples of Median / IQR or TrMean / IQR statistics is very close to the symmetric distribution. The probability making type 1 error will be preserved at 5% level when the critical table values given in Median / IQR and TrMean / IQR statistics are used in cases where the sample size is larger than 60.

**KEYWORDS:** Pre-post designs, Simple difference, Change measures, Ratio, Simulation
climate change has always been one of the most imperious topics regarding water resources. The assessment of meteorological parameters such as precipitation, temperature, wind speed, and relative humidity could be practically useful in risk management, water resource management, and making decisions on climate change. These four variables also have irrefutable effects on hydrological cycle, agriculture, and the environments. These variables are not only related to each other, but the variables are also dependent. In this study, we obtain a relationship between the selective meteorological variables such as pressure, humidity, wind speed and air pollutants including particular matter and sulphur dioxide in which contribute to the climate change in Nevsehir, Turkey. Causality analysis was applied to the variables. The results revealed that humidity is positively related to pressure and wind speed and there is a strong relation between particular matter and pressure, humidity, wind and sulfur dioxide. The instantaneous causality analysis between temperature and other variables concludes that there is one-way causality from the pressure and wind to the temperature; whereas, there are two-way causality between the air pollutants and the temperature.

**KEYWORDS**: Meteorological variables, climate change, time series analysis, causality analysis
A
A. Fırat ÖZDEMİR, 14, 24, 144
Abdelhamid Hamidi ALAOUI, 7, 87
ABDUL GHANOR HUSSEIN, 20, 200
Abdullah YALÇINKAYA, 13, 135
Adem KARATAŞ, 6, 75
Adnan KARAİBRAHİMOĞLU, 6, 71
Ahmet Cagdas ACARA, 21, 213
Ahmet Ergun KARAAGAOGLU, 3, 50
Ahmet SEZER, 7, 82
Alaattin YÜKSEL, 6, 71
Ali ERKOÇ, 9, 103
Ali İhsan BOYACI, 6, 79
Ali Zafer DALAR, 10, 110
Alican BOZYİĞİT, 5, 67
Aysen DENER AKKAYA, 9, 98
Ayşe Övgü KINAY, 11, 118
Ayşe ÖZGÜR GÜLER, 16, 165
B
Bahar YALÇIN, 24
Baki UNAL, 15, 154
Banu Esra ASLANERTİK, 24
Banu SARAR YÜREKLİ, 3, 20, 48, 202
Bariş Keçeci, 9, 104
Bariş Tekin TEZEL, 11, 118
Batuhan ÖZKAN, 19, 192
Bekir CETİNTAV, 14, 147
Berhan ÇOBAN, 24
Berna YAZICI, 7, 24, 82
Beste H. BEYAZTAS, 7, 88
Betül Zehra KARAÇALIK, 8, 91
Beyza DOĞANAY ERDOĞAN, 3, 52
Beyza KIZILKAYA, 5, 70
Birgül ŞENOĞLU, 5, 6, 12, 13, 18, 24, 68, 78, 134, 135, 137, 140, 181
Buğra SARAÇOĞLU, 17, 18, 176, 179, 183
Buğra VAROL, 3, 49
Buket COŞKUN, 17, 172
Burcu FELEKOĞLU, 11, 15, 117, 155
Burcu HUDAVERDİ UCER, 11, 122
Burcu MESTAV, 2, 41
Burcu Ozcan, 19, 196
Burcu UCER, 13, 24, 143
Bülent ALPTEKİN, 2, 9, 10, 40, 105, 108
Büşra SEVİNÇ, 14, 24, 147
C
Cagdas Hakan ALADAĞ, 5, 15, 68, 154
Can ATŞEŞ, 3, 52
Can ATILGAN, 5, 67
Caner TANIS, 17, 18, 176, 183
Cem KADILAR, 14, 22, 24, 148, 217
Cem KOCAK, 19, 191
Cem passwordARIKAN, 8, 90
Cem password ÇELİKOĞLU, 24
Ceren GEGE, 15, 157
Ceren Öcal TAŞAR, 19, 194
Ceren ÜNLÜ, 21, 148, 214
Ceren VARDAR ACAR, 16, 17, 170, 172
Claudio AGOSTINELLI, 24
Coşkun KUŞ, 6, 17, 18, 21, 73, 173, 178, 184, 207
Coşkun PARİM, 19, 192
Çağdaş Hakan ALADAĞ, 2, 9, 10, 24, 40, 105, 108
D
Demet SEZER, 4, 61
Deniz ALPTEKİN, 2, 10, 12, 40, 108, 133
Deniz OZONUR, 6, 7, 77, 80
Derya Eren AKYOL, 16, 166
Dietrich von ROSEN, 25
Dilek VEYSİKARANİ, 16, 165
Duygu KILIC, 7, 80
E
Ebru ÖZGÜR GÜLER, 16, 165
Ebru USTA, 21, 213
Eda KARAŞMAİLOGLU, 3, 50
Edward OMEY, 17, 25, 177
Ekin Can ERKUŞ, 8, 94
Elif ÇOKER, 5, 64
Emel ÇANKAYA, 9, 100
Emily CRAPARO, 12, 129
Emrah ALTUN, 7, 86
Emre KOCABAŞKAN, 2, 41
Enes FILIZ, 10, 111
Engin TAS, 9, 113
Engin YILDIRIZTEPE, 5, 13, 24, 70, 142
Eralp DOGU, 2, 43
Erdal ÇELİK, 24
Eren BAŞ, 10, 19, 11, 0, 189, 190, 191, 195
Erhan ÇENE, 8, 19, 89, 192
Engin YILDIZTEPE, 5, 13, 24, 70, 142
Erhan ÇENE, 8, 19, 89, 192
Erset AVCI, 3, 4, 47, 57
Ezgi AYYILDIZ, 3, 45
Ezgi NAZMAN, 5, 11, 65, 120
Fabrizio DURANTE, 25
Fatih ŞAHİN, 17, 14
Fatih TANK, 25
Fatma GÜRTÜRKÜN, 4, 56
Femin YALCİN, 13, 139
Filiz KARAMAN, 8, 89
Florabela CARAUSU, 21, 209
Francisco CARVALHO, 25
Fulden CANTAŞ, 3, 18, 20, 46, 188, 205
Gamze GÜVEN, 13, 139, 140
Gamze MUSLUOĞLU, 9, 98
Gamze ÖZEL, 21, 22, 214, 217
Georgy L. SHEVLYAKOV, 25
Gordana RADOJEVIĆ, 12, 127
Gökcan ALTUN, 20, 206
Gökhan KARAKÜLAH, 4, 58
Görkem SARIYER, 19, 194
Gözde NAVRUZ, 14, 24, 144
Gurkan OZTURK, 15, 16, 158, 160, 169
Gülay BAŞARIR, 25
Gülden HAKVERDİ, 3, 51
Gülşen AYDIN KESKİN, 7, 8, 83, 85, 97
Gültakin ÖZDEMİR, 10, 112
Güvenç ARSLAN, 6, 76
Sedat ÇAPAR, 24
Selen ÇAKMAKAYAPAN, 21, 214
Selim Orhun SUSAN, 11, 24, 122
Selma GÜRLER, 14, 24, 147
Selma TOKER, 9, 99, 102
Semiha ÖZGÜL, 3, 44
Semra ERBAŞ, 5, 11, 65, 120
Senay ÖZDEMİR, 12, 132
Senem VAHAPLAR, 24
Sercan GÜR, 15, 153
Sercan MADANLAR, 11, 15, 117, 155
Serdar KURT, 25
Seren ÖZMEHMET TAŞAN, 10, 115
Serhan DURAN, 9, 107
Serhan TANYEL, 18, 186
Serkan ERYILMAZ, 25
Serpil AKTAŞ ALTUNAY, 4, 5, 20, 54, 63, 206
Seyit ANKARALI, 21, 22, 215, 216
Sezi ÇEVİK ONAR, 8, 92, 93
Siddik KESKİN, 4, 20, 53, 204
SITI FATIMAH HASSAN, 20, 200
Sibel AL, 14, 149
Simo PUNTANEN, 25
Sinan ÇALIK, 25
Sinan SARACLı, 11, 119
Sukru ACITAS, 5, 6, 68, 78
Süleyman GÜNAY, 12, 133

Ş
Şenay ÇELİKÖRS, 6, 71
Şenay ÇETİN DOĞRUPARMAK, 7, 83, 85
Şengül CANGÜR, 21, 212

T
T. Metin Sezgin, 34
Tahir HANALİOĞLU, 25
Talha ARSLAN, 6, 78
Timothy O'BRIEN, 25
Timur KÖSE, 20, 202
Tolga BERBER, 6, 72
Tolga YAMUT, 13, 143
Tugay KARADAG, 10, 111