PROFESSIONAL STATISTICIANS SOCIETY OF NIGERIA

THE BOOK OF ABSTRACT

5TH PSSN INTERNATIONAL CONFERENCE

THEME: STATISTICS FOR NATIONAL SECURITY AND SOCIO-ECONOMIC DEVELOPMENT

VENUE: ZENITH ICT CENTRE, UNIVERSITY OF NIGERIA, NSUKKA, NIGERIA

JULY, 2021

APPLICATION OF SOME EXPONENTIAL RELATED DISTRIBUTIONS

[1]M. A. Umar ; [2]W. B. Yahya ; [3]J. O. Jimoh

[1] Department of Statistics, Bayero University, Kano, Nigeria

[2] University of Ilorin, Ilorin, Nigeria

[3] School of Preliminary Studies, Science Department, Kogi State Polytechnic, Lokoja

Corresponding Author: wbyahya@unilorin.edu.ng

Exponential family of distribution is a very popular family of distribution functions for analyzing any lifetime data with a lot of applications in different fields of knowledge. This family has distribution functions whose survival, hazard and mean residual life functions are simple and easy to study. As a result, this family has been generalized, modified and mixed with other density functions to give more flexible density functions to facilitate better modeling and analysis. This study is carried out to apply some of these distributions to real life datasets from clinical sciences, remission times of a Bladder cancer dataset. The behaviours of these distributions were illustrated graphically. The parameters of the distributions were estimated using the maximum likelihood method and their goodness-of-fits were examined. The distributions were found to provide satisfactory fits to the datasets considered. Exponentiated Exponential and Exponential-Gamma distributions were found to perform better than all the competing distributions.

Keywords: Exponential family, Gamma, Weibull, Lindley, Hazard rate, Mean residual life, Goodness-of-fit.

CHALLENGES OF GLOBAL INSECURITY IN ACHIEVING SUSTAINABLE DEVELOPMENT GOALS (SDGS)

[1]POPOOLA OSUOLALE PETER ; [2]Apla Abimiku John

[1] MATHS AND STATISTICS DEPARTMENT. THE IBARAPA POLYTECHNIC, ERUWA OYO STATE NIGERIA

[2] Nigeria Institute of Leather and Science, Zaira

Corresponding Author: osuolalepeter@yahoo.com

The sustainable development goals are a call for action by all countries – poor, rich, and middle-income – to promote prosperity while protecting the planet. They recognize that ending poverty must go hand-in-hand with strategies that build economic growth and address a range of social needs including education, health, social protection, and job opportunities while tackling climate change and environmental protection. More important than ever, the goals provide a critical framework for COVID-19 recovery. Security is concerned with how people live and breathe in a society, how freely they exercise their many choices, how much access they have to market, and social opportunities-and whether they live in conflict or in peace. Human security is adversely affected when economic marginalization of the general population is fueled by the rising militancy of local strongmen who have forfeited an economically advantageous patron-client infrastructure due to economic and political liberalization measures. Terrorism as a mega threat affecting the decline of global peace is exacerbated by the lack of concrete international frameworks governing cyber-security and the development of information and communication technology (ICT), which is fully autonomous systems that can select and fire upon targets on their own without any human intervention. Climate change challenges, civil unrest, environmental pollution and now the ravaging power of the Coronavirus pandemic (COVID-19) have added to the security threats of our time. These multi-facets securities challenge no doubt are increasingly becoming serious threats to the global socio-economic and political development of our world and more importantly to the fulfillment of the 17 SDGs before the 2030 targeted date. More so, with the increasing power of globalization, ICT to connect the world, availability of enormous data (big data) around us, and with the rapid changes in the climate, it is more important now than ever to establish the framework for sustainable approaches to mitigate these global security threats if the world wishes to achieve the SDGs. Hence, this research aims at surveying various challenges to achieving SDGs, their causes, and their effects on the world, and provides various frameworks for sustainable approaches to mitigate its effect on the global world.

Keywords: Sustainable Deployment Goals (SDGs), Security Threats, COVID-19, Big Data, Globalization

CHALLENGES OF OFFICIAL STATISTICS IN THE ERA OF COVID-19 IN NIGERIA.

[1]POPOOLA OSUOLALE PETER ; [2]Osuolale Peter Popoola ; [3]Osuolale Peter Popoola

[1] MATHS AND STATISTICS DEPARTMENT. THE IBARAPA POLYTECHNIC, ERUWA OYO STATE NIGERIA

[2] Mathematics and Statistics Department, Adeseun Ogundoyin Polytechnic, Eruwa. Oyo State, Nigeria

[3] Mathematics and Statistics Department, Adeseun Ogundoyin Polytechnic, Eruwa. Oyo State, Nigeria

Corresponding Author: osuolalepeter@yahoo.com

Official statistics are obtained when data are collected and processed into statistical information by a government institution or international organization. In Nigeria, Official Statistics are produced, collated, and disseminated by the federal governments of Nigeria through the Nigeria Bureau of Statistics (NBS). National Statistical Organizations (NSOs) in Africa have not been exempt from the ravages of the COVID-19 pandemic. At a time when the need for high-quality, timely data is more urgent than ever, NBS have been hobbled by restrictions on face-to-face survey interviewing due to social and physical distancing measures. However, the COVID-19 pandemic has brought data to the center of policymaking and public attention. Hence, there is an unprecedented need for data to not only enable decision-makers to inform policies and planning but also minimize the risk for all especially the vulnerable population groups in Nigeria. Data is an important driver to all-around development; Data is needed to bring about development in Nigeria. Nonetheless, governments, civil society, and the general public are increasingly turning to NBS for reliable data to understand the health, economic, and social impacts of the pandemic. This paper examines various method employed by NBS in obtaining official statistics before COVID-19; outline opportunities the new world order has an offer for obtaining official statistics needed for national development; examines challenges in obtaining official statistics in Nigeria and suggest ways NBS could partners with private and public data producers as the nation try to adjust to this 'new normal'.

Keywords: Official Statistics, Nigeria Bureau of Statistics, National Statistical Organization, COVID-19.

AN INVENTORY MODEL FOR DETERIORATING ITEMS WITH THREE-PARAMETER WEIBULL DISTRIBUTIONS UNDER INFLATION.

[1]Odoh Nnamdi Paschal ; [2]Akaninyene Udo Udom

[1] Enugu State University of Science and Technology (ESUT)

[2] University of Nigeria

Corresponding Author: akaninyene.udom@unn.edu.ng

This study deals with a deterministic inventory model for deteriorating items with three-parameter Weibull distribution, under trade credit. The model allows for shortages, which were partially backlogged at rate which varies exponentially with time. The objective of this study is to minimize the total relevant cost by finding the optimal order quantity and optimal replenishment cycle. The model is developed under two different circumstances; case 1: when the customer does not pay the supplier by time M, he will earn interest for the outstanding balance, and case 2: when the customer will be able to sell all his stock and earn interest on the sales proceeds until he has settled the account. Numerical examples and sensitivity analysis are given to illustrate the application and the performance of the model. Therefore, this model will be helpful to the retailer to find the optimal replenishment cycle under various situations and provides a new managerial insight that will help the industry to reduce total relevant cost.

Keywords: "Inventory", "Shortage", "Weibull Distribution", "Trade Credit" , Inflation.

A NEW GENERALIZED DISCRETE DISTRIBUTION

[1]Elebe E. Nwezza ; [2]F. I. Ugwuowo ; [3]Uchenna U. Uwadi

[1] Alex Ekwueme Federal University Ndufu, Alike Ikwo

[2] Department of Statistics, University of Nigeria, Nsukka

[3] Department of Mathematics and Statistics, Alex Ekwueme federal University Ndufu-Alike Ikwo

Corresponding Author: emma_nwezza@yahoo.com

In many areas of survival and reliability studies, data sets of non-negative integers are usually common. There is need for flexible discrete distributions that can take care of over(under)-dispersion and provide adequate fit for reliable statistical inference. In this paper, A new discrete distribution which generalizes the Kumaraswamy-Geometric distribution is introduced. Some of the properties of the new distribution such as the quantile function, moments, moment generating function are derived. Furthermore, the stochastic ordering of the new distribution is considered. The maximum likelihood approach is used to estimate the parameters of the distribution. Comparison with other existing distribution is made using a real-life data set to illustrate the applicability of the new distribution.

Keywords: Discrete distribution, Kumarswamy distribution, Stochastic ordering, Over-dispersion

A STUDY ON SEQUENTIAL PROBABILITY SAMPLING PLAN FOR A TRUNCATED LIFE TESTS CONSIDERING A DESIGNED DOUBLE SAMPLING PLAN

[1]Bello Zoramawa Aminu

[1] Usmanu Danfodiyo University, Sokoto

Corresponding Author: aminu.bello@udusok.edu.ng

This paper proposed a sequential probability sampling plan to a previously designed double sampling plans where the interest was to obtain the minimum sample size necessary to assure that the average life time of a product is longer than the default life time at the specified consumer's and producer's confidence level. Estimations of minimum sample, acceptance and rejection numbers obtained are analyzed and presented to explain the usefulness of sequential plans in relation to single and double sampling plan. Probability of acceptance (Pa), Average sample number and AOQ for the plans are computed. The three regions of acceptance, continue sampling and rejection were determined. The five points necessary to plot ASN curve were also computed and presented.

Keywords: Producer's risk, Consumer's risk, Acceptable Quality Level (AQL), Lot Tolerance Percent Defective (LTPD), Average Sample Number (ASN).

FERTILITY DETERMINANTS AMONG REPRODUCTIVE AGE WOMEN IN NIGERIA: EVIDENCE FROM SOME MODELLING TECHNIQUES.

[1]Mashood Lawal

[1] Air Force Institute of Technology, Nigerian Air Force Base, Mando, Kaduna, Nigeria.

Corresponding Author: maslaw008@gmail.com

This paper presented four parametric count distributions: Poisson (P), Negative binomial (NB), Poisson hurdle (PH) and Negative binomial hurdle (NBH) regression models. Data used was extracted from the 2018 National Demographic and Health Survey. The LRT, Vuong test, rootograms, Akaike Information Criteria (AIC) and Bayesian Information Criteria |(BIC) were used as goodness-of-fit and model selection measures. The objectives of this study were to examine the models for analyzing ideal number of children data exhibiting overdispersion, evaluate their performance and interpret the result of the best model selected that significantly assess some factors contributing to fertility preferences in Nigeria. It was revealed that Poisson-type (P and PH) models were more appropriate in handling of the overdispersion in the ideal number of children data than the NB-type (NB and NBH) models. The result further showed that there was no difference between the PH and NBH models (Z = 0.2435, p = 0.4038). According to both AIC and BIC of the four competing models, it shows that PH model provided a good fit to the ideal number of children data best than the other models (P, NB and NBH). The finding from this study was that mother’s current age, age at first birth, age at first intercourse, place of residence, region of residence except South-West; middle wealth quintile category and Muslim women were found to be significant factors for mothers choosing no child and at least a child as the ideal number of children to have for their whole life in Nigeria.

Keywords: Count data, Overdispersion, NDHS, Fertility, Vuong, and hurdle model.

ON CONSISTENCY OF RANK-SHAPEY VALUE

[1]Eze Chinonso Michael ; [2]Ugwuowo Ifeanyi Fidelis ; [3]Asogwa Oluchukwu Chukwuemeka

[1] University of Nigeria Nsukka

[2] Department of Statistics, University of Nigeria Nsukka

[3] Department of Mathematics & Statistics, Alex Ekwueme Federal University Ndufu Alike

Corresponding Author: chinonso.eze@unn.edu.ng

In this work, the consistency of Rank-Shapley value is justified. The reduced game of Hart and Mascollel (1989) is adopted with the verification that the Rank-Shapley value admits consistency only in a certain class of games. However, a modification of Hamiache's framework is introduced for the establishment of the consistency of Rank-Shapley value in all class of TU-Games. The modification reflects the essence of recognizing a unique proportion for each coalition in sharing the surplus of every possible coalition with isolated coalition (player) in the associated game. In addition, the splitting player/rank property of the value is presented as a property that preserves consistency even when a player is split into many players.

Keywords: Rank, consistency, associated game, Rank-Shapley value, reduced game

A NOTE ON GENE SELECTION OF TISSUE SAMPLES USING GENETIC ALGORITHM

[1]Morolake Oladayo LAWRENCE ; [2]Waheed Babatunde YAHYA ; [3]Rasheed Gbenga Jimoh

[1] Department of Computer Science, Baze University, Abuja

[2] Department of Statistics, University of Ilorin, Ilorin Kwara State

[3] Department of Computer Science, University of Ilorin, Ilorin, Kwara State

Corresponding Author: rolex.oladayo@gmail.com

Cancer is one of the most devastating health conditions in the world. Several studies have shown that early detection of the tumour by clinical methods usually takes a long time. This informs searching for an alternative non-clinical diagnosis of cancer cases using microarray technology. Identification and selection of the most relevant genes have been a cause of challenge to statistician and computational biologists. This is due to the complex nature of high dimensional microarray cancer data. This study proposes a gene selection technique using a filter/wrapper technique. The efficiency of any classification technique depends on the nature of the genes selected by the feature selection technique. T-Statistics was used to reduce the dimension space before the most relevant subset of genes was selected by the Genetic Algorithm. The efficiency of the techniques was evaluated on four publicly available microarray datasets. The result of the analysis shows that the proposed technique can compete favourably with some other gene selection techniques.

Keywords: High-Dimensional, Microarray, Gene Selection, Genetic Algorithm ,Classification

ENHANCEMENT OF RATIO ESTIMATORS OF FINITE POPULATION MEAN

[1]Muili Jamiu Olasunkanmi ; [2]Ahmed Audu ; [3]James Tolu O. ; [4]Ibrahim Yunusa Adamu

[1] Kebbi State University of Science and Technology Aliero

[2] Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria

[3] Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria

[4] Department of Mathematics and Statistics, Federal Polytechnic Nasarawa, Nasarawa State, Nigeria

Corresponding Author: jamiunice@yahoo.com

Hodges-Lehmann measures the median of the pairwise Walsh averages which is robust against outliers. A family of ratio-type estimators of finite population mean has been proposed. The study deals with the improvements on ratio estimators for the estimation of finite population mean of the variable of interest by using known value of hodges-lehmann auxiliary information. The properties of the proposed estimators: bias and mean square error were obtained using Taylor's Series method. Empirical study is established in order to evaluate the performance or merit of the proposed estimators with respect to the conventional and other existing estimators. The results of the empirical study reveals that the proposed estimators are more efficient than existing estimators.

Keywords: Study variable, Mean Square Error, Auxiliary information, Hodges-lehmann estimator, Population mean.

ESTIMATING THE PARAMETERS OF LINEAR REGRESSION MODEL WITH ERRORS DRIVEN BY SHAPE MIXTURES OF SKEW T NORMAL DISTRIBUTION

[1]Nduka Uchenna

[1] Professor J. O. C. Ezeilo Building

Corresponding Author: uchenna.nduka@unn.edu.ng

Multiple linear regression model based on normally distributed and uncorrelated errors is a popular statistical tool with applications in various fields. But this assumption of normality and no serial correlation are hardly met in real life. Hence, this study considers the linear regression time series model for series with outliers and autocorrelated errors. These autocorrelated errors are represented by a covariance-stationary autoregressive process where the independent innovations are driven by shape mixture of skew-t normal distribution. The shape mixture of skew-t normal distribution is a flexible extension of the skew-t normal with an additional shape parameter that controls skewness and kurtosis. With this error model, stochastic modeling of multiple outliers is possible with an adaptive robust maximum likelihood estimation of all the parameters. An Expectation Conditional Maximization Either algorithm is developed to carry out the maximum likelihood estimation. We derive asymptotic standard errors of the estimators through an information-based approximation. The performance of the estimation procedure developed is evaluated through Monte Carlo simulations and real-life data analysis.

Keywords: Autoregressive process; Expectation Maximization algorithm; Linear regression model; Robust estimation; Skew-t distribution

VOLATILITY MODEL SELECTION FOR TIME SERIES COUNT DATA WITH OUTLIERS

[1]Imam Akeyede ; [2]Saleh Ibrahim Musa ; [3]Mustafa Babagana Abubakar ; [4]Ibi G. Fanen

[1] Department of Statistics, Federal University of Lafia, PMB 146, Lafia, Nasarawa State

[2] Department of Statistics, Federal University of Lafia, Nigeria

[3] Department of Statistics, Federal University of Lafia, Nigeria

[4] Department of Statistics, Federal University of Lafia, Nigeria

Corresponding Author: imamakeyede@gmail.com

The specification of appropriate volatility model for capturing the features exhibited by financial series is of significance to policy and economic managers. It is a type of model that captures variation of return volatility in high fluctuating financial markets that exhibit temporal structure in time series. A time series data sometimes contains different data standards from other data and do not reveal the characteristics of a set of data. These data standards are called Outliers. A problem occurs when count data have to be modeled. The number of counts in a certain period can only be an integer and that is why the commonly used ARMA-model, which assumes stationarity, seems not very useful anymore, simply because there are some associated problems, like outlier that can be encountered in the count data, which may actually lead to violation of stationarity assumption of the ARMA model. This problem was addressed by using GARCH and EGARCH model. The orders of the models were studied in respect to the levels of outlier and sample sizes through Mort-Carlo simulation. The results obtained are desirable.

Keywords: Hetroscedastic, Count data, Outlier, Simulation

GAINS IN SUBGROUP CONFIGURATION OF SYSTEM STRUCTURES IN STATISTICAL HUMAN RESOURCE PLANNING AND CONTROL IN MARKOV MODELS

[1]Ossai Everestus

[1] Department of Statistics, University of Nigeria, Nsukka

Corresponding Author: everestus.ossai@unn.edu.ng

Many decades of works in statistical human resource planning and control have witnessed manpower systems modeled as a homogeneous set of employees having internal flows along cadre or length of service only. In some recent developments a manpower system is considered as a union of k subgroups of employees having additional internal flow across subgroups, with each subgroup having the potential for entertaining peculiar membership characteristics and managerial treatments. In this paper the gains in the sub grouping of manpower structures in Markov models and its prospects in handling heterogeneous structures are presented. A case of two classes of subgroups is also highlighted.

Keywords: Statistical manpower control, Markov model, subgroups, transition probability

STATISTICAL ANALYSIS OF NIGERIAN INFANT MORTALITY RATE USING LINDLEY INVERSE EXPONENTIAL DISTRIBUTION

[1]Musa sani Makama ; [2]Hassan Allahde ; [3]Jirayi Joshua magaji ; [4]Umar sahalu

[1] Department of Statistics Shehu Idris collage of health sciences and technology makarfi

[2] Department of Statistics Shehu Idris collage of health sciences and technology makarfi

[3] Department of Statistics Shehu Idris collage of health sciences and technology makarfi

[4] Department of Statistics Shehu Idris collage of health sciences and technology makarfi

Corresponding Author: makamamusa3610@gmail.com

Infant mortality has become more prevalent due to lack of access to health care before, during and after delivery. This contributes to high infant mortality rates both in developing and under-developed countries which is the case in Nigeria. Probability models, their properties and interrelationships are very significant in dealing with naturally occurring phenomena. A large number of distributions have been used over the past decades for modeling data in several fields, however, extending these standard distributions has produced several compound distributions that are more flexible compared to the baseline distributions. In this paper, a Lindley-inverse exponential distribution has been proposed and used to describe infant mortality rate in Nigeria. The properties, estimation of parameters and application of the new distribution are presented and discussed in this paper. An application and investigation of the new model was done using a dataset on the rate of infant mortality in Nigeria and the result was compared with that of other related models.

Keywords: "LINDLEY G FAMILY" ,APPLICATION TO INFANT MORTALITY",LINDLEY INVERSE EXPONENTIAL DISTRIBUTION

ON THE PROPERTIES AND APPLICATION OF THE INVERSE LOMAX-EXPONENTIATED EXPONENTIAL DISTRIBUTION

[1]Falgore Jamilu

[1] Ahmadu bello University Zaria

Corresponding Author: jamiluyf@gmail.com

This article introduces a new extension of Exponential distribution with Inverse Lomax-Exponentiated Family of distributions by Falgore and Doguwa (2020). The new model is called Inverse Lomax-Exponetiated Exponential (IL-EE) distribution. Some of the basic properties of the new distribution such as the quantile function, moment generating function, moments, and order statistics are derived. A Monte Carlo Simulation studies was conducted by setting the initial parameters as (?=0.5, ?=0.6, ?=0.8, and ?=0.9) and with 500 replications. The estimation of the parameters of the IL-EE distribution was done using method of maximum likelihood estimation (MLEs). At the end, IL-EE distribution was applied to Airbone Communication Transceiver data. The application was in favor of the IL-EE distribution as it outperforms the other comparators.

Keywords: Inverse Lomax G; Exponential distribution; Inverse Lomax Exponentiated G; Monte Carlo Simulation; Inverse Lomax-Exponetiated Exponential distribution.

LIKELIHOOD OF INSURANCE COVERAGE ON DAMAGES DUE TO LEVEL OF INSECURITY IN NIGERIA: LOGISTIC MODELLING APPROACH

[1]Ukamaka Cynthia, Orumie ; [2]BARTHOLOMEW, DESMOND CHEKWUBE ; [3]Chukwudi Paul, Obite

[1] Department of Mathematics and Statistics, University of Port Harcourt, Nigeria.

[2] DEPARTMENT OF STATISTICS, FEDERAL UNIVERSITY OF TECHNOLOGY OWERRI, OWERRI, IMO STATE

[3] Department of Statistics, Federal University of Technology Owerri, Nigeria.

Corresponding Author: desmond.bartholomew@futo.edu.ng

Insurance serves as a protection against the unexpected, and it is one of the most effective risk management tools that protect individuals from being bankrupt due to various contingencies. The binary logistic regression model approach was employed to model the described dataset; the model so obtained was statistically significant. All the levels of education were statistically significant in predicting the odds of having insurance cover except for primary education level. Also employment status and age were statistically significant in predicting the likelihood for insurance cover in Nigeria. The results showed that individuals who move from no formal education to obtain Higher education level are 21.66 times more likely to obtain an insurance cover and individuals who move from no formal education to obtain Secondary education level are 2.63 times more likely to obtain an insurance cover. The odd ratio is not significant for moving from no formal education to Primary education and therefore should not be interpreted. Further, individuals who move from being unemployed to being employed are more likely to obtain an insurance cover. Education has the highest impact in predicting the likelihood for one to have an insurance cover in Nigeria. We therefore recommend overhauling of the educational system in order to revamp this sector.

Keywords: Insurance, Risk Management, Binary Logistic Regression Model, Likelihood Test

IMPROVED PRODUCT TYPE ESTIMATOR IN SAMPLE SURVEY

[1]Zakari Yahaya ; [2]Isah Muhammad

[1] Department of Statistics, Ahmadu Bello University, Zaria

[2] Department of Statistics, Binyaminu Usman Polytechnic, 013 Hadejia

Corresponding Author: yahzaksta@gmail.com

In this paper, a modified product type estimator in simple random sampling without replacement by using information on an auxiliary variable is proposed. The properties of the proposed estimator such as bias and mean square error are derived. Theoretically, the efficiency condition of the proposed estimator compared with some of the conventional product estimators are obtained. The proposed estimator and some of the conventional product estimators are applied to real data sets, and the proposed estimator performs better than some of the conventional product estimators considered.

Keywords: "Product estimator, bias, mean square error, auxiliary variable, efficiency"

CENTER POINTS RUNS EFFECTS ON THE PREDICTION VARIANCE OF AUGMENTED ORTHOGONAL UNIFORM COMPOSITE DESIGNS FOR THE THIRD-ORDER MODEL

[1]Yankam Mbouamba Brenda ; [2]Abimibola Oladugba Victoria

[1] University of Nigeria, Nsukka

[2] University of Nigeria, Nsukka

Corresponding Author: brenda.yankam.pg03685@unn.edu.ng

In this paper, the effects of center points runs on the prediction variance of augmented orthogonal uniform composite designs for the third-order model in the spherical region are shown graphically using the variance dispersion graph and the fraction of design space plot. These effects are also evaluated by using some single value criteria such as relative A-, D- and G. From the obtained results, the augmented orthogonal uniform composite designs were found to have a consistent prediction variance in the entire design region. This consistency improved as the center points runs increased from . The relative A-, D- and G-efficiencies proved to increase as the number of center points runs increased hence, center points runs have a positive effect on the prediction variance as well as the relative A-, D- and G-efficiencies of augmented orthogonal uniform composite designs.

Keywords: Prediction variance, variance dispersion graph, fraction of design space plot, A-efficiency, D-efficiency, G-efficiency

PHASE II CONTROL CHART FOR LOCATION WHEN ONE PARAMETER IS ESTIMATED: FURTHER DISCUSSION ON RUN LENGTH DISTRIBUTION AND PERCENTILES

[1]Ugwu Samson ; [2]Akaninyene Udo Udom ; [3]Nnaemeka Martin Eze

[1] University of Nigeria Nsukka, Nsukka Campus

[2] University of Nigeria, Nsukka

[3] University of Nigeria, Nsukka

Corresponding Author: offorma.ugwu@unn.edu.ng

It is well known in the literature that the median run length measures chart performance better than average run length. Some authors already called for more representative measures like the percentiles for assessing chart performance and encouraged examination of the percentiles of the entire run length distribution. Some authors have studied the percentiles of Shewhart x ?-chart when the process mean and variance are known (Case KK) and when the process mean and variance are unknown (Case UU). Our work here considers when only the process mean is unknown (Case UK) and when only the process variance is unknown (Case KU) by evaluating their cumulative distribution function (dfs) via conditioning-unconditioning approach and obtaining their percentiles and plotting their exact run length cdfs for some reference samples (m) each of size 5 at a given false alarm rate. We compare the results with those of the earlier works in the literature. The cdf curves for the case when mean is estimated are stochastically ordered relative to that of the geometric distribution and the dominance is a function of the shifts. However, for the case when variance is estimated, the cdf curves cross that of the geometric distribution at some points. Also, for reference samples of 500 and above each of size 5, the curves for the two cases converge with that of the previous studied cases.

Keywords: Average run length (ARL), False alarm rate (FAR), Median run length (MRL), Stochastic ordering, Phase II control chart.

ON MULTIVARIATE WILCOXON-MANN-WHITNEY TEST

[1]Mbaeyi George ; [2]Chijioke Nweke ; [3]Kelechi Ojide

[1] Alex Ekwueme Federal University Ndufu-Alike, Nigeria

[2] Alex Ekwueme Federal University Ndufu-Alike, Nigeria

[3] Alex Ekwueme Federal University Ndufu-Alike, Nigeria

Corresponding Author: figuresconsults@gmail.com

Multivariate approach for testing independence between pairs of vectors and distributional affinity in a two-sample location problem was examined. In terms of the test statistic, two approaches which differ in their evaluation was studied. Implications of this variation was evaluated using available covid-19 data.

Keywords: Multivariate, independence, location, Wilcoxon-Mann-Whitney

EFFICIENT DISTANCE FUNCTION FOR CALIBRATION ESTIMATORS IN STRATIFIED RANDOM SAMPLING

[1]Muili Jamiu Olasunkanmi ; [2]Ran Vijay Kumar Singh ; [3]Gerald Ikechukwu Onwuka ; [4]Afolabi Wasiu Babayemi ; [5]Ahmed Audu

[1] Kebbi State University of Science and Technology Aliero

[2] Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria

[3] Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria

[4] Department of Mathematics, Kebbi State University of Science and Technology Aliero, Nigeria

[5] Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria

Corresponding Author: jamiunice@yahoo.com

Calibration weights are derived by minimizing distance function which satisfied a set of constraints in line with the auxiliary information. In this study, existing different distance functions are put into consideration and new calibration weights and new calibration estimators of population mean are obtained in stratified random sampling. A simulation study is carried out to ascertain the most efficient among the distance functions. The results reveal that distance function performed better than other distance functions considered.

Keywords: Distance function, Calibration weight, Stratified sampling, Auxiliary information.

CHI – SQUARE – GAMMA {LOG – LOGISTIC} DISTRIBUTION: PROPERTIES AND APPLICATION TO SURVIVAL DATA

[1]OKORAFOR UNEKE ; [2]Eno. E. E. Akarawak ; [3]Hamadu Dallah

[1] YABA COLLEGE OF TECHNOLOGY, YABA, LAGOS, NIGERIA.

[2] University of Lagos Akoka

[3] University of Lagos Akoka

Corresponding Author: uneke.okorafor@yabatech.edu.ng

Akarawak, E. E. E., Okorafor, U.and Hamadu, D. Department of Mathematics1, University of Lagos. Department of Statistics2, Yaba College of Technology, Yaba, Lagos. Department of Actuarial Science and Insurance3, University of Lagos. Correspondence Author: uneke.okorafor@yabatech.edu.ng 08150750279 Abstract The aim of this research is to develop a new continuous probability distribution Chi – Square – Gamma {Log – Logistic} and apply it to survival data. The framework on which this research will be carried out is the T-R{Y} framework. The T-R{Y} distribution obtained has a connection with the hazard function and survival function. The variations that abound in survival data need to be analyzed, as well distribution should be fit on them, as most existing distributions might not be able to capture all the variability that are in these data. The proposed new probability distribution that will be developed in this research work will be able to capture the variability in survival.

Keywords: Key Words: "Chi – square – Gamma distribution", "T-R{Y}", "Cumulative Density Function"

STATISTICAL ANALYSIS OF THE IMPACT OF INFORMATION COMMUNICATION TECHNOLOGY ON POPULATION CENSUS IN IKORODU LOCAL GOVERNMENT AREA, LAGOS STATE.

[1]Are, S. O. ; [2]OKORAFOR UNEKE

[1] Federal Polytechnic, Ilaro, Ogun State Nigeria.

[2] YABA COLLEGE OF TECHNOLOGY, YABA, LAGOS, NIGERIA.

Corresponding Author: uneke.okorafor@yabatech.edu.ng

Are, S. O. and Okorafor, U. Mathematics1 and Statistics Department, Federal Polytechnic, Ilaro, Ogun State, Nigeria. Department of Statistics2, Yaba College Of Technology, Yaba, Lagos, Nigeria. Correspondence Author: uneke.okorafor@yabatech.edu.ng Abstract The concept of information communication technology has-direct link to the people in the nation which brings about easy transfer of information from one person to another. ICT has direct link to population census also along brings about the easy computation of census record in order to be used as references later in the future. Primary data was used. Chi square test was the statistical tool employed. The results show that the participation of persons in the census has a mighty rate of non-participants amongst the respondents than the participants. About 46% of the respondent has participated while 54% of the respondents have not participated in any of the census done in Nigeria. They have a positive effect on the people and the society at large. 95% succeeded to the use of ICT for easy transfer while 50% disagreed to the use of ICT for easy transfer. From the analysis, it was discovered that through the effectiveness of ICT E-Census can be done easily. The following recommendations were made: the government should de-politicize the use of census figure by embarking on nationwide orientations on the importance of census and National Population Commission should retrain their staff on massive application of new technologies such as the ICT devises.

Keywords: Key words: "Statistical Analysis", "Impact", "Information Communication Technology", "Population Census".

PERFORMANCE CRITERIA OF MODIFIED MULTIHALVER TECHNIQUE FOR DETECTING OUTLYING VALUES OF BODY MASS INDEX (BMI): A HIGHER RISK FACTOR AND PROGNOSIS OF COVID – 19 INFECTIONS

[1]USMAN YAHAYA ; [2]USMAN, Yahaya Baba ; [3]ALOYSIUS, Morris T. ; [4]ALIYU, Aminu Ohindase ; [5]IGHO, Victor Gbare

[1] FEDERAL POLYTECHNIC, NASARAWA, NASARAWA STATE, NIGERIA

[2] Mathematics and Statistics department, Federal Polytechnic, Nasarawa, Nasarawa state, Nigeria

[3] Life and Health Science Department, Nutrition and Dietetics, University of Nicosia, Nicosia, Cyprus

[4] Mathematics and Statistics department, Federal Polytechnic, Nasarawa, Nasarawa state, Nigeria

[5] Mathematics and Statistics department, Federal Polytechnic, Nasarawa, Nasarawa state, Nigeria

Corresponding Author: ybusman_2005@yahoo.com

The detection of outlier(s) is often thought of as a means to eliminate such aberrant observation(s) from a set of data to avoid anomalies or further analysis. But outliers can be interesting observations in themselves as they can be leading information to certain abnormal condition(s) just like the case of the recent discovery of obesity being a higher risk factor and prognosis to the novel COVID – 19 pandemic. This study is focused on detecting outlying values of Body Mass Index (BMI) which are higher risk factors and prognosis of COVID – 19 infections using the modified Multihalver technique for detecting outliers. The data of weights in kilogram (kg) and heights in meter (m) was collected from the records of a hospital and the Body Mass Index (BMI) was calculated as weight in kilogram divided by height in meter squared. The distribution of the computed BMI was fitted to the robust criteria of the modified Multihalver technique and the outlying values which are indices of underweight, overweight and obesity were detected as outliers. The study encourages obesity prevention at all ages and recommendation is on appropriate dieting to stay off the possible underlying diseases and in particular the novel COVID – 19 pandemic.

Keywords: COVID – 19 infections, Body Mass Index, Obesity, Modified Multihalver Technique, Outliers

INFORMATION AND COMMUNICATION TECHNOLOGY (ICT) NEEDS AND COMPETENCE LEVEL OF BASIC TECHNOLOGY TEACHER IN ILORIN METROPOLIS

[1]ADETUNJI KAMALDEEN ; [2]ALANAMU TAOHEEDAT ; [3]ADEFILA JUSTINA ENILARA
[1] UNIVERSITY OF ILORIN, ILORIN, NIGERIA.
[2] KWARA STATE COLLEGE OF EDUCATION, ILORIN.
[3] KWARA STATE COLLEGE OF EDUCATION, ILORIN.
Corresponding Author: taoheedatalanamu@yahoo.com

ABSTRACT Information and Communication Technology (ICT) has become an important issue, as an electronic that have come to reshape the world in all aspects of human endeavors with its stronghold in the provision of education for all. However, despite the importance of (ICT) in education, it was argued that majority of teachers do not know how to use the computers to promote educational efficiency, and they are not adequately trained and competence enough in the use of ICT to use modern information media. Hence this study examined Information and Communication Technology (ICT) Needs and Competence level of Basic Technology Teacher in Ilorin Metropolis. A descriptive research design was employed to address the research problem. The target populations for this study are all Basic technology teachers in Kwara State. One hundred teachers was randomly select from public and private schools in Ilorin metropolis as samples. The instrument used for this study was structured questionnaire, data was analyzed using Simple percentage, total mean and T-test ,while statistical packages for social sciences (SPSS) at 0.05 level of significant was used for research hypothesis. The findings of this study revealed that ICT skills have a significant effect on basic technology teachers' effective teaching, the result also proved that there is no significance difference between male and female basic technology teachers in the use of ICT. Based on the findings, the study therefore, recommends that ICT should be Integrate into teaching and learning process, it will bring about a lot of benefit to basic technology 'teacher to student' or 'teacher to teachers' communication, which in turn may translate to greater learning outcome.

Keywords: ICT - INFORMATION COMMUNICATION TECHNOLOGY, BASIC SCIENCE, TEACHER.

CORRECTING MULTICOLLINEARITY DEFICIENCY USING RIDGE AND LASSO REGRESSIONS

[1]Adeshina Ibrahim ; [2]Amobi Abiodun Qasim ; [3]Amusan Ajitoni Simeon
[1] amusan.ajitoni@federalpolyede.edu.ng
[2] Federal Polytechnic Ede Osun State
[3] Federal Polytechnic Ede Osun State
Corresponding Author: adeshinaibrahim10@gmail.com

One of the most important assumptions to consider before the adoption of a multiple regression model is the presence or otherwise of multicollinearity in the predictor variables. A Multicollinearity-free model will yield a reliable result as well as strike balance between the biasness of the estimator and the extent of variation in the predictive power of such model. This study adopts real sector data that are highly correlated to predict the Gross Domestic Product (GDP) of Nigeria over a period of thirty-five years. The Variance Inflation Factor (VIF) from the Ordinary Least Square (OLS) regression shows that five variables out of nine predictor variables are highly correlated, the condition which renders the regression coefficient of the OLS unreliable. Ridge regression (L2) was adopted using a shrinkage value (?) of 3.4 to penalize each of the regression coefficients. The Least Absolute Selection and Shrinkage Operation (LASSO) regression (L1) is employed to select the most significant coefficients to be included in the model. The best model from the LASSO regression indicates that industrial activities, construction, food index, population, and inflation positively affect gross domestic product of Nigeria while Electricity rate has a Negative impact on the GDP. The model is given as: GDP=-0.24indus+0.455constr+0.327food+0.474Pop+0.070Inf-0.290Elect

Keywords: Keywords: Mulitcollinearity, Ordinary Least Square, Ridge regression, LASSO regression.

K-MEANS CLUSTERING METHOD THAT GENERATES INITIAL CLUSTER CENTERS AND ALSO UPDATES CLUSTER CENTROIDS.

[1]OTI ERIC ; [2]Michael O. Olusola ; [3]Wariebi K. Alvan ; [4]Godwin Oberhiri-Orumah

[1] Federal Polytechnic, Ekowe Bayelsa State

[2] Department of Statistics, Nnamdi Azikiwe University, Awka Anambra State

[3] Department of Physics electronics, Federal Polytechnic Ekowe, Bayelsa State

[4] Federal Polytechnic Library, Federal Polytechnic Ekowe, Bayelsa State

Corresponding Author: eluchcollections@gmail.com

In this paper, a new k-means clustering method is proposed which addresses the initial cluster center problem in k-means algorithm based on binary search techniques, and it also updates cluster centers (centroid). In the initialization phase, the initial cluster centers are generated using the modified binary search property approach, while in the updating phase, cluster centroids are updated using an algorithm depending on if a point is added to a cluster or a point is removed from a cluster. It was observed in the simulation and also using real-life data like iris data set and breast cancer Wisconsin (diagnostic) data set which are all multivariate data, the proposed method performed favorably in comparison with the existing methods (Lloyd, MacQueen, Faber, and Astrahan) in terms of accuracy, efficiency and minimization of the within-cluster sum of squares for k clusters.

Keywords: K-Means Clustering, Binary Search, Centroids, Euclidean Distance, Data-points.

PURE JUMP PROCESS OF SECURITY TURNOVER (A CASE STUDY OF NIGERIA STOCK EXCHANGE)

[1]AKINKUNMI WASIU BABATUNDE ; [2]EKEH JOSEPH ERIM ; [3]OKE JAMIU ADEBOWALE

[1] FEDERAL SCHOOL OF STATISTICS, ENUGU

[2] FEDERAL SCHOOL OF STATISTICS, ENUGU

[3] FEDERAL SCHOOL OF STATISTICS, IBADAN

Corresponding Author: takinkunmi3@gmail.com

ABSTRACT The study was carried out to examine existence of jump, jump pattern, average jump in value and volume, and the distribution of the arrival pattern of the jump that exist in the Nigeria Stock Exchange Securities data that include stocks, bounds etc. The study examines the understanding of jumps to be the size in each arrival of the real valued variables which can be either discrete or continuous. The data used were the weekly activities on the Nigerian Stock Exchange trading floor. The trading days on the floor of the NSE is usually five-working days with the exemption of public holidays which made some days to be less than five trading days in a week. It was found out that jumps actually exist as the deals arrive randomly and with different sizes. The arrival pattern showed that the data follows Poisson process while that of the jumps (size) follow Fisher distribution. It was concluded that the Nigerian Stock Securities data can actually be modeled with Compound Poisson since summarily Pure Jump process follows Compound Poisson Distribution which is the combination of Poisson and Fisher distributions

Keywords: PURE, JUMP, PROCESS, SECURITY, TURNOVER

THE NEW REFLECTED MINIMAX DISTRIBUTION ON A BOUNDED DOMAIN: THEORY AND APPLICATION

[1]Falgore Jamilu

[1] Ahmadu bello University Zaria

Corresponding Author: jamiluyf@gmail.com

ABSTRACT In this paper, we introduce a new flexible family of distributions with bounded support, called reflected minimax distribution (RMD), obtained by reflecting Minimax distribution (Jones, 2009) about the y-axis and shifting its units to the right. We proposed standard reflected minimax distribution (SRMD) from this family which includes, as special cases, some distributions on (0, 1) support such as the one-parameter minimax distribution, the power distribution, and uniform distribution. This work is an attempt to partially fill a gap regarding the issue of tractability of continuous distributions with bounded support, which appear to be very useful in many real contexts, including that of finance and reliability. Some properties of the family, including moments, hazard rate and quantile are investigated. Moreover, the maximum likelihood estimators of the parameters are examined and some applications to real data are reported.

Keywords: minimax distribution, reflected minimax distribution, failure rate function, cumulative moments, maximum likelihood estimation.

DUMMY REGRESSION ANALYSIS AS AN ALTERNATIVE TO ONE-WAY ANALYSIS OF VARIANCE

[1]Oke Jamiu Adebowale ; [2]Fatoki Felicia Motunrayo ; [3]Ogunsola Obaseye Amos

[1] Federal School of Statistics Ibadan

[2] Federal School of Statistics, Ibadan

[3] Federal School of Statistics, Ibadan

Corresponding Author: okejamiade@yahoo.com

This study compares the techniques of Dummy Regression Model Analysis with that of one-way Analysis of Variance (ANOVA). The underline function of one-way analysis of variance is to determine significant difference among the means of levels of a single categorical factor as related to the response variable in the ANOVA model. Similar to one-way ANOVA, the Dummy Regression Analysis regress a dependent variable on a categorical independent variable to model a relationship among the two variables. In this study, common estimates were found in the computation of the F-ratio of the two models having subjected the two techniques to real life data. The estimates of the Dummy Regression coefficients explain better than the estimates of ANOVA model parameters.

Keywords: Dummy variable, Dummy regression model, one-way ANOVA model and F-ratio

VECTOR AUTOREGRESSIVE AND LOGISTIC REGRESSION; MODELLING AND FORECASTING WIND SPEED ON RAINFALL, TEMPERATURE AND RELATIVE HUMIDITY AS A SOURCE OF POWER GENERATION, IN NASARAWA STATE, NIGERIA.

[1]AHMED IBRAHIM ; [2]Nweze N. O ; [3]Alhaji Ismaila S

[1] DEPARTMENT OF STATISTICS NASARAWA STATE UNIVERSITY KEFFI

[2] NASARAWA STATE UNIVERSITY KEFFI

[3] NASARAWA STATE UNIVERSITY KEFFI

Corresponding Author: ibrahimloko75@gmail.com

Wind speed modelling and forecasting is pertinent in order to have a power system that is reliable and secured. The Vector Autoregressive model, variance decomposition and the logistic regression model were applied on a 23-year monthly data obtained from the Meteorological Agency Federal Capital Territory Abuja Nigeria. The Vector Autoregressive model shows that some of the meteorological variables namely; Monthly rainfall, Monthly temperature, and Monthly relative humidity, significantly affect Monthly wind speed. It was also found that each of the meteorological variables have varying effects on Wind speed over a future time horizon as depicted by the variable. The Wind speed was also model using the logistic regression where it was found that there is no statistically significant difference between Wind speed and the meteorological variables. To this end, this study posits that wind speed is fairly significant being influenced by some of the meteorological variables. That is to say most of this metrological variable will not be able to generate enough heat in other to influence the wind speed for a sufficient power supply.

Keywords: Modelling, Vector Autoregressive , logistic regression, wind speed and meteorological variables.

ON THE NORMAL APPROXIMATION FOR BINOMIAL DISTRIBUTION

[1]Oke Jamiu Adebowale ; [2]Phillips Samuel Ademola ; [3]Fadayomi Akinwumi Festus

[1] Federal School of Statistics Ibadan

[2] Federal School of Statistics, Ibadan

[3] Federal School of Statistics, Ibadan

Corresponding Author: okejamiade@yahoo.com

This paper revisits the minimum number of trial (n) require for Normal approximation of Binomial Distribution. Thirty (30) iterations of simulated binomial values at different number of trials were subjected to test of one-sample proportion and Chi-square test of goodness of fit for Normal distribution. Proportion of success were examined at different fixed n to check if sample proportion equals population proportion and fit of the empirical distribution for normal. The outcome of the test suggested minimum number of trials of binomial variable for Normal approximation to begin from 15 as fit of Normal became stable at this point upward.

Keywords: Binomial distribution, Normal distribution, one sample proportion test, and test of goodness of fit

ADAPTIVE REGRESSION MODELING OF COUNT DATA: AN APPLICATION TO HEALTH INSURANCE DATA.

[1]Alakija Temitope ; [2]Ajiboye, Y.O ; [3]Okorafor, U. ; [4]Ikegwu, E.M ; [5]Okafor, C. J

[1] Yaba College of Technology, Yaba, Lagos.

[2] Yaba College of Technology, Yaba, Lagos.

[3] Yaba College of Technology, Yaba, Lagos.

[4] Yaba College of Technology, Yaba, Lagos.

[5] University of Lagos, Akoka

Corresponding Author: lawaltope2003@yahoo.com

The need to identify a most suitable model to fit count data is of great interest to statisticians, as the ordinary least square is limited due to its inability to take care of heteroscedasticity inherent in count data. Several models have been developed and one of such recently introduced model is the Discrete Weibull (DW) model. The model is considered able, to handle, under-dispersed and over-dispersed data effectively. This article is concerned with comparing of some of these competing models such as DW, Tweedie regression model, Zero-truncated etc. by applying the models to Health Insurance data. The efficiency of the compared models was ascertained by simulation study. Results from the statistical analysis using the AIC and BIC indicates that Tweedie regression model outperformed DW model, though in general, Tweedie regression, Zero-truncated and DW performed creditably well. Using the best performed model (Tweedie), results further show that the number of encounter of patients is affected by number of diagnosis, class of diagnosis (check-up), age, and sex in that order than whether a patient is on admission or not. These models are therefore suggested to researchers and practitioners for application in fitting models when response variables are count data.

Keywords: Count Data, Discrete Weibull Model, Tweedie Regression, Health Insurance Data, Simulation Study

A COMPARISON OF THE RIDGE REGRESSION AND LEAST ABSOLUTE SHRINKAGE SELECTOR OPERATOR(LASSO) METHODS IN SOLVING COLLINEARITY PROBLEM:

[1]Eze Chinonso Michael ; [2]Chisom Karem Anaeme

[1] University of Nigeria Nsukka

[2] Department of Statistics, University of Nigeria Nsukka

Corresponding Author: chinonso.eze@unn.edu.ng

Prediction accuracies and model interpretability are often inhibited by the presence of multi-collinearity, which is evident among the predictor variables. Such anomalies when discovered are handled by Ridge Regression, Least absolute shrinkage selector operator (LASSO), subset selection method etc. This study compares two known models of overcoming collinearity which are the Ridge Regression and Lasso methods by introducing bias in the estimates. In essence, simulated datasets which are conditioned to exhibit collinearity were used in this research. The Mean Square error and R square adjusted measures were used as benchmarks for comparison. It was observed that Ridge Regression performs better than least absolute shrinkage selector operator (LASSO).

Keywords: Ridge, LASSO, Multicollinearity, ordinary least square method

ON THE EFFECT OF A FIXED COST ON AN INVENTORY MODEL WITH TIME –DEPENDENT PARAMETERS

[1]Nwakobi, Micheal nnamdi

[1] University of Nigeria, Nsukka

Corresponding Author: nnamdi.nwakobi@unn.edu.ng

An economic order quantity model with a time-dependent demand, a fixed cost and a time-dependent holding cost is developed. It provides quantitative insight into a serious practical problem where costs are incurred even when an order is not placed. The effect of a fixed cost on the inventory model is examined. Previous models incorporating time-dependent demand rate assume that the holding cost is constant for the entire inventory cycle. The holding cost is considered as an increasing function of time spent in storage. Differential calculus is used for finding the optimal solution. A numerical example is used to validate the proposed model. Sensitivity analysis is carried out to analyze the effect of changes in the optimal solution with respect to changes in various parameters.

Keywords: Inventory model, Time-dependent demand, Time-dependent holding cost, Fixed cost, Optimization, fixed cost effect

PROBABILISTIC PROGRAMMING AND BAYESIAN INFERENCE FOR THE DYNAMICS OF SPATIAL DEPENDENCE OF SELECTED FOOD PRICES IN NIGERIA

[1]Okonkwo, Stephen Johnson ; [2]Ukoha, Peace Amarachi ; [3]Adedoyin, Ekundayo David ; [4]Adewoye, Adeyinka Ralph

[1] Forestry Research Institute of Nigeria

[2] Forestry Research Institute of Nigeria

[3] Forestry Research Institute of Nigeria

[4] Forestry Research Institute of Nigeria

Corresponding Author: sj.okonkwo@gmail.com

This study investigates the temporal behavioral patterns of the spatial dependence of selected food prices in Nigeria from January 2017 to April 2021. The data was extracted from the national bureau of statistics (NBS) website (https://www.nigerianstat.gov.ng/). Data analyses were carried out on two levels: between states and then between geopolitical zones. The dynamics of spatial dependence were calculated for each month using Moran's I. Augmented Dickey-Fuller (ADF) test was used to test stationarity at a 0.05 confidence level. Data were analyzed using PySAL (a Python library for exploratory spatial analysis) and PyMC3 (a Python probabilistic programming language) to implement Bayesian analysis and inference for time series forecasting.

Keywords: Bayesian Analysis, Forecasting, Probabilistic Programming, Spatial Autocorrelation

STATISTICS OF UNEMPLOYMENT RATE IN NIGERIA

[1]Badmus Idowu ; [2]Faweya Olanrewaju

[1] Department of Mathematics, University of Lagos, Akoka, Nigeria

[2] Department of Statistics, Ekiti State University, Ekiti State, Nigeria

Corresponding Author: idoscomus2011@gmail.com

*Badmus, N. I and Faweya Olanrewaju *Department of Mathematics, University of Lagos, Akoka, Nigeria Department of Statistics, Ekiti State University, Ekiti State, Nigeria Corresponding email address: idoscomus2011@gmail.com Abstract Globally, security and socio-economic variable is a vital instrument in the development of any nation. Lack of adequate security in any nation of the world results in developmental challenges in all areas including: health, education, employment opportunity, economic, social amenities and many more. Both theories and applications of vital statistics are useful for national security and socio-economic development, therefore in this research work, we assess the rate of unemployment and health which constitute part of the instruments for national security and socio-economic development in Nigeria. Data on unemployment from 1999 to 2019 is collected from world bank website for the analysis of the work. We use graphs and charts to fit the data sets. It was revealed from the graphs that unemployment has an upward trend yearly.

Keywords: Charts, Graphs, Variable, Security, Socio-economic, Theories

ON THE MODIFIED AILAMUJIA DISTRIBUTION

[1]OSOWOLE OYEDEJI ; [2]OYEDEJI ISOLA OSOWOE ; [3]FELICIA FATOKI ; [4]RITA NWAKA

[1] DEPARTMENT OF STATISTICS,UNIVERSITY OF IBADAN

[2] UNIVERSITY OF IBADAN

[3] FEDERAL SCHOOL OF STATISTICS, IIBADAN

[4] UNIVERSITY OF DELTA STATE

Corresponding Author: academicprofessor2013@gmail.com

This study considered the modified Ailamujia distribution. Some essential properties of the new distribution were derived to showcase its flexibility and tractability. The result indicates that the Modified Ailamujia distribution is a good addition to the existing family of generalized distributions in literature.

Keywords: Modified Ailamujia distribution, Moment generating function, Cumulant generating function, Hazard function, Order statistics

DEVELOPMENT OF EXPONENTIATED WEIBULL FRECHET DISTRIBUTION

[1]Bukoye Abdulwasiu

[1] Auchi Polytechnic Auchi

Corresponding Author: wasiubukoye@yahoo.com

Bukoye A,1, Igozie M.2, Jimoh M. A.3 Ganiyu A.4 1,2,3,4Department statistics Auchi Polytechnic Auchi) _____ Abstract This article presents a new five-parameter continuous model, called the exponentiated weibull frechet distribution based on exponentiaeted Weibull-G family (Hassan and Elgarhy, 2016). The new model contains some existing distributions as its sub distributions. Various mathematical properties of this distribution were studied. General explicit functions for the expressions for the quantile function, density functions, moments generating function, and other statistics were obtained. The estimation of the model parameters is discussed using maximum likelihood method.The result revealed that the proposed Exponentiated Weibull Frechet Distribution is more efficient in fitting some real data set when compared with Exponential, Weibull, Frechet, Weibull Frechet and Exponential Frechet distributions.

Keywords: Exponentiated, Weibull , Frechet,Distribution, Quantile

ESTIMATION OF POPULATION MEAN IN POSTSTRATIFIED SAMPLING SCHEME IN THE PRESENCE ON NONRESPONSE

[1]Izunobi Chinyeaka ; [2]Onyeka Aloysius Chijioke ; [3]Iwueze Iheanyi Sylvester ; [4]Ogbonna Chukwudi Justin

[1] Federal University of Technology Owerri, Imo State

[2] Federal University of Technology Owerri, Imo State

[3] Federal University of Technology Owerri, Imo State

[4] Federal University of Technology Owerri, Imo State

Corresponding Author: chiyeaka2007@yahoo.com

Abstract: Nonresponse is a major challenge in surveys. The present study focuses on the development of adequate sampling strategies for the handling of the problem of nonresponse under the poststratified sampling scheme. In developing the sampling strategies, the use of auxiliary information is employed, where there is nonresponse on both the study and auxiliary variables. Based on a subsampling design of the nonresponse group, three (3) separate-type estimators of the population mean of the study variable were proposed. These are the customary-type sample mean estimator, the difference-type estimator, and a class of ratio/product-type estimators. Properties of the proposed estimators, including an optimum class of estimators, were obtained both for an achieved sample configuration (conditional argument) and for repeated samples of fixed size n (unconditional argument). Efficiency conditions under which the proposed estimators would perform better than the usual poststratified sample mean estimator were obtained, as well as the best (optimum) estimators among the proposed estimators. The theoretical results were verified using numerical illustration.

Keywords: "Nonresponse," "Poststratification", "Auxiliary information", "Sampling strategies"

ADDRESSING MULTICOLLINEARITY IN REGRESSION MODELS: A CASE STUDY OF NIGERIA ECONOMY.

[1]ORANYE HENRIETTA EBELE ; [2]ARUM KINGSLEY CHINEDU ; [3]UGWUOWO FIDELIS IFEANYI

[1] UNIVERSITY OF NIGERIA NSUKKA

[2] UNIVERSITY OF NIGERIA, NSUKKA

[3] UNIVERSITY OF NIGERIA, NSUKKA

Corresponding Author: henrietta.oranye@unn.edu.ng

Multicollinearity is a common problem in time series data because time series data grows together. The inefficiency of the ordinary least square estimator for the parameter estimation of a linear regression model with multicollinearity problem has led to the development of various shrinkage estimators. In this study, we examine the effect of some economic variable on gross domestic product in Nigeria using of the shrinkage estimators. The results show that unemployment rate has a negative impact on economic growth and it will affect inflation rate and investments in the country.

Keywords: OLS, Ridge,multicollinearity, estimators, GDP

DETERMINATION OF FACTOR(S) INFLUENCING THE ACTIVITIES OF PR INTERVAL ON THE ELECTROCARDIOGRAM (ECG OR EKG) IN AN EXPERIMENT OF THE FACTORIAL TYPE

[1]SAKA ADISA

[1] OBAFEMI AWOLOWO UNIVERSITY, ILE-IFE.

Corresponding Author: sakajamiu@gmail.com

Human heart is a strong muscle that pumps blood to the body. A normal, healthy adult heart is about the size of human clenched fist and it is like an engine that makes a car works, moves and functions properly as such, the heart keeps the body running. A healthy heart supplies the body with just the right amount of blood at the right rate for whatever the body is doing at that time. The flow of blood to the heart could be reduced by plaque build-up or blockage by a plaque suddenly ruptures, these could cause angina (chest pain or discomfort) or a heart attack. When the heart muscle does not get enough oxygen and blood nutrients, the heart muscle cells will die (heart attack) and weaken the heart, diminishing its ability to pump blood to the rest of the body. Remark that, if one is observing some unusual feeling in the heart, chest or in some related areas, this could result to carrying out an Electrocardiogram (ECG) test or examination. In this study, six factors (Sex, Weight, Height, Systolic Blood Pressure (SBP), Diastolic Blood Pressure(DBP) and Heart Rate(HR)) were considered and investigated using factorial experiments to see those main effect(s) and interaction effect(s) that are significant in determining the activities of PR-interval in Electrocardiogram examination. The effect of Sex is the most significant of all the treatment effects considered, followed by the effects of Height and HR, while other factors were shown to have very little effect. Hence, it could be concluded that Sex, Height and HR are the most important factors influencing PR Interval in the ECG examination in order to evaluate the metabolic disorders, effects and side effects of pharmacotherapy, and the evaluation of primary and secondary cardiomyopathic processes, among others.

Keywords: Electrocardiogram, PR-interval, QRS-complex, Heart rate, Blood pressure, Factorial experiments, Experimental runs, Treatment effect.

A METHOD OF NON PARAMETRIC STATISTIC FOR TESTING THE SIGNIFICANT RELATIONSHIP BETWEEN HELICOBACTER PYLORI AND NON STEROIDAL ANTI INFLAMMATORY DRUGS ON PEPTIC ULCER.

[1]MUSA ALIYU

[1] KOGI STATE UNIVERSITY ANYIGBA

Corresponding Author: musaajogealiyu12@gmail.com

Peptic ulcer is a major disorder in the digestive tract which associated with a very high risk of mortality and morbidity. The major causes of peptic ulcer which damage the gastric epithelial cell are helicobacter pylori infection and NSAIDS. The researcher aim is to analyze statistically the suitable statistics tests to analyze the effect of drugs on peptic ulcer patient especially those with helicobacter pylori type, and also to test if there exist significant relationship between the two causes of peptic ulcer. In this work, both four non-parametric test were used to identify which of the best statistical tools suitable for the analysis. A further test were carried out between chi square analysis, Mann Witney, Wilcoxon test and spearman for non- parametric test and the result shows that spearman non parametric test is suitable because it show a very strong relationship between the two causes of peptic ulcer. The spearman test indicate a value of 711.0 for helicobacter pylori positive and 79.0 for non -steroidal anti-inflammatory drugs higher than the other three statistical tools which have a low rate value. The spearman ranking shows a very strong relationship between the two causes of peptic ulcer.

Keywords: Peptic Ulcer, Non-Parametric, Chi Square, Mann Witney, Wilcoxon, Spearman Ranking.

MODELLING AND FORECASTING EXCHANGE RATES VOLATILITY USING ARFIMA-GARCH MODEL BY 1AKINTUNDE M. O., 2ACHA, C. K., OLAWALE, A.O. AND 1MASOPA, A.N.

[1]AKINTUNDE Mutairu

[1] Federal polytechnic, Ede, Osun State

Corresponding Author: waleakintunde2004@gmail.com

Volatility modeling and forecasting is a very important financial Time series research whose importance is enormous and could not be easily quantify as it possesses both theoretical and practical importance to the government, investors, academia and private practitioners (modelers and forecasters) but has received little attention. The present study therefore, serves as a link that bridges the existing gap in stock of knowledge of long memory in exchange rates volatility modelling and forecasting. ARFIMA-GARCH model is applied to the monthly exchange rates data of Federal Republic of Nigeria and Botswana Republic from 2000 to 2019. Literature has established that GARCH model is inadequate for modeling and forecasting financial and economic time series data. The study reveals that Nigeria and Botswana exchange rates are heterosecadastic and fractionally integrated process, Results obtained further revealed the evidence of persistence, mean reverting tendency and asymmetric effects in nature. The results obtained are an invaluable policy documents for the two countries under study and to the well being of the economics of the two countries, it is also an invaluable policy documents needed by private practitioners, academia and policy makers. It should be noted that linear models will be grossly inadequate to solving problems in the direction of long memory.

Keywords: Exchange rate, GARCH, Volatility, ARFIMA-GARCH, long memory, ADF and KPSS

MEASURING THE IMPACT OF LOCATING TERTIARY INSTITUTION ON SOCIO-ECONOMIC DEVELOPMENT OF THE AFFECTED AREA(S) BY 1AKINTUNDE M. O., 1OLAWALE, A.O. 1MASOPA, A.N. AND 1AGBONA, A.A

[1]AKINTUNDE Mutairu

[1] Federal polytechnic, Ede, Osun State

Corresponding Author: waleakintunde2004@gmail.com

The establishment of Tertiary institutions of recent has generated a lot of acrimony and political imbroglio among the Communities in the specific areas they are located. This is because of the benefits attracted in locating tertiary institutions in a particular community. Such benefits include boosting of development economically, socially and productive co-existence. This is typical of Ibadan land where several tertiary institutions are located (University of Ibadan, University College hospital, Technical university, Kola Daisi university the Polytechnic Ibadan, Tower Polytechnic, Ibadan, Ibadan city Polytechnic, Mufutau Lanihun college of Education, Al-Ibadan college of Education, School of Nursing and Hygiene, Eleyele, and School of Nursing (UCH) just to mention the few). The objective of this paper is to evaluate the impact which locating tertiary institutions have brought to socio-economic development of Ibadan land. The use of questionnaire was employed to obtain data used for the study. Statistical Package for Social Sciences (SPSS) was used for the analysis of the data collected from the field. .Results revealed that locating tertiary institution in Ibadan has contributed significantly to improved standard of living of people within the towns they are located and brought people from diverse areas together and also created employment opportunities for harmonious coexistence. In conclusion locating Tertiary institution(s) in any area(s) is a catalyst to the rapid growth, provided the activities of such Tertiary institution are well guided and monitored.

Keywords: Tertiary institutions, impact, community, development, questionnaire and socio-economic.

CALIBRATION ESTIMATION OF POPULATION MEAN IN STRATIFIED SAMPLING USING STANDARD DEVIATION

[1]Babatunde Oluwagbenga ; [2]Oladugba Abimibola

[1] University of Nigeria Nsukka

[2] University of Nigeria, Nsukka

Corresponding Author: babatundegbenge03@gmail.com

Calibration estimation is an approach in sample survey which involves the use of auxiliary variables with the purpose of improving the estimates of the population parameters. A new class of calibration estimators for estimating the population mean in the stratified random sampling using standard deviation of the auxiliary variable is proposed in this paper. The proposed estimators were compared with the estimators proposed in Tracey et al. (2003) and Garg and Pachori (2019) using mean square error and absolute relative bias. Also, the relative efficiency of the proposed estimators with respect to the existing estimators was computed using the percent relative efficiency proposed in Tracey et al. (2003).

Keywords: Calibration, Estimator, Stratified sampling, Standard deviation, Mean square, Bias, Relative efficiency

THE EFFECT OF EXCHANGE RATE FLUCTUATION ON ECONOMIC GROWTH OF NIGERIA

[1]Acha Chigozie

[1] Michael Okpara University of Agriculture, Umudike Abia State.

Corresponding Author: acha.kelechi@mouau.edu.ng

This paper looked at the impact of exchange rate fluctuation on economic growth of Nigeria from 1986 -2019. Exchange rate is the price at which one country exchanges its currency for other currencies. It also signals the competitiveness of a country's exchange power with reference to the rest of the world in a pure market. Besides, it also serves as an anchor which supports sustainable internal and external macroeconomic balances over the medium-to-long term. This paper established that there were fluctuations in exchange rate, gross domestic product, government expenditure, inflation rate, interest rate and money supply in Nigeria from 1986 to 2019. The exchange rate, gross domestic product, government expenditure, inflation rate, interest rate and money supply in Nigeria for the period under study do not have unit root and were stationary at first difference with probability values less than 0.05 level of significance. There is long-run relationship among exchange rate, gross domestic product, government expenditure, inflation rate, interest rate and money supply in Nigeria which means that the variables under study are co-integrated. The error correction (ECT) lagged one period with coefficient of -0.409501 is appropriately significant. Thus, a short run disequilibrium in exchange rate is corrected at a speed approximately 41% per annum. It obvious from the result that exchange rate fluctuation has positive impact on the economic growth of Nigeria.

Keywords: Exchange Rate, Inflation Rate, Interest Rate, Money Supply, Economic Growth

OIL PRICE INSTABILITY AND INDUSTRIAL SECTOR IN NIGERIA

[1]Acha Chigozie ; [2]Akintunde Mutairu Oyewale ; [3]Blessing James Uke

[1] Michael Okpara University of Agriculture, Umudike Abia State.

[2] Department of Statistics, School of Applied Sciences the Federal Polytechnic, Ede, Osun State, Nigeria

[3] Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria.

Corresponding Author: acha.kelechi@mouau.edu.ng

This study provides a useful insight into the impact of oil price instability on industrial sector based on the trend of oil price and macroeconomic variables in Nigeria for the period of 1983 to 2020. The result established that there is long run relationship between oil price and gross domestic product, exchange rate, government expenditure and unemployment rate in Nigeria. The trend analysis showed that there was unstable increase in five variables. The Maximum Eigenvalue statistics which reveals the existence of a long-run relationship between the variables, which implies that they move together in the long run. The vector error correction model revealed that oil price instability of –0.159580 and not statistically significant at 5% level of significance which implies that the model corrects its short run disequilibrium by about 15.95% speed of adjustment in order to return to the long run equilibrium. The long run relationship estimate of the impact of oil price instability in industrial sector shows an indication of a strong relationship between oil price and industrial sector in Nigeria. Oil price had 0.02398, 0.70420, 0.077787 insignificant effect on the gross domestic product (GDP), Exchange rate (EXR), and inflation rate (INF) respectively in Nigeria at 0.05 level of significance. The result also showed that Oil price had -0.548854 significant impact on unemployment rate (UNER) in Nigeria, suggesting that a unit decrease in oil price will influence on unemployment rate (UNER) in Nigeria by 54.88 percent. This implies that the variables under study were all cointegrated. The vector error correction estimate showed that instability in oil price had influence on the industrial sector in Nigeria.

Keywords: Cointegrated, Economic growth, Oil price, Industrial sector, Macroeconomic variables

THE APPLICATION OF PRINCIPAL COMPONENT ANALYSIS (PCA) ON EDUCATION DEVELOPMENT INDEX (EDI) IN A FRAGILE CONTEXT

[1]Salawu Saheed

[1] Air Force Institute of Technology

Corresponding Author: salahus74@gmail.com

Education Development Index (EDI) is the use of a homogenous index to explain heterogeneous indicators as it applies to education. The development index for assessing the investment in education and its outcomes is challenging and seemingly not available for use across the country. The average and the PCA methods were used to build the EDI models using the exogenous variables (access, equity, infrastructure) and the endogenous variable (Outcome). The indicators were normalized for standardization. The PCA was applied, and four components were developed accounting for 73.24 percent of the variation. A correlation matrix was used for the average and the PCA methods, the results indicated that in each of the models, the pair of equity, access, and outcomes are significantly related at a 5 percent level of significance. The stepwise multiple regression model was developed for the average and the PCA model variables, the model indicated a high significance level in the access and equity variable at a 5 percent level of significance. The coefficient of determination for the models is 34.5% and 97.8% for mean and PCA, respectively. This showed that the model developed using the PCA explained most of the variation in the model, with access and equity being the most significant variables.

Keywords: Correlation, Education Development Index (EDI), Indicators, Regression, Normality, Principal Component Analysis.

FORECASTING OF STOCK MARKET RETURNS OF THE NIGERIAN STOCK EXCHANGE: AN ARFIMA MODELING APPROACH

[1]OGUNJOBI EYITAYO OLUWOLE ; [2]ODUSINA, MATTHEW TAIWO

[1] ADESEUN OGUNDOYIN POLYTECHNIC, ERUWA OYO STATE

[2] ADESEUN OGUNDOYIN POLYTECHNIC, ERUWA OYO STATE

Corresponding Author: eyitayo_ogunjobi@yahoo.com

The stock market serves as a veritable tool in the mobilization and allocation of saving among competing uses which are critical to the growth and efficiency of the economy. This paper modeled and forecasted the stock market returns of the Nigerian Stock Exchange using the Autoregressive Fractionally Integrated Moving Average (ARFIMA) p,d,q model. To authenticate the accuracy, adequacy, and predictive efficiency of the model, six criteria based on precision errors were employed: Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Scale Error (MASE), and Mean Error (ME). Having analyzed properly ARIMA (3, 1, 1) and ARIMA (2, 1, 2) model appeared to be better for making accurate forecasting. Another result of the study shows that ARFIMA (4, 0.28.1) was the best model of all the models selected by KPSS

Keywords: Long memory, ARIMA Model, Forecasting, Nigerian Stock Exchange, ARFIMA Model

SIRD STOCHASTIC MARKOV MODELLING OF COVID-19 PANDEMIC IN SOME SELECTED WEST AFRICAN COUNTRIES

[1]Mba Emmanuel ; [2]Ifeoma Christy Mba
[1] Department of Statistics
[2] Department of Economics, University of Nigeria, Nsukka
Corresponding Author: emmanuel.mba@unn.edu.ng

The world is currently besieged by the COVID-19 pandemic. The activities and the global economies are also affected so also are the West African countries. COVID-19 is in the family of the Coronaviruses. Coronavirus is one of the main pathogens that fundamentally aims at the human respiratory system. This study, therefore, examined the rate of COVID-19 transmission in some selected West African countries via the Stochastic Markov analysis by using the Susceptible, Infective, Recovered, and Death (SIRD) states. The stochastic Markov model was applied to the WHO COVID-19 data and it was discovered that the transition probabilities of the SIRD state's series were 0.61 (61%) for Liberia, that is, from a susceptible state to an infective state, 0.53 (53%) for The Gambia and 0.45 (45%) for Nigeria respectively. The findings of the study show that the recovery rate was higher with The Gambia (53%) and lower with Liberia (37%). The infective state to death (absorbing) state had Nigeria top the list with 44%, and The Gambia with 7% respectively. The findings in the study capitalize more on the need to adhere strictly to the COVID-19 laid down rules so as to stay safe.

Keywords: COVID-19, WHO, SIRD, Pandemic, Nigeria, Liberia, The Gambia, West Africa

AN EXACT TEST FOR OUTLIERS IN LINEAR REGRESSION MODELS

[1]Ugah Tobias Ejiofor
[1] DEPARTMENT OF STATISTICS UNIVERSITY OF NIGERIA NSUKKA
Corresponding Author: tobias.ugah@unn.edu.ng

A test statistic involving the maximum absolute internally studentized residual statistic for detecting a single outlier in linear regression models is considered. The maximum absolute internally studentized residual is a basic test statistic for a single outlier detection in a least squares analysis based on a general linear model. Exact critical values of this test statistic are hard to obtain because the associated distribution of this test statistic is complex. Published critical values of this test statistic are based on either the Bonferroni upper bound or scale simulations. In this paper, the distribution of this test statistic is derived and its exact critical are obtained. The critical values of this statistic obtained herein compare favorably well with approximate critical values obtained by other authors via the Bonferroni inequality. The closeness of the exact critical values to the approximate values seem attests to the accuracy and goodness of the first-order Bonferroni upper bound.

Keywords: Bonferroni inequality, Critical values, regression diagnostics, Hat matrix, Outlier, Statistic, Studentized residua.

COMPUTERIZED PREMIUM QUOTATION SYSTEM FOR INSURANCE COMPANY

[1]IZUCHUKWU UZO ; [2]EZUGWU ASSUMPTA
[1] UNIVERSITY OF NIGERIA NSUKKA
[2] UNIVERSITY OF NIGERIA NSUKKA
Corresponding Author: izuchukwu.uzo@unn.edu.ng

Property theft often has adverse effect on insurance companies cum the insured. Over the last decade, Nigeria has witnessed several cases of theft especially car theft in various region of the country, this in turn has led to several claims by clients over their insured properties. One cannot over emphasize the impact of these claims on the sustainability of these companies when inadequate computerized solution is not in place to address all the necessary features needed to be considered prior to premium quotation. This research work will try to simulate these features and provide a guide to car insurance company in Nigeria before making quotation for premium. The Object Oriented Analysis and Design (OOAD) methodology is used for designing and implementing the computerized premium quotation system for insurance company. The software application is modeled using standard notations and techniques of the Unified Modeling Language (UML).

Keywords: insurance, premium, quotation, theft, simulation

ITEM RESPONSE THEORY MODELS: ON PROPENSITY OF ENDORSING ITEM-OPTIONS IN MULTI-CATEGORY TEST ITEMS AND RESPONSE CATEGORIES ANALYSIS

[1]ADETUTU OLAYIWOLA ; [2]H. B. Lawal

[1] DEPARTMENT OF STATISTICS, FEDERAL UNIVERSITY OF TECHNOLOGY MINNA

[2] Kwara State University Malete, Ilorin, Nigeria

Corresponding Author: adetutuolayiwola@gmail.com

The use of unordered multi-category responses have been on increase virtually in our Nigeria universities since the advent of COVID-19 to solve problems encountered by increasing number of examinations candidates without sacrificing the validity of test items. In most cases, lopsided test item(s), and defectives item-options (items' distractors) in a multiple choice items have been endangered the quality, integrity, and awards of certificates. A 35 multiple choice questions in a compulsory undergraduate statistics course where options A, B, C, and D are coded as 1, 2, 3, and 4 respectively were used to illustrate this framework. A nominal item response theory model is focused here to investigate and analyse the propensity of endorsing each of the item-options as a function of the attribute being measured, often described as a category probability curve for each of the four option responses. Each item-option curve shows how candidates perceived each of the item-options thereby helped in identified poor, marginal and defectives category options. The analysis was done using Stata 16SE on window 7.

Keywords: item-options, responses, category, unordered, propensity

STATISTICAL MODELLING OF TUBERCULOSIS DIAGNOSTICS

[1]OGUNJOBI EYITAYO OLUWOLE ; [2]ODUSINA, MATTHEW TAIWO

[1] ADESEUN OGUNDOYIN POLYTECHNIC, ERUWA OYO STATE

[2] ADESEUN OGUNDOYIN POLYTECHNIC, ERUWA OYO STATE

Corresponding Author: eyitayo_ogunjobi@yahoo.com

Tuberculosis (TB) remains one of the top ten causes of death worldwide; presently great progress has been made around the world in the prevention and control of Tuberculosis. Thus, this study aims to determine the total number of tuberculosis cases and the mean differences in patients affected with given data of ten years. T-Test was used to determine the mean difference between alive and dead patients with Tuberculosis (TB) and ARIMA was used to forecast the survival of patients. From the analysis, it was observed that males have a greater risk of Tuberculosis than the female.

Keywords: Survival, Tuberculosis, ARIMA Model, Patients, Alive

FITTING BINARY LOGISTIC REGRESSION MODEL ON BMI STATUS OF YOUTHS

[1]BELLO ANDREW

[1] AUCHI POLYTECHNIC, AUCHI

Corresponding Author: belloao2002@yahoo.com

1 Bello Andrew Ojutomori (belloao2002@yahoo.com) 2 Momoh Beshiru (bashmoh2002ng@yahoo.com) 1&2 Department of Statistics, School of Information and Communication Technology, Auchi Polytechnic, Auchi An Abstract submitted to Professional Statisticians Society of Nigeria (PSSN) 5th International Virtual Conference. UNN 2021 holding at Zenith ICT Centre, University of Nigeria, Nsukka. On 5th – 8th July 2021. ABSTRACT Body Mass Index (BMI) is a standard measurement used to classify body weight, it gives an indication of health status of an individual. Overweigh and obesity is conditions of excess body fat accumulation which might lead to an adverse effect on healthy living. In this study BMI status was classified into two (dichotomous) - Normal weight and Non-normal weight. Logistic regression is useful for examining situations in which predicting the presence or absence of a characteristic based on values of a set of predictor variables. The dependent variable is BMI status. The independent variables considered were age, weight and height of the observations understudy. This paper aimed at fitting a Binary Logistic Regression model examined normal weight and non-normal weight among youths with respect to age, weight and height factors and to determine the odds of youths been normal weight. Primary source of data collection was used to gather data on age, weight and height from one hundred and thirty two (132) volunteer youths; their BMI status was estimated and classified into two categories. Maximum Likelihood method was used to fit the logistic regression, automatic selection technique was used to select the predictor variables, Wald test used to establish significance of each predictor variables. Hosmer and Lemshow test used to determine the goodness-of-fit of the model.

Keywords: Normal-weight, Non-normal weight, BMI Status, Predictor, Binary Logistic Regression

PREDICTIVE MODEL OF ARRIVAL AND SERVICE RATE OF PATIENTS AT HEALTH CARE DELIVERY CENTRE IN NIGERIA

[1]Abdullahi Shitu Umar ; [2]Ayuba Godwin Aniah ; [3]Aliyu Nuhu ; [4]Hassan Idris ; [5]Kassim Abdulganiyu

[1] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi, Kaduna State

[2] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi, Kaduna State

[3] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi, Kaduna State

[4] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi, Kaduna State

[5] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi, Kaduna State

Corresponding Author: umarabdullahishitu@gmail.com

In this research, the average inter-arrival time (11 seconds) and service time (10 seconds) of patients at health care delivery centre examine by W.B Yahaya etal, PSSN, 3rd international conference, Zaria, 2019 have been used as information required for modeling/predicting the future values of Arrival and service rates to proffer solutions to enhance the management of the aforesaid Health care delivery centre. Chi-squire distribution was used to test the goodness of the fitted model/predictive models and the result showed that both arrivals and service rate follow Gamma distribution while the fraction of time in the interval of time untill arrival follow Beta distribution; t-distribution was used to investigate if there is significance difference between the fraction of time until arrival arrives and arrival arrives. The researcher used the models in predicting the arrivals and service rate for 1(one) minute and found that the number is 5 patients after one minute and the service rate is 6 patients per minute, also the average sum of arrival rate and service rate is 11 patients per minute, proportion of time to event is 0.4 which shows that there is no significant difference in fraction of time.

Keywords: "Gamma", "Beta", "Chi-square distribution ", "Arrival rate ", "Service rate ".

SPATIAL COMPARISON OF GEOGRAPHICALLY WEIGHTED REGRESSION AND KRIGING METHODS FOR ESTIMATING THE AVERAGE RAINFALL IN SOME SELECTED TOWNS IN THE NORTH WEST OF NIGERIA.

[1]Abdulkadir Yasir ; [2]Umar Usman

[1] Umar Musa Yaradua University Katsina

[2] Usmanu Danfodiyo University Sokoto

Corresponding Author: uusman07@gmail.com

Rainfall data presents a vital input for many hydrological investigation, as such the study seeks to analyze the average rainfall data for the period of one decade obtained from NIMET Kaduna for 30 selected towns in the North western part of Nigeria. The aim is to compare the efficiency of Geo-statistic and Regression model to interpolate the neighboring locations from our obtained data and determine the best out performing model in the study area base on the validation methods of assessing the best interpolator use in the research, i.e. Method of RMSE and the Coefficient of determination R2. The supportive variables which have been used to predicting neighboring rainfall in the in the region are; Humidity, Temperature, Wind Speed, and Pressure.

Keywords: "Geographically Weighted regression", "Universal Kriging", "Rainfall ".

GOAL OPTIMIZATION OF A BANK'S INVESTMENT.

[1]EZRA PRECIOUS

[1] UNIVERSITY OF NIGERIA, NSUKKA

Corresponding Author: precious.ezra@unn.edu.ng

Abstract Goal programming is an extension of linear programming in which the decision maker's objectives are treated as goals or targets that must be attained as closely as possible within the given problem constraints. Most goal programming problems are with multiple objectives and the objective functions are not maximized or minimized directly; rather, the unwanted deviations from the set goals are minimized. These unwanted deviations could be positive or negative, denoted by d+ and d- respectively. The objective of this study is to develop a goal programming model that could be used to optimize a bank's investment, indicating whether the decision makers' set goals are exactly achieved, underachieved or overachieved. The three goals considered here are: maximization of profits, minimization of capital-adequacy risks and minimization of illiquid asset risk. Three methods of solutions were discussed and applied, namely: The weighted sum approach, the e-constraint approach and the merged (weighted –sum/e-constraint) approach. Data were collected on the bank's investment and analyzed by applying the simplex algorithm of linear programming on the three methods of solutions and the result showed that only the e-constraint approach minimized all the deviational variables to zero, implying exact achievement of the bank's investment goals. .

Keywords: Keywords: "Goal optimization", "deviational variables", "simplex method", "linear programming", "Investment risk".

A NEW TEST FOR MULTIVARIATE NORMALITY BASED ON MOMENT GENERATING FUNCTION OF PRINCIPAL COMPONENTS

[1]Mbanefo S. Madukaife ; [2]Uchenna C. Nduka ; [3]Everestus O. Ossai ; [4]Tobias E. Ugah

[1] Department of Statistics, University of Nigeria, Nsukka

[2] Department of Statistics, University of Nigeria, Nsukka

[3] Department of Statistics, University of Nigeria, Nsukka

[4] Department of Statistics, University of Nigeria, Nsukka

Corresponding Author: mbanefo.madukaife@unn.edu.ng

In this paper, we propose an alternative generalization of a recent test for univariate normality which is based on the empirical moment generating function to the multivariate case. We show, among other properties, that the proposed weighted L2-class of statistics is affine invariant and consistent. The empirical critical values of the proposed test are evaluated for different sample sizes, variable dimensions and values of the smoothing parameter through large scale simulations. The empirical power comparison of the test with a strong competitor shows that the test has a considerable high power performance, especially at large sample sizes as well as under heavy-tailed alternative distributions.

Keywords: Moment generating function, principal components, multivariate normality, empirical critical value, weighted L2 – statistic

THE IMPLICATION OF CASE AND INFECTIOUS FATALITY RATES OF COVID-19 WITH SOME MODELLING APPROACHES

[1]Oyamakin Oluwafemi

[1] Room 005, Department of Statistics, University of Ibadan, Nigeria

Corresponding Author: fm_oyamakin@yahoo.com

Coronavirus disease 2019 (COVID-19) is the latest infectious disease to develop rapidly worldwide. The public opinion within central Nigeria was that COVID-19 is a "big man disease" and most infected people show symptoms within five to six days. However, infected patients can also be asymptomatic. Two measures used to assess the proportion of infected individuals with fatal outcomes were modelled in this study. The first is infection fatality ratio (IFR), which estimates this proportion of deaths among all infected individuals. The second is case fatality ratio (CFR), which estimates this proportion of deaths among identified confirmed cases. Key mortality indicators, such as crude death rate, case fatality rate and infection fatality rate and the leading causes of death in a population were considered as a measure for population health assessment during the pandemic. COVID-19 mortality measures based on the daily reported data were modelled using probability distributions such as Weibull, Lognormal, Exponential and Gamma. Case Fatality Rate (CFR), Infectious Fatality Rate (IFR) and Crude Death Rate (CDR) for the Nigeria States and all the six continents of the world were also computed and some inferences were made

Keywords: COVID-19, Case Fatality Rates, Infectious Fatality Rates, Crude Death Rates.

TIME SERIES ANALYSIS ON ANNUAL PRODUCTION OF CRUDE OIL IN NIGERIA: A BILINEAR AND VAR MODEL APPROACHES

[1]Oyenuga Iyabode ; [2]Agbona Anthony Adisa ; [3]Agboluaje, Saheed Abidemi ; [4]Olayiwola Barakat Dolapo

[1] Department of Mathematics and Statistics, The Polytechnic, Ibadan

[2] Federal Polytechnic, Ede

[3] The Polytechnic, Ibadan

[4] The Polytechnic, Ibadan

Corresponding Author: ioyenuga2007@polyibadan.edu.ng

Crude oil is a major source of energy in Nigeria and the world in general. Oil being the mainstay of the Nigeria economy plays a vital role in shaping the economy of the country. This research explores the monthly production of the crude oil per barrel in Nigeria from 2000 to 2019 and was extracted from the Central Bank of Nigeria Bulletin. The comparison was made between the two Time Series models namely: Bilinear and VAR models in terms of their performances. Forecast was made with the two models for the next five years. The result showed that Bilinear model performed better than the VAR model.

Keywords: Bilinear, VAR model, Crude oil, Time Series

ANALYSIS OF CLUSTERED DATA USING DUAL LINEAR MIXED MODEL

[1]Bello Yusuf ; [2]Bayo Lawal

[1] Federal University Dutsin-Ma

[2] Kwara State University, Malete

Corresponding Author: ybello@fudutsinma.edu.ng

Clustered data are a commonly occurring phenomenon. The presence of clustering brought a dependency of the observations within clusters, such that if we ignore the dependency of the observations within clusters the estimated standard errors of the regression coefficients will be biased downward, leading to too narrow estimated confidence intervals and smaller p-values. Linear mixed models are often used to account for the clustering effects. In this paper we proposed a 'dual linear mixed models' that restructure the initial clusters to surface some new latent clusters that maximize the dependency of the observations within clusters, the goodness-of-fit, the estimated standard errors of the regression coefficients, the confidence intervals and the p-values for good parameter estimation and valid statistical inference.

Keywords: Clustered data, Estimated standard errors, Dual linear mixed models, Goodness-of-fit

STATISTICAL ANALYSIS ON COVID – 19 PANDEMIC CONFIRMED CASES, RECOVERIES AND FATALITIES IN NIGERIA FROM MARCH - DECEMBER, 2020

[1]Ikegwu Emmanuel ; [2]Ayodeji Adekunle A. ; [3]Etisioro, Caroline O. ; [4]Muhammed, Idris A.

[1] Yaba College of Technology, Yaba Lagos

[2] Yaba College of Technology, Yaba Lagos

[3] Yaba College of Technology, Yaba Lagos

[4] Yaba College of Technology, Yaba Lagos

Corresponding Author: macioe2002@yahoo.com

The study investigated the distributions of confirmed, recoveries and fatality cases of Covid - 19 pandemic in Nigeria in 2020. The distributions were compared over the months to identify the patterns and differences observed. The data used for the study were daily reports obtained from the National Centre for Diseases Control (NCDC) from March – December, 2020 and was analysed using Analysis of Variance and Jonckheere - Terpstra test with the aid of SPSS 26. The results show that December has the highest mean confirmed cases (663) and recoveries (466.3) while March has the least (confirmed, 7 and recoveries, 1), June has the highest mean fatalities and March the least. The Anova test showed that there was a statistically significant difference in mean of confirmed, recoveries and fatality cases over the months ($p < 0.05$). Also, a Jonckheere-Terpstra test for ordered alternatives showed that there was a statistically significant trend of higher median of confirmed, recoveries and fatality cases from March to December ($p < 0.05$). The study concluded that there was overall increase in the confirmed, recoveries and fatality cases of Covid – 19 pandemics in Nigeria over the period studied.

Keywords: Pandemics, fatalities, Jonckheere-Terpstra test, confirmed cases, Median

METHODS OF ESTIMATING VARIANCE COMPONENTS IN A COMPLETE RANDOMIZED DESIGN.

[1]Abdullahi Shitu Umar ; [2]Godwin Ayuba Aniah ; [3]Aliyu Nuhu ; [4]Kassim Abdulganiyu ; [5]Hassan Idris

[1] Department of Statistics, Shehu Idris College of Health Sciences and technology, Makarfi.

[2] Department of statistics, Shehu Idris College of Health Sciences and technology Makarfi.

[3] Department of statistics,Shehu Idris College of Health Sciences and Technology

[4] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology, Makarfi, Kaduna State.

[5] Department of Statistics,Shehu Idris College of Health Sciences and Technology, Makarfi.

Corresponding Author: anuhu@sichstnkr.edu.ng

In our everyday life, there are differences or variations in every repeated thing we do. No matter how negligible these differences may be, the most important thing is that differences exist. Measuring or estimating these differences are known simply as estimating the measure of differences or estimating variance component. A Complete Randomized design with unequal replications was used in this study with an intention to obtain the most efficient and precise method of estimating or Measuring variance component in a compele Randomized Settings. The Methods under study are Stratified Sample Mean and One way ANOVA Sample mean .The Mean and Mean Square error of the two estimtors were obtained and the empirical studies revealed that the distributions of stratified sample mean and ONE WAY ANOVA Sample mean are the same, except that the location parameter for Stratified Sample mean is larger than the location parameter for ONE WAY ANOVA Sample mean. Since, the distribution of Stratified Sample mean is shifted to the right of the distribution of one way ANOVA sample mean and the amount that the distribution of Stratified Sample mean is shifted to the right of the distribution of one way ANOVA Sample mean exist. We then conclude that the two estimators do not differ significantly but the sratified Sample mean is more efficient in estimating or measuring variance components in a Complete Randomised design.

Keywords: "Variance Component","Estimator", "Efficient"," Strafied Sample Mean","One way ANOVA model".

GEOADDITIVE REGRESSION MODELLING WITH HIGH-DIMENSIONAL DATA

[1]Saheed Olalekan JABARU ; [2]Waheed Babatunde YAHYA ; [3]Kamoru JIMOH

[1] AL-HIKMAH UNIVERSITY, ILORIN, KWARA STATE, NIGERIA

[2] University of Ilorin, Ilorin, Nigeria

[3] AL-HIKMAH UNIVERSITY, ILORIN, KWARA STATE, NIGERIA

Corresponding Author: saheedjabar@alhikmah.edu.ng

Linear models are perhaps most widely used statistical models to investigate the influence of a set of predictors on a response variable, in which the number of predictors (p) of the data must be less than or equal to the number of observations (n); otherwise the design matrix will be singular and the parameters in the regression model cannot be uniquely estimated. This is a case of high-dimensional data where the conventional methods used to analyze low-dimensional data ($p < n$) are not directly applicable thereby making model's estimation and selection become more challenging. High-dimensional problems arise in many fields such as medicine, network security, biometrics, e-commerce, etc and could lead to models that are very complex. This research work aims at developing a geoadditive modelling technique with high-dimensional data focusing on epidemiological cases in Nigeria. Simulation of high dimensional data with n = 50, 100 and 150; p = 200, 250 and 1000 with varying parameters shall be performed to fit geoadditive models with Poisson distributed responses. Model regularization will be achieved using ridge regression, elastic net, LASSO variants (fused, group and adaptive LASSO) and Smoothly Clipped Absolute Deviation (SCAD). The AIC and BIC will be used for model selection while the Leave-one-out and k-fold cross validation techniques will be used for model's validation and evaluation on the test samples. In addition, real life data will be collected from the record of Nigeria Centre for Disease Control (NCDC) on the risk factors of a selected epidemic disease to validate the model.

Keywords: High-dimensional data, Geoadditive model, Epidemiology, LASSO, SCAD, Regularization.

ESTIMATING THE DIFFICULTY AND DISCRIMINATING INDICES OF SOME SELECTED COURSES IN THE DEPARTMENT OF MEDICAL STATISTICS, SICHSTM, KADUNA STATE, 2021.

[1]Abdullahi Shitu Umar ; [2]Aliyu Nuhu ; [3]Hassan Idris ; [4]Godwin Ayuba Aniah ; [5]Kassim Abdulganiyu

[1] Department of statistics,Shehu Idris College of Health Sciences and Technology, Makarfi.

[2] Department of statistics, Shehu Idris College of Health Sciences and Technology, Makarfi

[3] Department of Statistics, Shehu Idris College of Health Sciences and Technology, Makarfi.

[4] Department of Statistics, Shehu Idris College of Health Sciences and Technology.Makarfi.

[5] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology, Makarfi, Kaduna State.

Corresponding Author: kaganiyu@sichstmkr.edu

This study examined Difficulty and Discriminating Indices of some selected Coureses in the department of Medical Statistics, SICHSTM, Kaduna State. Five Courses were selected through purposive sampling from the list of courses available for 100 level Medical Statistics students, SICHSTM. The sample of the study consisted of 102 students that made up the population of the study. Pre and post- test was developed to validate the method used for data collection and analysis. The data collected were analyzed using difficulty index and discriminating index or Correlation Analysis. The findings revealed that the Expected difficulty index for STA 121 was almost fair , while the expected difficulty of STA 122 ,STA 124 ,MTH 122 and COM 124 were good. Hence ,STA 121 requires modification as the questions might have been unclear or misunderstood by the students. Also, discriminatory indexes for all the four couses revealed that the descrimination tend to be in opposite direction and considered inappropriate. We therefore conclude that the four courses (STA 122 ,STA 124 ,MTH 122 and COM 124) requires modification as the questions might have been too easy so that almost every student get it right with low discrimation, while STA 121 is difficult so that more than 50% of the students get it Wrong with Fair discrimination.

Keywords: "100L DMS","SICHSTM","Difficulty index","Discrimination index","Correlation Analysis"

VARIANCE DISPERSION GRAPH OF SOME OPTIMAL CENTRAL COMPOSITE DESIGNS IN CUBOIDAL REGION.

[1]ONWUAMAEZE CHARITY UCHENNA ; [2]NWAFOR CYNTHIA NDIDIAMAKA

[1] DEPARTMENT OF STATISTICS, UNIVERSITY OF NIGERIA, NSUKKA

[2] DEPARTMENT OF STATISTICS, UNIVERSITY OF NIGERIA, NSUKKA

Corresponding Author: uchenna.onwuamaeze@unn.edu.ng

Variance Dispersion Graph (VDG) is use to assess the performance of more than one design. In this work, the variance dispersion graph was used to assess the central composite design (CCD), small composite design (SCD) and Minimum run resolution V (MinResV) design in cuboidal region to give a comprehensive picture of the performance of the prediction variance throughout the region of interest. The CCD, SCD and MinResV were assessed for k = 3, 4, and 5 factors with 2, 3 and 5 center points. The result showed that CCD performed better than SCD and MinResV designs.

Keywords: "scaled prediction variance", "factorial portion", "axial points", "variance dispersion graph"

EFFECTS OF LENGTH OF DATA ON PREDICTIVE PERFORMANCE OF DYNAMIC TIME SERIES ANALYSIS: A CASE STUDY OF COVID-19 PANDEMIC IN NIGERIA

[1]Afolabi Saheed ; [2]Olanrewaju I. Shittu ; [3]Kehinde A. Bashiru

[1] University of Ibadan

[2] University of Ibadan

[3] Osun State University

Corresponding Author: Afsab4ever@yahoo.com

Most Analysists have always resulted into time series techniques to analyse time series data without consideration for the length of data. More often than not, forecast values from such analysis are not reliable. Evidence has shown that reliable forecast from time series data has to consider the length greater or equal to 100. The focus of this work is to illustrate the effects of length of data on parameter estimates of dynamic time series model with varying length of data. In this work, an attempt was made to compare forecast performance of two data sets on Covid-19 cases (Confirmed, Active and Death) in Nigeria from Worldometer using ARIMA(p,d,q) model. The First Data of 63 points (28th Feb., 2020 to 30th April, 2020) and the Second Data of 109 points (28th Feb., 2020 to 15th June, 2020) were used for the analysis. The forecast values were found to be increasing for the two data sets but the Second Data (Set N=109) with longer lengths were very close to the observed values from Woldometer with the minimum mean square error of forecast compare to data with shorter lengths. The result is in agreement with previous research.

Keywords: Time Series, Coronavirus, Variables, Modeling, Forecast performance

APPLICATION OF SOME SELECTED SKEWED DISTRIBUTIONS TO WIND SPEED IN IKEJA STATION, LAGOS STATE, NIGERIA

[1]Afolabi Saheed ; [2]Olalude, G. A. ; [3]Olaleye, O. A.

[1] University of Ibadan

[2] Federal Polytechnic, Ede

[3] Federal Polytechnic, Ede

Corresponding Author: Afsab4ever@yahoo.com

Wind Speed is a fundamental atmospheric quantity caused by air moving from high to low pressure, usually due to changes in temperature. This research aimed at modeling the wind speed (m/s) of Ikeja, Lagos State data between 2000 and 2020 using statistical techniques – descriptive statistics, data visualization, Anderson Darling and Cramer-Von Mises Statistics, Kolmogorov Smirnov and chi-square via three skewed distributions which are Gamma, Weibull and Log-Normal distribution. Descriptive analysis revealed that majority of the data set was all skewed to the right (? > 0). The test of Goodness of Fit conducted affirmed that wind speed at Ikeja Lagos follows Weibull distribution (p = 0.391). It can therefore be concluded that among the three distributions fitted of the wind speed (m/s), Weibull distribution fits the best with ? > 1, followed by Log-Normal Distribution and then Gamma Distribution. The implication of this study is that Ikeja may likely to experience more high wind speed in the future. Hence, certain control measures should be taken to avoid destruction of properties and valuables in Ikeja. This study will serve as information for the Meteorologists in the state to prepare ahead for a likely climate change in that part of the city.

Keywords: Modeling, Wind speed, Data visualization, Skewed distributions, Goodness of fit

MODEL SELECTION FOR TIME SERIES COUNT DATA IN PRESENCE OF OVER-DISPERSION AND EXCESS ZEROS

[1]Saleh Ibrahim Musa ; [2]N. O. Nweze

[1] Federal University Lafia, P.M.B. 146

[2] Department of Statistics, Nasarawa State University Keffi, PMB 122 Keffi, Nigeria

Corresponding Author: obininweze@yahoo.com

Time series of count with over-dispersion and excess zeros is the reality often encountered in many biomedical and public health applications. Rottenly, the Poisson and negative binomial distributions have been widely used in practice for discrete count time series data, their forms are too simplistic to accommodate features such as over-dispersion and excess zeros. Unable to account for these associated features while analyzing such data may result in incorrect and sometimes misleading inferences as well as detection of spurious associations. Therefore, the need for further investigation of count time series models suitable to fit count time series with over-dispersion of different level and excess zeros with different features becomes necessary. The study proposed a best model that can fit and forecast time series count data with different levels of over-dispersion and excess zeros. Simulation studies were conducted using R statistical package, to investigate the performances of INAR, ACP and PAR models at different sample sizes. The predictive ability of the models were observed at h-steps ahead for INAR, ACP and PAR models forms of autoregressive. The relative performance of each model was examined using MSE, AIC, BIC, and HQIC. Conclusively, in count time series with over dispersion data structures, the best model to fit was ACP at different sample sizes and the performance of INAR model competed with ACP models in zero inflated count series as the sample size increases. The predictive abilities of the four fitted models increased as sample size and number of steps ahead were increased.

Keywords: ACP, INAR, PAR, overdispersion, excess zeros

A THREE LEVEL SPLIT-LOT DESIGN IN SUDOKU SQUARE DESIGN STRUCTURE

[1]Umar Mustapha Anka ; [2]Nasiru Sani Dauran ; [3]Ahmed Audu ; [4]Ahmad Kamba ; [5]Sulyman Muhammed

[1] Usmanu Danfodiyo University, Sokoto.

[2] Usmanu Danfodiyo University Sokoto

[3] Usmanu Danfodiyo University Sokoto

[4] Usmanu Danfodiyo University Sokoto

[5] Usmanu Danfodiyo University Sokoto

Corresponding Author: mustinhoanka@gmail.com

ABSTRACT In this research, a situation where blocking is required for a split lot design in order to tease out noise from the dependent variable (prevent factors other than that of interest influencing the outcome) was considered. Blocking for every stage of the design is administered, an elaborate construction procedure for the design was developed by infusing the ordinary split-lot design into a Sudoku Square design structure, the hybridisation gave it a convenient structure for the research. The linear model and the sum of squares of the design were derived, the ANOVA table was constructed and the table was used to analyse the whole system. The clear advantage of this design has been observed to be the additional source of variation, because the introduction of the Block Sum of Squares have reduced the Error Sum of Squares as a result makes it more efficient, but the fact that precision and cost are both functions of the number of sublots per step as well as the total number of items, comparison is complicated. Which design provides greater precision for main effects, however, depend on the ratio MSXiE/MSE.

Keywords: Split lot, sub lots, design of experiments.

EFFICIENT DATA-MINING ALGORITHM FOR PREDICTING HEART DISEASE BASED ON ANGIOGRAPHIC TEST

[1]Banjoko Alabi ; [2]Yahya Waheed Babatunde ; [3]Garba Mohammed Kabir ; [4]Afolayan Razak Bayo ; [5]Abdulazeez Kawthar Opeyemi

[1] University of Ilorin

[2] University of Ilorin

[3] University of Ilorin

[4] University of Ilorin

[5] University of Ilorin

Corresponding Author: banjokoalabi@gmail.com

Background: The computerized classification and prediction of heart disease can be useful for medical personnel for the purpose of fast diagnosis with accurate results. This study presents an efficient classification method for predicting heart disease using a data-mining algorithm. Methods: The algorithm utilizes the weighted Support Vector Machine method for efficient classification of heart disease based on a binary response that indicates the presence or absence of heart disease as the result of an angiographic test. The optimal values of the Radial Basis Function kernel parameters for a heart disease classification were determined via a ten-fold cross-validation method. The heart disease data was partitioned into training and testing sets using different percentages of the splitting ratio. Each of the training sets was used in training the classification method while the predictive power of the method was evaluated on each of the test sets using the Monte-Carlo Cross-Validation resampling technique. The effect of different percentages of the splitting ratio on the method was also observed. Results: The Misclassification Error Rate was used to compare the performance of the method with three selected machine learning methods and also some already established result in the literature using the same heart disease data and it was observed that the proposed method performs best over others among the selected machine learning methods and also compete favourably well among the already established results. Conclusion: Finally, the results illustrate that the classification algorithm presented can effectively predict the heart disease status of an individual based on the results of an angiographic test.

Keywords: Weighted Support Vector Machine, Biserial Correlation, Cross-Validation, Heart Disease, Splitting Ratios

STATISTICAL PROPERTIES OF EXPONENTIATED EXTENDED EXPONENTIAL DISTRIBUTION WITH APPLICATIONS TO SURVIVAL DATA

[1]Ismail Kolawole

[1] Federal Polytechnic,Kaduna

Corresponding Author: talktoismak@yahoo.co.uk

Statistical Properties of Exponentiated Extended Exponential Distribution with Applications to Survival Data Kolawole Ismail Adekunle[1]*,Ibrahim Sule[2], Olalekan Akanji Bello3 *Corresponding author email: talktoismak@yahoo.co.uk [1]Department of Mathematics and Statistics, Kaduna Polytechnic, Kaduna, Nigeria. 2 & 3 Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria Abstract Many researchers in the field of distribution theory has main focus on generalizing or extending existing models in order to improve the flexibility of the classical distributions. This research paper introduced a new probability distribution called the exponentiated extended exponential distribution with three shape parameters and one scale parameter. Some of its mathematical and statistical properties such as quantile function, moments, moment generating function, cumulative hazard function, odds functions and reliability functions. The distribution of order statistic for this distribution was also obtained. Monte Carlo simulation was carried out to see the performance of MLEs of the exponentiated extended exponential distribution. We performed estimation of parameters by using the technique of maximum likelihood estimate. The proposed model was applied to two real datasets and the result showed that the new model provides better fit than other well-known distributions considered. KEY WORDS: EtEEtx, Survival data, and simulation

Keywords: "EtEEtx", "Survival data", and "simulation"

THE PROPERTIES OF KUMARASWAMY TYPE I HALF-LOGISTIC-G FAMILY OF DISTRIBUTIONS WITH APPLICATIONS TO REAL-LIFE DATA SETS

[1]Ismail Kolawole

[1] Federal Polytechnic,Kaduna

Corresponding Author: talktoismak@yahoo.co.uk

The properties of Kumaraswamy Type I half-logistic-G family of distributions with Applications to Real-life Data sets Ibrahim Sule1, Ismail Kolawole Adekunle2*, Olalekan Akanji Bello3 *Corresponding author email: talktoismak@yahoo.co.uk 1& 3 Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria [2]Department of Mathematics and Statistics, Kaduna Polytechnic, Kaduna, Nigeria. Abstract In this paper we present a new family of continuous distributions called Kumaraswamy Type I half-logistic family of distributions. Some signi?cant mathematical properties of the proposed family of distributions are determined. Estimates of the unknown parameters of the proposed family of distributions are obtained using the maximum likelihood method. Two sub-models are considered to assess the fit and flexibility of the family. Monte Carlo simulation is carried out in order to evaluate the estimation performance. Two practical datasets are considered to analyze the usefulness and ?exibility of the proposed distribution. The proposed Kumaraswamy Type I half-logistic exponential distribution outperformed other well-known distributions, showing its great adaptability in the context of real data analysis. Key words: Kumarswamy, Type II half-logistic and Monte Carlo,

Keywords: "Kumarswamy", "Type II half-logistic" and "Monte Carlo"

FOOD WEB ANALYSIS THROUGH INFORMATICS APPROACH- A DEEP LEARNING IMPLEMENTATION IN ECOLOGY

[1]BELLO ADESHINA ; [2]Yaun Chang li

[1] FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA

[2] Institute of Bioinformatics, Academia Sinica, Taiwan

Corresponding Author: oyedele.bello@futminna.edu.ng

Machine learning (ML) can build predictive models for the networks of trophic links in food webs to explaining ecosystem structure and dynamics. Logic-based machine learning has been successfully used to build insect food webs; using ANN with either MLR or SOM approaches. However, the observations on previous work's implementations of ML in the food web is that it is difficult to track their models, their architecture, and internal parameters. The ANN was presented, as a black box and this approach is now more trackable with the evolved knowledge of deep learning (DL). Which we have demonstrated in our implementation using DL with TensorFlow. Thinking the future-ecological big data, we have demonstrated why to consider DL in Ecology research. Our work explored insight into the future of the big data from various fossil sites that will help to identify species present, dead, or already extinct in environmental samples by sequencing. The need of implementing various protocols in DL for handling ecological big data was the motivation. In our DL, we incorporated prior knowledge to enhance our DL to be robust in extrapolation prediction in an unsupervised manner. Prior knowledge was combined from Globi meta-database. The trait-based approach computation space was achieved by ML forward-backwards propagation of gradient descent. The best positioning in food web trait-based was achieved by parameters twerking. We broadly classify node to taxa; classification link to trophic; Prediction-direct or undirected edges. NN input layer encodes the adjacency matrix values for the nodes. The output layer encodes the probability of species relationship type and function group learned by the machine from network data. Hidden layers are functions of the input and are used to efficiently encode through forward-backwards propagation. Results obtained, showed greater ecological realism in representing community can be realized in DL.

Keywords: deep learning ,ecology, food webs , TensorFlow

BAYESIAN CHANGEPOINT MODELLING OF A MULTIVARIATE GAUSSIAN DISTRIBUTION WITH APPLICATION TO SPANISH STOCK MARKET.

[1]Adegoke Taiwo

[1] University of Ilorin, Kwara State

Corresponding Author: adegoketaiwom@hotmail.com

Adegoke T. M., Yahya, W. B. Department of Statistics, University of Ilorin, Nigeria. e-mail: adegoketaiwom@gmail.com, wbyahya@unilorin.edu.ng In this work, we develop a single changepoint model for a multivariate normally distributed datasets using Bayesian techniques with an application to Spanish stock market datasets. The datasets was modelled by multivariate Gaussian density with an informative prior probability distribution specified for both the vector mean and the covariance of the distribution before and after the changepoint. The respective posterior distributions were derived using Markov Chain Monte Carlo simulation method based on the simulated and real life data. The results obtained from simulation studies proves that the Bayesian method developed was efficient to identify the point at which change actually occurred in multivariate normally distributed data sets. In the Spanish market dataset, the method was able to identify the period at which there was a change in the Spanish stock market which was Friday 22nd, 2018.

Keywords: Changepoint analysis, Bayesian method, Multivariate Gaussian distribution, Spanish stock market, Gibb sampling, Posterior Distribution.

TWO STATES VOLATILITY MODEL FOR FINANCIAL TIME SERIES MODELLING AND FORECASTING

[1]Kane Ibrahim

[1] Umaru Musa Yar'adua University

Corresponding Author: ibrahim.lawal@umyu.edu.ng

Volatility is an inevitable property of most financial time series and has since dominates literatures on financial time series modelling and foresting. Volatility analysis is an important aspect in many financial decisions for constructing less risky portfolios or obtaining higher profits. Many volatility models have been put in place. Most of the popular and successful volatility models are the family of autoregressive conditional heteroskedasticity (ARCH) models, however, volatility comes in two states, high state and low state of which the family ARCH models are unable to capture. This paper developed a two states volatility model for the low and high states using hybrid of hidden Markov model (HMM) and GARCH the generalized autoregressive conditional heteroskedastic model (GARCH). The model was applied to tax revenue data and was able to model the volatility in both states. Therefore the hybrid HMM-GARCH model has proved successful in separating and modelling volatility states in tax revenue data, hence recommended for future application of the model in other financial time series data.

Keywords: Volatility, Hidden Markov Model, GARCH, Tax Revenue, Low volatility, high volatility.

PERTURBED MANN ITERATION FOR STOCHASTIC SPLIT FEASIBILITY PROBLEM IN HILBERT SPACE

[1]Chijioke Joel Nweke ; [2]Akaninyene Udo Udom

[1] Alex Ekwueme Federal University, Ndufu-Aloke, Ikwo, Ebonyi State.

[2] University of Nigeria, Nsukka

Corresponding Author: cj_nweke@yahoo.com

Abstract This study examines the split feasibility problem in stochastic domain. The lipschitzian and pseudo-contractive mapping of the operators in Hilbert space was used. Almost surely convergence of iteration based on perturbed Mann-type methods for solving stochastic split feasibility problem was proved.

Keywords: Perturbed Mann iteration, split feasibility, convergence, stochastic

OPTIMAL CONTROL OF A MULTI-LEVEL HIERARCHICAL MANPOWER SYSTEM WITH A DEPARTMENTALIZED MARKOV DECISION ALGORITHM

[1]UDOM AKANINYENE

[1] DEPARTMENT OF STATISTICS, UNIVERSITY OF NIGERIA, NSUKKA

Corresponding Author: akaninyene.udom@yahoo.com

A recurrent problem in manpower control is how to attain a desired structural configuration, when this is possible, in an optimal way, since a desired structural configuration can be reached using different control inputs. The major aim of this paper is to develop a departmentalized solution algorithm based on Markov Decision Process (MDP) for optimal control of a Multi-level Hierarchical Manpower System (MHMS) by promotion or interdepartmental transfers or both. The problem is solve by decomposing the global problem into smaller local problems and the solutions of the local problems are combined, to generate a global solution.

Keywords: Markov decision process; multilevel hierarchical manpower system; optimal control; promotion control

PROBABILITY GENERATING FUNCTIONS OF OCCURRANCE OF DISEASES IN HEALTH CARE DELIVERY CENTRE IN NIGERIA.

[1]Godwin Ayuba Aniah ; [2]Abdullahi Shitu Umar ; [3]Kassim Abdulganiyu ; [4]Hassan Idris ; [5]Aliyu Nuhu

[1] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi, Kaduna State

[2] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology Makarfi,Kaduna State

[3] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology, Makarfi, Kaduna State.

[4] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology, Makarfi, Kaduna State.

[5] Department of Medical Statistics, Shehu Idris College of Health Sciences and Technology, Makarfi, Kaduna State

Corresponding Author: kaganiyu@sichstmkr.edu.ng

In some real life problems, we count the occurrences of some types of events in a given time. For example, you might have a random process Nt that shows the number of patients who arrive at a particular Hospital in a given time t starting from time 0. For such a processes, we usually assume N(0)=0, so as time passes and patients arrive, N(t) takes positive integer values(0,1,2 , , , , ,), for t in (0, &). There fore The number of patients arriving at a particular Hospital in a given time t can be modeled by a Poisson process with intensity ?t Patients per disease per time , When patient (X) arrives, he/she has probability P of having Malaria independently from what happens with any other patient, the PMF of a patient (X) having a particular disease is that of binomial with parameter(1,P). The theoretical study revealed that the probability generating function (PGF) of total cases of a particular disease 'M' in a given time G(s)=G(G(s) is that of poison random variable with intensity ?tP. Since the distribution of Total cases of a particular disease is Poisson with intensity ?tP then the distribution of; * Time to first reported case of that disease M in the Health center follow exponential distribution with parameter (1/?tP) and PGF G(S). *Time to Total reported cases of that disease M in the Health care center follow Gamma distribution with parameter (M, 1/?tP) and PGF G(S).

Keywords: "Probability Generating Function","Diseases",Occurance"," Health care Delivery center","Distributions".

THE ATTITUDE OF WOMEN TOWARDS THE UTILIZATION OF ANTENATAL CARE SERVICES AMONG REPRODUCTIVE AGE GROUP WOMEN (15-49 YEARS) IN KINKINAU, KADUNA SOUTH LOCAL GOVERNMENT OF KADUNA STATE.

[1]ABUBAKAR UMAR

[1] KADUNA STATE COLLEGE OF NURSING AND MIDWIFERY, POLYTECHNIC R0AD, TUDUN WADA KADUNA.

Corresponding Author: umarturaki@ymail.com

TITLE: THE ATTITUDE OF WOMEN TOWARDS THE UTILIZATION OF ANTENATAL CARE SERVICES AMONG REPRODUCTIVE AGE GROUP WOMEN (15-49 YEARS) IN KINKINAU, KADUNA SOUTH LOCAL GOVERNMENT OF KADUNA STATE. Abubakar, Umar Department of Midwifery, School of Midwifery, Kaduna State College of Nursing and Midwifery, Polytechnic Road Tudun Wada Kaduna. ABSTRACT This research work aimed at investigating the attitude of women towards the utilization of antenatal care services among reproductive age group women (15-49 years) in kinkinau, Kaduna south local government of Kaduna state. The study had been undertaken to provide opportunities for health education. Antenatal care makes an effort to recognize women who are at risk of complication. The study design was the descriptive survey, three (3) research questions were formulated to guide the study, a sample size of 100 was taken for the study, a well-developed questionnaire was used for data collection. The result showed that 95% of women have a positive attitude toward antenatal care services while 5% have a negative attitude, 89% of the respondents utilized antenatal while 11% did not, 36% said lack of awareness is a hindering factor for the utilization of antenatal care services, 49% said nurses attitude. From the above analysis, it was concluded that women have a positive attitude towards antenatal care services. Effort should be made in improving the quality of the services at the facilities, community education should also emphasize the importance of antenatal care in order to change the attitude of women and maximize the utilization of the services. Keyword: antenatal, utilization, attitude, Reproduction, women.

Keywords: Keyword: antenatal, utilization, attitude, Reproduction, women.

OPTIMAL DESIGN OF RECTIFYING SINGLE SAMPLING (RSS) PLAN WITH INSPECTION ERRORS

[1]IORKEGH SIMON

[1] UNIVERSITY OF ILORIN,ILORIN NIGERIA

Corresponding Author: simkegh@yahoo.com

Optimal Design of Rectifying Single Sampling Plan with Inspection Errors Iorkegh Simon Tyozua1*, Osanaiye P. A.2 1,2Department of Statistics, University of Ilorin, Ilorin-Nigeria *Corresponding Author: Tel: 08069659752; Email: simkegh@yahoo.com ABSTRACT Two types of errors are inherent in every acceptance sampling process. These include type I inspection error and type II inspection error. Type I inspection error occurs when a non-defective unit is classified as defective while type II inspection error occurs when a defective unit is classified as non-defective. In this paper, the effect of type I and type II inspection errors on the performance measures of rectifying single sampling (RSS) plan is investigated. Kumar's (2018) economic acceptance single-sampling plan model is adopted and modified to incorporate inspection error and additional objective functions introduced. Optimal sampling plans that minimize the total cost and satisfy both the producer's and the consumer's risk requirement is determined using both the existing model and the modified model. Comparison between the existing model and the modified model is made using the Operating characteristic (OC) curve, Average Total Inspection (ATI) and Total Cost (TC). Results showed that type I inspection error decreased the Average Outgoing Quality (AOQ) and Average Outgoing Quality Limit (AOQL) but increased Average Total Inspection (ATI). On the other hand, as type II inspection error increased and type I inspection error is kept at zero, the AOQ and AOQL increased but the ATI decreased. Also the optimal sampling plan in the modified model showed minimum total cost with lower producer's and consumer's risk as compared to the optimal sampling plan of the existing model. Therefore the modified model is found to performed better and is more economical than the existing model.

Keywords: Keywords: Average Outgoing Quality (AOQ), Average Outgoing Quality Limit,(AOQL) Average Total Inspection(ATI), producer's risk. Consumer's risk

POSYNOMIAL PROGRAMMING APPROXIMATION TO DYNAMIC PROGRAMMING

[1]Amuji Harrison

[1] Department of Statistics, FUTO

Corresponding Author: amujiobi@yahoo.com

Abstract: Dynamic programming has limitation in application due to the "course of dimensionality". Dynamic programming makes use of recursive relationship and principle of optimality; this makes the solution to explode because some variables are repeatedly carries along in the subsequent solutions. For this reason, there is an exponential increase in the state parameters thereby making a small problem to occupy a large memory space. The exponential increase in state variables (course of dimensionality) increases the dimension of the problem and makes it impractical to solve large problems. Hence, dynamic programming is restricted to solving small problems while the real life scenario demands more that. The researchers have found that this limitation can be solved by Posynomial programming approximation to Dynamic programming. This technique eliminates the restriction in the application of dynamic programming to solving problems of various sizes.

Keywords: Dynamic programming, Course of dimensionality, Recursive relationship, Posynomial programming, Optimal objective function

SARIMA MODELS FOR MODELING AND FORECASTING OF GOLD PRICES

[1]CHAKU SHAMMAH ; [2]Abdulrazaq Aminat Adebukola ; [3]Adehi Mary Unekwu ; [4]Aliyu Sani Aliyu ; [5]Timnan Bindap Ndam

[1] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

[2] NASARAWA STATE UNIVERSITY, KEFFI

[3] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

[4] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

[5] DEPARTMENT OF ECONOMICS, UNIVERSITY OF JOS

Corresponding Author: chakushammah@nsuk.edu.ng

Application of SARIMA model in modeling and forecasting average monthly gold prices was carried out in this study. Data on gold from January 2015 to December 2020 was obtained. Monthly adjusted close prices were used for the analysis. The gold price data was stationary after first difference (D = -3.8426, P = 0.02183< 0.05). SARIMA(0,0,0)(0,1,1)[12] was identified as the best model that fit the gold price data with minimum AIC and BIC. Forecast of gold prices from January, 2021 to December 2025 was obtained. Forecast values will be tested against actual values for January, 2021 to August, 2021 to note if actual values fall within forecast limits.

Keywords: SARIMA, Gold, Time Series, Forecast.

SURVIVAL ANALYSIS OF HIV PATIENTS DATA USING KAPLAN MEIER AND COX PROPORTIONAL REGRESSION

[1]CHAKU SHAMMAH ; [2]Bernard Okpara ; [3]Ahmed Ibrahim ; [4]Bilkisu Maijamaá

[1] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

[2] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

[3] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

[4] DEPARTMENT OF STATISTICS, NASARAWA STATE UNIVERSITY, KEFFI

Corresponding Author: chakushammah@nsuk.edu.ng

The purpose of this study was to assess the size of covariate, compare the Kaplan-Meier and Cox regression analysis and survival and predictors of mortality among adult patients starting highly active antiretroviral therapy at Federal Medical Centre Keffi, Nasarawa state. An institution- based retrospective study was conducted among the medical record of 384 study subject selected using simple random sampling from 1st January, 2021 to 31st December 2021. The data was analyzed using the Kaplan-Meier statistics and Cox regression Models and R package was used. Among 384 adult patients, we have a total of 128 male and 256 female,19 patients (4.95%) died, 3765 active (95.05%),we can conclude that the variable sex with Pr (>|z|) = 0.752 is not statistically significant likewise age > 50, with Pr (>|z|) = 0.330, hence the survival plot shows that the probability for both male and female HIV contracted patient to survive is 0.87 (87%), however, the female patients have higher risk on contraction than the male patients.

Keywords: Survival Analysis, Cox Regression, Kaplan Meier, HIV

ON D-OPTIMALITY CRITERIA FOR BLOCKING RESPONSE SURFACE DESIGNS: AN ALGORITHM

[1]UMELO-IBEMERE NGOZI

[1] FEDERAL UNIVERSITY OF TECHNOLOGY, OWERRI

Corresponding Author: ngoziumeloibemere@gmail.com

Ngozi C. Umelo-Ibemere Department of Statistics, Federal University of Technology, Owerri, Nigeria ngoziumeloibemere@gmail.com This article presents an algorithmic approach to constructing D-optimal designs for linear polynomial models with fixed block effects. The procedure begins by the selection of non-singular initial design points from the experimental space and arrangement of these design points in different blocks with pre-specified block size . A search is carried out to locate a point of maximum prediction variance which replaces the design point with minimum prediction variance thereby improving the initial design. The algorithm converges rapidly and its application to designs with equal and unequal block sizes are shown.

Keywords: D-optimal design, exchange algorithm, fixed block effect, unequal block size

A NEW MODIFIED BIASING PARAMETER ESTIMATORS FOR RIDGE REGRESSION MODELS

[1]Ibrahim Suleman

[1] Al-Hikmah University, Ilorin

Corresponding Author: adeshinas2010@alhikmah.edu.ng

[1] [2] S. A. Ibrahim and [1]W.B. Yahya [1]Department of Statistics, University of Ilorin, Ilorin Nigeria [2]Department of Physical Sciences, Al-Hikmah University, Ilorin, Nigeria Corresponding Author: e-mail: adeshinas2010@alhikmah.edu.ng, Tel: +234(0)8052262175
Abstract In this study, some new modified biasing parameters estimators for modelling linear regression in the presence of multicollinearity were proposed and their efficiencies were compared with some of the existing ones in the literature. Data were simulated from R statistical software at varying sample sizes n and correlation levels among the predictor variables. The analyses were performed at 1000 replication to ensure models' stability. The various estimators were assessed for efficiency using the criteria of Absolute bias (AB) and Mean Square Error (MSE). The results showed that the proposed biasing estimators are more efficient than the Least squares estimator and some other existing estimators when the data were plagued with the problem of multicollinearity. In conclusion, the newly proposed methods for biasing parameter in ridge regression are therefore recommended to be used when multicollinearity is the issue in multiple linear regression estimation and prediction. .

Keywords: Keywords: Multicollinearity, Ridge regression, Biasing parameter, and MSE

STATISTICAL ANALYSIS OF THE MEDIAN TEST AND THE MANN-WHITNEY U TEST

[1]OTI ERIC ; [2]Michael O. Olusola ; [3]Wariebi K. Alvan ; [4]Justin Ikirigo

[1] Department of Statistics, Federal Polytechnic, Ekowe Bayelsa State

[2] Department of Statistics, Nnamdi Azikiwe University, Awka Anambra State

[3] Department of Physics electronics, Federal Polytechnic Ekowe, Bayelsa State

[4] Department of Physics electronics, Federal Polytechnic Ekowe, Bayelsa State

Corresponding Author: eluchcollections@gmail.com

The median test and Mann-Whitney U test are nonparametric methods designed to handle two samples problem, their data are continuous which consist of two mutually independent random samples. They are used to test whether two (or more) independent samples have been drawn from populations with the same median. In this paper, we discussed and analyzed these two methods using the same illustrative example, testing the null hypothesis at 5 percent significance level and it was observed that both the median and Mann-Whitney U test were statistically significant indicating that the two samples of scores earned by students in Statistics department were drawn from populations with equal median scores in the course.

Keywords: Chi-Square, Contingency Table, Population, Rank, Two Samples.

MODELING OF NIGERIA STOCK EXCHANGE MARKET USING THRESHOLD AUTOREGRESSIVE MODEL.

[1]Oyenuga Iyabode ; [2]Agboluaje, Saheed Abidemi ; [3]Adeleye, James Adekola

[1] Department of Mathematics and Statistics, The Polytechnic, Ibadan

[2] The Polytechnic, Ibadan

[3] The Polytechnic, Ibadan

Corresponding Author: ioyenuga2007@polyibadan.edu.ng

The accuracy in forecasting the stock market prices or at least predicting the trend correctly is of crucial importance for any future investment in a dynamic global economy. Over the years, economists and finance analysts have consistently maintained that an unregulated market price is the best yardstick reflecting the true scarcity or worth of a commodity. Therefore, the study investigates the impact of the Nigeria Stock Exchange on Nigeria economy using Threshold Autoregressive Moving Average models. The data used for the analysis were called Nigeria Stock Exchange Market Capitalization monthly data published by the Central Bank of Nigeria Statistical bulletin between year 1985 and 2019. The non-stationary data was differenced to attain stationarity as well as reducing the season effect. An appropriate model was obtained to make reliable forecasts that would pave way for an improved economy if properly implemented.

Keywords: Threshold Autoregressive, non-stationary, differenced, Nigeria Stock Exchange.

ON THE RIGHT MODEL FOR COMPUTING THE GDP GROWTH ON SELECTED PETROLEUM PRODUCTS AND ITS TOTAL CONSUMPTION IN NIGERIA; A CASE STUDY OF NNPC PORT HARCOURT

[1]ONU Obineke Henry ; [2]Inamete, Emem Ndah H. ; [3]Pepple, Mc-Kelly Tamunotena

[1] Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Rivers State.

[2] Statistics Technology department, Federal Polytechnic of Oil and Gas, Bonny, Rivers State, Nigeria.

[3] Electrical/Electronics Engineering Technology department, Federal Polytechnic of Oil and Gas, Bonny, Rivers State, Nigeria.

Corresponding Author: onuobinekehenry@gmail.com

By 1. Onu, Obineke Henry, 2. Inamete, Emem Ndah H., 3.Pepple Mc-kelly Tamunotena 1Mathematics and Statistics department, Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Rivers State, Nigeria. 2 Statistics Technology department, Federal Polytechnic of Oil and Gas, Bonny, Rivers State, Nigeria. 3 Electrical/Electronics Engineering Technology department, Federal Polytechnic of Oil and Gas, Bonny, Rivers State, Nigeria. Corresponding email:onuobinekehenry@gmail.com Abstract The study considered choosing the right model for the computation of the GDP Growth on selected petroleum products in Nigeria, with the effect of the Total consumption of these products also considered. We built three multiple linear regression models. (a) having four interactions, (b) having no interactions and (c) having three interactions. The four interaction model contains the interaction terms of AGO*COKE, AGO*RF, AGO*LPFO, and AGO*TFC, while the three interaction model only excluded AGO*COKE. The paper considered ANOVA, P-value, F-value and R-squared test statistic in the analysis. It was revealed in the four interaction model that AGO and COKE are not significant and that all the interactions are not significant in the analysis, which suggested the use of model that have all the interactions omitted and in this model, AGO became significant while the insignificance of COKE reduced by 0.511. all the other variables are significant. Though, the model with four interactions is better by the value of R-squared and F-value. For the model with three interactions, it was revealed that all the variables including the three interactions are significant. It gave the R-square value of 95.67% with 0.06% slightly lower than the model with four interactions. Since the difference is negligible, and because all the variables in the three interaction model both the interaction terms are significant, hence the model with three interactions is recommended as the best model for this analysis.

Keywords: 'Multiple Regression' 'Interaction' Petroleum Products' 'ANOVA; R-squared'.

CHOICE OF PRINCIPAL FITTED COMPONENTS REGRESSION DIMENSION WITH INFORMATION COMPLEXITY CRITERION

[1]Olorede Kabir ; [2]Waheed Babatunde Yahya ; [3]Hamparsum Bozdogan

[1] Department of Statistics & Mathematical Sciences, Kwara State University, Malete

[2] Department of Statistics, P.M.B. 1515, University of Ilorin, Ilorin Nigeria

[3] Department of Business Analytics & Statistics, The University of Tennessee Knoxville, TN 37996 USA

Corresponding Author: kabir.olorede@kwasu.edu.ng

At the core of principal fitted component (PFC) regression applications is the choice of optimal number of dimensions, d, to consider, the choice of the basis function and the structure of PFC covariance. The premise of this study is to propose information complexity criterion as a data-adaptive criterion for these PFC requirements. An information theoretic entropic criterion called Bozdogan's (1987) Information Complexity (ICOMP) criterion is proposed to further address arbitrariness and anomalies that have confronted applications of PCF methodologies with other information criteria including Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) and the Likelihood Ratio Test (LRT). Specifically, Bozdogan's consistent information complexity (CICOMP) criterion is proposed as a new information theoretic criterion for selecting dimension cardinality d, PFC basis function cardinality and covariance structure without misspecification in Principal Fitted Component Regression (PFCR) model fitting. The new proposal in PFC applications promises to be an efficient information criterion for selecting optimal d cardinality in high-dimensional PFCR model especially when other criteria are known to misspecify the model.

Keywords: Akaike Information Criterion, Bayesian Information Criterion, Information Complexity, Consistent Information Complexity Criterion, Covariance Complexity

SPARSE SUFFICIENT DIMENSION REDUCTION USING MAXIMUM ENTROPY COVARIANCE ESTIMATOR WITH INFORMATION COMPLEXITY

[1]Olorede Kabir ; [2]Waheed Babatunde Yahya ; [3]OluFemi A. Omitaomu

[1] Department of Statistics & Mathematical Sciences, Kwara State University, Malete

[2] Department of Statistics, P.M.B. 1515, University of Ilorin, Ilorin Nigeria

[3] Computational Systems Engineering and Cybernetics, Oak Ridge National Laboratory; Department of Industrial and Systems Engineering University of Tennessee, Knoxville.

Corresponding Author: kabir.olorede@kwasu.edu.ng

In this paper, a new shrinkage sufficient dimension reduction (SSDR) method that contains full information on the entire features in high and ultrahigh-dimensional data sets is proposed. A new cost-economy unified estimation strategy in regression- and classification-type formulations of SDR methods is presented. The procedure utilised the maximum entropy covariance hybridization and maximal covariance complexity-based shrinkage estimation to produce sparse and accurate solutions. The novel contribution here includes the hybridization of the newly proposed maximum entropy covariance (MEC) estimator by Olorede and Yahya (2019) with existing smoothed covariance estimators (SCEs) to data-adaptively generate shrunken estimate of the inverse MEC estimator. To our knowledge, utilization of SCEs and Hybridized Smoothed Maximum Entropy Covariance Estimator (HSMEC) in SDR formulation has not been presented in the literature. Our method can be applied to most of the existing SDR methods such as sliced inverse regression, sliced average variance estimation and principal hessian directions. We demonstrated the utility, versatility, and effectiveness of the proposed method with Riboflavin (vitamin B2) production rate data and all the results obtained showed that the proposed method is quite efficient with high predictive performance.

Keywords: Sufficient Dimension Reduction, Maximum Entropy, Smoothed and Hybridized Covariance Estimators, Maximum Entropy Covariance (MEC) Estimator, Information Complexity

BURR X EXPONENTIAL – EXPONENTIAL DISTRIBUTION: PROPERTIES AND APPLICATION. A. A. SANUSI, S.I.S DOGUWA, A. YAHAYA AND Y. M. BARAYA

[1]Sanusi Abdulmuahymin

[1] Kashere, Gmbe State.

Corresponding Author: abdulmuahymin81@gmail.com

A new distribution called Burr X Exponential – Exponential (BXEE)distribution is developed in this paper. The respective density and distribution functions of this distribution were shown. Some structural properties of this new distribution were derived such as moment, moment generating function, quantile function, probability weighted moment, Renyi entropy and order statistics. The parameters of this distribution were estimated using Maximum Likelihood Estimate (MLE) methods. Finally, two real data sets were used to validate the results obtained from MLE and Monte Carlos simulation was carried out. The results shown that BXEE distribution provide better fit in the data sets than some other commonly known distributions. Perhaps, this new distribution may be used to modeling positive real life data sets.

Keywords: Burr X Exponential – Exponential distribution, Moment generating function, Quantile function, Renyi entropy, Order statistics, probability weighted moment, maximum likelihood estimation.