



CHARTERED INSTITUTE OF STATISTICIANS OF NIGERIA (CISON)

Established by CISON Act 2022

(Publishers of the Journal of the Chartered Institute of Statisticians of Nigeria)

BOOK OF ABSTRACTS





CHARTERED INSTITUTE OF STATISTICIANS OF NIGERIA (CISON)

Established by CISON Act 2022

(Publishers of the Journal of the Chartered Institute of Statisticians of Nigeria)

2ND CISON INTERNATIONAL CONFERENCE 2025

Theme: Big Data and Analytics in Nigeria:
Developing a National Data Ecosystem for
Enhanced National Prosperity

Date: September 22–26, 2025

Venue: The Dome Event Centre, Asaba, Delta State



TABLE OF CONTENTS

PAPERS ON STATISTICAL THEORIES

- T001.25** A Compound Distribution for Modelling Discrete Failure Events - Elebe E. Nwezza, Uchenna U. Uwadi, Chukwunyenye I. Okonkwo, Kelechi E. Arua, Ikenna E. Chimezie
- T002.25** A Modified Gaussian Kernel Weights for Improving Goodness-of-Fit of Local Linear Regression - Efosa Edionwe and Omo Eguasa
- T003.25** A Modified New Weighted Rayleigh Distribution - Kole, Emmanuel and Ilori, Adetunji K.
- T004.25** Efficiency of the Variance Estimators in the Presence of Random Nonresponse Under Two-Phase Sampling - Isah Muhammad, Muftaba Suleiman, Adamu Zakar Adamu and Samaila Manzo
- T005.25** Parameterized Extension of Exponential Discriminant Analysis for Enhanced Classification - F. Meka, J. E. Osemwenkhae and A. Iduseri
- T006.25** Simulation And Application of the Transmuted Unit Burr XII Distribution in Real-world Data Analysis - Abdulhameed Ado Osi, Yusuf Bello and Muftahu Zubairu Ringim
- T007.25** On Parameters Estimation of Sine Type II Topp-Leone Frechet Distribution - M. A. Mahmud, W. B. Yahya, Abubakar Awwal and Ahmed Abdullahi Makada
- T008.25** Power Unit Garima Distribution: Properties, Comparison of Estimation Procedures and Diverse Applications - Emmanuel Wilfred Okereke
- T009.25** Inverse Power Hazard Rate Distribution: Properties and its Application - Elem-Uche, O., Nsude, F. I., Dike, A. O. and Madu, C.
- T010.25** A Generalized Two-Step Calibrated Ratio Estimators of Population Mean in Stratified Random Sampling - Ekpenyong, Emmanuel John and Nelson, Theophilus Obijuru
- T011.25** Optimal Control Model Using Probability Distribution - Chinelo U. Chikwelu, J. I. Mbegbu and F. Ewere
- T012.25** An Efficient Generalized Estimator of Population mean in Two-Phase Successive Sampling Under Random Nonresponse - Ewere, F., Mbegbu, J. I. and Agbebia, C. S.
- T013.25** The Odd Teissier Kumaraswamy Distribution: Properties and Applications Mbegbu Julian I., Ekhosuehi N. and Nzei Lawrence C.

PAPERS ON STATISTICAL APPLICATIONS

- A001.25** A Comparative Analysis of Selected Forecasting Models Using Inflation Rate - Onyeka-Ubaka, J. N., Arowolo O. T. Adeniyi, S. A.
- A002.25** A Comparative Power Analysis of Heteroscedasticity Detection Tests Using Simulated Data - Obabire, Akinleye A. Olumi, Toba Timothy, Agadaga, Gregory Okiemute
- A003.25** A Comparison of Bayesian Regression and Classical Ordinary Least Squares Methods for Multiple Linear Regression Using Post-utme Data - C. O. Odijie and N. Ekhosuehi



- A004.25** Advanced Machine Learning Algorithms for Cardiovascular Disease Prediction: Comparative Evaluation with Clinical Data Integration - Abubakar F. Kabara, Ishaq O. Olawoyin, Auwalu Ibrahim, Yakubu Husaini, Hamisu A. Hamisu, Salma S. Suleiman, Nafisat Yusuf, Fatima U. Muhammad, Umar A. Huguma and James S. Ogheneovo
- A005.25** Comparative Analysis of Stochastic Models and Machine Learning Algorithms for Inflation Rate Prediction In Nigeria - Edesiri Bridget Nkemnole and Abiodun Oyelami
- A006.45** Comparative Analysis on Effects of Early Treatment of Lassa Fever and Symptoms in Nigeria - Kole, Emmanuel and Ilori, Adetunji, K.
- A007.25** Counts Data Regression Modeling of Lassa Fever in Nigeria in the Presence of Over-dispersion - Abiodun Obaromi, Wadai Mutah, Omoha Joy, Grace Abakpa
- A008.25** Determinants of Minimum Wage Compliance in Nigerian States: A Panel Logistic Regression Analysis - Abiola T. Owolabi, Zainab T. Abdullahi, Olasunkanmi J. Oladapo, Janet I. Idowu and Kayode Ayinde
- A009.25** Dynamics of Crude Oil Price, Production and Exportation in Nigeria (2006 – 2025): A Time-series Analysis - G. C. Ibeh, J. C. Ajaraogu and C. E. Onyenekwe
- A010.25** Educational Curriculum Development for Data Science and Analytics: A Strategic Imperative for Nigeria's Digital Future - H. R. Bukari, M. Kaka and F. W. Usman
- A011.25** Waste Management in Nigeria: A Vector Error Correction Model Analysis of Population and Employment Impacts - Shehu Abdullahi
- A012.25** Fuzzy Modelling on Malaria Disease Dynamics - Enwere Kelachi Precious and Ogoke Uchenna Petronilla
- A013.25** Multivariate Approach to Investigate Risk Factors Predisposing Under-Five to Severe Broncho-Pneumonia - S. S. Ahmed, B. Maijama'a and M. O. Adenomom
- A014.25** Predictive Analytics for Early Detection of Disease Outbreaks in Urban Slums: A Machine Learning Approach - Joseph A. Akinyemi, Matthew I. Ekum, Olatunji T. Arowolo and Bolarinwa O. Ajala
- A015.25** Predictive Modelling of Health Status Awareness among Academic Staff in Tertiary Institutions in Kogi State Using Random Forest - Adenike Oluwafunmilola Olubiyi, Oluwamayowa Opeyimika Olajide, Michael Sunday Olayemi
- A016.25** Spatial Assessment of Flooding Impact on Food Security in Coastal Nigeria Using a Geographically Weighted Regression Approach - Odior, K. A. and Elugwu, F.
- A017.25** Comparison of Arima-e with Arima-n Performance in Modeling Nigeria's GDP, 1960 – 2024 - Salisu Shehu Umar, Muhammed Adamu Obomeghie and Bello Andrew Ojutomori



A018.25 Modeling Insurance Claim Data: An Infinite-sum Representation of Aggregate Claim Approval Distribution - Friday I. Agu

A019.25 The Application of Ordered Logit Model on Blood Pressure (BP) of Patients. (a Case Study of Dalhatu Araf Specialist Hospital Lafia) - Ehada, M. I., Rilwanu Y. M. and Gona, S. I.

A020.25 Empowering The Next Generation of Data Scientists: A Comprehensive Curriculum Framework for Data Science and Analytics Education - Kamaldeen Olarewaju Adetunji

A021.25 Leveraging Predictive Analytics for Enrollment Trend and Curriculum Innovation in Nigerian Statistics Programs - Chidimma Florence Ejiofor, Ijeoma Eberechukwu Okechukwu, Emeka Henry Chukwueloka

A022.25 A Generalized Bridge Regression on Infectious Disease Modelling - Enwere Kelachi Precious, Ogoke Uchenna Petronilla and Ethelbert C. Nduka

A023.25 Development of Machine Learning Algorithms for Predicting Prostate Cancer in Benue State, Nigeria - Edwin Hart Ogwuche and Rotimi Felix Afolabi

A024.25 Evaluating Sphericity Tests Under Multivariate Normal and Heavy-Tailed Distributions: A Simulation Study of Type I Error and Power Across Dimensions - A. D. Akinboro and G. M. Oyeyemi

A025.25 Impact of Monetary Policy Using Machine Learning-Based Counterfactual Analysis - Itiveh, F. E. and Adams, J. M.

A026.25 Improving Agricultural Efficiency with Artificial Intelligence and Data Analytics: An Analysis of Corn Yield Prediction - Kenenisa Abdisa Kuse

A027.25 On the Prediction Variance Performances of Space-filling Design in some Response Surface Models Under Cuboidal Space - J. C. Nwanya, H. I. Mbachu, R. Efeizomor

A028.25 Unemployment And Its Role In Crime Rates In Nigeria: An Empirical Investigation Using ARDL - Adigun K. Abimbola Adigun A. Oladele, and Olubiyi, O. A

A029.25 Nigeria's Population Growth Analysis Using Ridge and Xgboost Regression - Uchenna J. Okeke, Cornelius Ayantse, OlaOluwa S. Yaya and Evelyn N. Okeke

A030.25 Statistical Analysis of Life Expectancy Data Using a New Parametric Model - F. C. Opone, D. Enegesele, A. J. Otunomeruke, P. E. Omosioni, J. C. Ehiwario and N. Ekhsosuehi

A031.25 Logistic Quantile Regression Model Using Outlier Distributed Data Set: A Comparative Robustness Analysis - Besta Okey Onyegbuchulem¹, Chialuka Adline Onyegbuchulem, and Joy Chioma Nwabueze

A032.25 Leveraging AI and Big Data Analytics for Enhanced National Security Intelligence - Taoheedat Alanamu

- A033.25** Macroeconomic Shocks and Income Growth in Africa: Evidence from a Panel Structural Var Analysis - Umeokwobi Richard, Dalhatu Bello, Adams Joshua, and Anih David
- A034.25** A Re-appraisal of the Supply Side Determinants of Banking Credit in Nigeria - G. U. Ebuh, I. B. Ezike, J. O. Obeta, B. Dalhatu, R. O. Umeokwobi, D. S. Yusuf and B. A. Umemezia
- A035.25** Statistical Method for the Analysis of Results from Controlled Comparative Diagnostic Screening Tests and Clinical Trials for a Condition in a Population - Matthew Chukwuma Michael, Oyeka, I. C. A., Ajibade, Bright, Onyiorah, A. A. and Efobi, C. C.
- A036.25:** Determinants of Financial Inclusion in Nigeria: Evidence from the 2023 Nationally Representative Household Survey on Access to Financial Inclusion - Adanegbe Peter, and Zakari Isiaka Osheku
- A037.25:** Impact of Fuel Subsidy Reforms on Nigerians' Purchasing Power - Nasir Bukola Rasheed
- A038.25:** Socioeconomic Status and Primary Health Care Utilization in Nasarawa State, Nigeria - Job Ozoga
- A039.25:** Statistical Analysis of Retail Prices of Staple Foods in Nigeria Based on National Bureau of Statistics (NBS) Data - Nuga Kehinde Adetayo and Omotola Adeniyi Andrew
- A040.25:** Statistical Assessment of Trends, Prevalence and Socio-Demographic Correlates of Some Selected Cancer Types at University of Nigeria Teaching Hospital (UNTH), Ituku-Ozalla, Enugu - Okanchi, Louis Chukwunweike, Ezeani, Chinyere Maria and Igwenagu, Chinelo Mercy
- A041.25:** Statistical Investigation of Some Determinants of High Blood Pressure (HBP) among Women of Reproductive-Age in Chukwuemeka Odumegwu Ojukwu University Teaching Hospital (COOUTH), Amaku, Anambra State, Nigeria - Ezeani, Chinyere Maria, Okanchi, Louis Chukwunweike and Igwenagu, Chinelo Mercy
- A042.25:** The Implementation of Machine Learning in the Prediction and Development of a System for Improved Diet Recommendation for Cardiovascular Diseases - Turman Henry, Christiana J. Bakut, Vincent Kanwai and Turman Felicia

PAPERS ON STATISTICAL THEORIES



T001.25: A COMPOUND DISTRIBUTION FOR MODELLING DISCRETE FAILURE EVENTS

Elebe E. Nwezza, Uchenna U. Uwadi, Chukwunenye I. Okonkwo, Kelechi E. Arua and Ikenna E. Chimezie

¹Department of Mathematics and Statistics, Alex Ekwueme Federal University Ndufu-Alike, Ebonyi State

Corresponding Email: elebe.nwezza@funai.edu.ng

ABSTRACT

In modeling discrete outcomes, classical discrete distributions are usually associated with constant parameter defining the probability of occurrence, as in the case of Bernoulli, binomial, geometric, etc., distributions, or rate of occurrence, as in the case of Poisson distribution. In this paper, we generated a new distribution for modeling the occurrence of discrete failure events with varying probability. Some of the properties of the new distribution such as the moments about the origin and the probability generating function were studied.

T002.25: A MODIFIED GAUSSIAN KERNEL WEIGHTS FOR IMPROVING GOODNESS-OF-FIT OF LOCAL LINEAR REGRESSION

Efosa Edionwe¹ and Omo Eguasa²

¹Department of Statistics, Federal University of Petroleum Resources, Effurun, Delta State, Nigeria;

²Department of Physical Sciences, Benson Idahosa University, Benin City, Edo State, Nigeria
Corresponding Email: edionwe.efosa@fupre.edu.ng

ABSTRACT

The flexibility of nonparametric regression models such as the local linear regression (LLR) is credited to the kernel weights generated from some decreasing functions such as the simplified Gaussian kernel function (GKF). It is observed that as the bandwidth (smoothing parameter) tends to 1 and above, the kernel weights all tend approximately to the same (constant) value for which the LLR produces essentially the same goodness-of-fit results as those of the ordinary least squares (OLS) method. This implies that a measure of the flexibility of LLR over the OLS is a function of the degree of dispersion of the kernel weights. In this paper, a more general form of the GKF is obtained from which a modified version that guarantees kernel weights of higher measure of dispersion (flexibility) is proposed. Comparisons of the goodness-of-fit of four responses from a widely-analyzed response surface problem from the literature show substantial improvements in the performance of LLR that utilizes the kernel weights generated from the proposed GKF.

T003.25: A MODIFIED NEW WEIGHTED RAYLEIGH DISTRIBUTION

Kole¹, Emmanuel and Ilori², Adetunji K.

¹Department of Mathematics and Statistics, Kaduna Polytechnic, Kaduna

²Statistics Programme, National Mathematical Centre, Abuja

Corresponding Email: koleemmanuel@kadunapolytechnic.edu.ng



ABSTRACT

This paper introduces the Weighted Rayleigh (WR) distribution by inducing inverted weight function into the existing Rayleigh distribution. Statistical and mathematical expressions of its properties such as Survival Function, Hazard Function, Moments, Moment Generating Function, Mean Deviation and Renyi entropy were explicitly derived. The model's parameter was estimated using maximum likelihood method of estimation. Two real life data sets on cancer and waiting time before service were considered to assess the flexibility of the Weighted Rayleigh distribution over existing distributions. The distributions performances were compared using Log-likelihood and Akaike Information Criteria (AIC). The Weighted Rayleigh distribution fits the real life data better than the Rayleigh, Inverse Weibull (IW) and Weighted Inverse Weibull (WIW) distributions.

T004.25: EFFICIENCY OF THE VARIANCE ESTIMATORS IN THE PRESENCE OF RANDOM NONRESPONSE UNDER TWO-PHASE SAMPLING

Isah Muhammad¹, Mujtaba Suleiman², Adamu Zakar Adamu¹ and Samaila Manzo¹

¹Department of Statistics, Binyaminu Usman Polytechnic, Hadejia, Nigeria.

²Department of Health Information Management, Federal Polytechnic, Daura, Nigeria.

Corresponding Email: isahsta@gmail.com

ABSTRACT

In this paper, a generalized two-phase estimator of population variance under two realistic situations of random non-responses is proposed. The properties of the estimators such as bias and mean square error were derived and compared theoretically with some existing estimators. The performances of the proposed and existing estimators were assessed using four real data sets. Evidence from the empirical study revealed that the proposed estimators performed better compared to usual variance estimator and other existing estimators considered based on the criterion of mean square error and percentage relative efficiency. Thus, the proposed estimators in this study are more efficient than some existing estimators and can be used to estimate variations that exist in real-world problems when there is random non-response.

T005.25: PARAMETERIZED EXTENSION OF EXPONENTIAL DISCRIMINANT ANALYSIS FOR ENHANCED CLASSIFICATION

F. Meka¹, J. E. Osemwenkhae² and A. Iduseri²

¹Department of Mathematics and Statistics, University of Delta, Agbor.

²Department of Statistics, University of Benin, Benin City

Corresponding Email: fortune.meka@unidel.edu.ng

ABSTRACT

Exponential Discriminant Analysis (EDA) has emerged as a powerful technique for dimensionality reduction and classification, particularly in scenarios involving high-dimensional data. Traditional EDA employs the matrix exponential of scatter matrices in the form $y=e^x$, where x represents a scatter matrix. However, this fixed exponential form lacks flexibility and may not fully capture the intrinsic structure of diverse datasets. In this paper, we propose a New Improved Exponential Discriminant Analysis (NIEDA) framework that generalises EDA by incorporating a parameterised

exponential function of the form $y=ae^{kx}$, where a and k are scalar parameters controlling the scale and growth rate of the transformation, respectively. This formulation allows adaptive scaling of the scatter information, enabling better separation of classes. Experiments on benchmark datasets demonstrate that NIEDA outperforms its counterparts in terms of classification accuracy.

T006.25: SIMULATION AND APPLICATION OF THE TRANSMUTED UNIT BURR XII DISTRIBUTION IN REAL-WORLD DATA ANALYSIS

Abdulhameed Ado Osi¹, Yusuf Bello² and Muftahu Zubairu Ringim³

¹Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Nigeria

²Department of Statistics, Federal University Dutsin-Ma, Nigeria

³Department of Mathematics, Aliko Dangote University of Science and Technology, Wudil, Nigeria

Corresponding Email: aaosi@kustwudil.edu.ng

ABSTRACT

This study introduces the Transmuted Unit Burr XII (TUB-XII) distribution, a flexible probabilistic model for analyzing bounded data with heavy tails and skewness. By incorporating a transmutation parameter into the unit support Burr XII framework, the proposed distribution generalizes several existing models while offering enhanced versatility in modeling (0,1)-interval data. We derive its key statistical properties, including moments, quantile functions, and hazard rate behavior. A maximum likelihood estimation procedure is developed for parameter inference, supported by a Monte Carlo simulation study to assess finite-sample performance. The utility of the TUB-XII distribution is demonstrated through empirical applications in reliability engineering (failure time analysis) and financial risk modeling (proportional loss data). Comparative results with Beta and Kumaraswamy distributions confirm its superior fit for datasets exhibiting extreme values or asymmetric patterns. The proposed model provides practitioners with a robust tool for proportional data analysis, implemented via reproducible R algorithms.

T007.25: ON PARAMETERS ESTIMATION OF SINE TYPE II TOPP-LEONE FRECHET DISTRIBUTION

M. A. Mahmud¹, W. B. Yahya², Abubakar Awwal³ and Ahmed Abdullahi Makada⁴

^{1,4}Department of Mathematics and Statistics, Niger State Polytechnic, Zungeru, Nigeria.

²Center for Statistical Genomics and Computational Biology, Department of Statistics, University of Ilorin, Nigeria.

³Department of Statistics and Data Analytics, Nasarawa State University, Lafia, Nigeria

Corresponding Email: mustaphamahmud746@gmail.com

ABSTRACT

The Sine Type II Topp-Leone Frechet (STIITLF) distribution is a probability model that blends the periodic nature of the sine function with the extreme-value characteristics of the Frechet distribution. This study focuses on methods for estimating the parameters of the STIITLF distribution. Specifically, it aims to assess parameter estimation techniques, including Maximum Likelihood Estimation (MLE) and Maximum Product of Spacing Estimation (MPSE), using Root Mean Square Error (RMSE) as the main criterion for comparing performance. The results indicate

that as sample size increases, the parameter estimates from all methods converge toward the true values, with reductions in bias and RMSE. Among the methods considered, MLE consistently yields parameter estimates that are closest to the true values, showing the lowest bias and RMSE across small, moderate, and large sample sizes, thereby establishing it as the most effective estimation approach for the STITLF distribution.

T008.25: POWER UNIT GARIMA DISTRIBUTION: PROPERTIES, COMPARISON OF ESTIMATION PROCEDURES AND DIVERSE APPLICATIONS

Emmanuel Wilfred Okereke

Department of Statistics,

Michael Okpara University of Agriculture, Umudike, Nigeria

Corresponding Email: okereke.emmanuel@mouau.edu.ng

ABSTRACT

This paper proposed a new statistical distribution named the power unit Garima (PUG) distribution, which is an extension of the unit Garima distribution found through a power transformation. Properties of this new model, among which are mode, moments, entropy, and Bonferroni and Lorenz curves were considerably studied. Additionally, several estimation techniques for the parameters of the PUG distribution were examined through a Monte Carlo simulation procedure, including maximum likelihood, ordinary least squares, weight least squares, Anderson-Darling, and Cramer-von Mises methodologies. The simulation results obtained are in favour of the maximum likelihood approach. Goodness of fits of the distribution to three data relating to various fields was determined relative to that of the fits of unit power Lindley, Kumaraswamy and beta distributions to the data. Consequently, the PUG distribution proved to be the best in modelling the data among the distributions under consideration.

T009.25: INVERSE POWER HAZARD RATE DISTRIBUTION: PROPERTIES AND ITS APPLICATION

Elem-Uche, O., Nsude, F. I., Dike, A. O. and Madu, C.

Akanu Ibiam Federal Polytechnic Unwana Afikpo

Corresponding Email: urchstat@gmail.com

ABSTRACT

A new distribution called the inverse power hazard rate (i-PHR) distribution has been introduced. The statistical characteristics such as the probability density function, cumulative distribution function, survival function, hazard rate function, quantile function, moments, mode, median, quantile function, stochastic ordering, order statistics, entropy and stress-strength reliability for i-PHR distribution are obtained. The maximum likelihood estimators is derived and employed to estimate the parameters of the distribution. Real life data set were used to discuss the application of i-PHR distribution to check its suitability over some other existing models. The goodness-of-fit measures such as Akaike information criterion (AIC), Bayesian information criterion (BIC) and Kolmogorov-Smirnov test was used. The result shows that the new proposed distribution is more flexible and provides a better fit than the other distributions considered based on the data used.

T010.25: A GENERALIZED TWO-STEP CALIBRATED RATIO ESTIMATORS OF POPULATION MEAN IN STRATIFIED RANDOM SAMPLING

Ekpenyong, Emmanuel John and Nelson, Theophilus Obijuru

Department of Statistics, Michael Okpara University of Agriculture, Umudike,
Abia State Nigeria

Corresponding Email: ej.ekpenyong@mouau.edu.ng

ABSTRACT

This paper develops a family of two-step calibration ratio estimator for estimating the finite population mean in stratified random sampling. The statistical properties of this family are derived up to first order approximation using Taylor's series approximation method. An existing calibration estimator have shown to be a member of the generalized two-step calibration ratio estimator for a case of unit constant of proportionality. An empirical study has been carried out through simulation studies and real data application to ascertain the efficiency of the proposed estimators over other existing estimators considered in the study; and the results revealed that some members of the proposed estimators outperform the existing estimators, while some compete favourably with them.

T011.25: OPTIMAL CONTROL MODEL USING PROBABILITY DISTRIBUTION

Chinelo U. Chikwelu¹, J. I. Mbegbu² and F. Ewere²

¹Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Nigeria

²Department of Statistics, Faculty of Physical Sciences, University of Benin, Benin City, Nigeria

Corresponding Email: uc.anyadiogwu@unizik.edu.ng

ABSTRACT

In this article, we introduce a novel probability distribution, the Type I Heavy Tail Topp Leone Rayleigh (TI-HT-TLR) distribution, and apply it to an optimal control model for the transmission dynamics of an infectious disease. We derive several key statistical properties of the proposed model, including its quantile function, moments and moment-generating function, Renyi entropy, and order statistics. The parameters of the distribution are estimated using the maximum likelihood estimation (MLE) method. A comprehensive simulation study is conducted to evaluate the performance of the MLEs based on their bias and root mean square error (RMSE). The optimal control properties of the model are also derived to determine effective disease management strategies. The results demonstrate the practical utility and flexibility of the new distribution. When applied to real-world data, the TI-HT-TLR distribution provides a superior fit compared to several existing distributions. Furthermore, the findings from the optimal control model highlight the critical importance of early intervention in controlling infectious diseases.

T012.25: AN EFFICIENT GENERALIZED ESTIMATOR OF POPULATION MEAN IN TWO-PHASE SUCCESSIVE SAMPLING UNDER RANDOM NONRESPONSE

Ewere, F., Mbegbu, J. I. and Agbebia, C. S.

Department of Statistics, University of Benin, Benin City, Nigeria

Corresponding Email: friday.ewere@uniben.edu

ABSTRACT

This paper develops an efficient generalized estimator of population mean in two-phase successive sampling under random non-response, known as the ratio-regression-exponential estimator.



T013.25: THE ODD TEISSIER KUMARASWAMY DISTRIBUTION: PROPERTIES AND APPLICATIONS

Mbegbu Julian I.; Ekhosuehi N. and Nzei Lawrence C.
Department of Statistics, University of Benin, Benin City, Nigeria
Corresponding Email: julian.mbegbu@uniben.edu

ABSTRACT

In this article, we present a novel distribution on a unit interval, called the Odd Teissier Kumaraswamy (OTKw) distribution for modelling real life time data with unit interval bound. The proposed model is a special case from the Odd Teissier generated family of distributions. Some statistical properties of the proposed model such as the quantile function, the moments and moment generating function, probability weighted moments, Renyi entropy and order statistics are derived. The parameters of the proposed model are estimated through the maximum likelihood estimation method. A simulation study is carried out to examine the performance of the MLEs with respect to their biases, standard errors and root mean square errors (RMSE). Finally, to illustrate the practical importance and flexibility of the proposed distribution in modelling real data application, three data sets were considered, and the results showed that the new distribution performs better than some other known existing distributions. Using this estimator in successive sampling under random non-response, assumptions that the number of non-respondents follows a probability distribution has been used. Corresponding optimum properties like the bias and mean square error has been established and comparison between this proposed estimator with other existing estimators is carried out. An empirical study included for illustration using real-life data and R-software simulation analysis reveals a significant gain in efficiency and dominance over some existing estimation procedures and is therefore deemed to provide a better substitute whenever priority is placed on efficiency.

PAPERS ON STATISTICAL APPLICATIONS



A001.25: A COMPARATIVE ANALYSIS OF SELECTED FORECASTING MODELS USING INFLATION RATE

Onyeka-Ubaka, J. N.¹, Arowolo O. T.² Adeniyi, S. A.¹

¹Department of Statistics, University of Lagos, Nigeria

²Lagos State University of Science & Technology, Ikorodu

Corresponding Email: jonyeka-ubaka@unilag.edu.ng

ABSTRACT

Inflation is a critical economic indicator with far-reaching implications for both policymakers and market participants. Accurate forecasting of inflation is crucial for economic planning, investment decisions, and the overall stability of a nation's economy. The study, therefore evaluates the performance of selected forecasting techniques (ARIMA and GARCH family models) and provide insights into their effectiveness in capturing the complex dynamics of inflation from January 2004 to May 2025. The results show that ARIMA model outperforms other models in short term forecasting while ARIMAX models is better in long term forecasting. Hybrid models, which combine elements of different techniques, aim to harness the strengths of each approach and provide evidence-based inflation forecasts to stakeholders.

A002.25: A COMPARATIVE POWER ANALYSIS OF HETEROSCEDASTICITY DETECTION TESTS USING SIMULATED DATA

¹ Obabire, Akinleye .A ² Olumi, Toba Timothy, ³Agadaga, Gregory Okiemute

Department of Statistics, The Federal Polytechnic, Orogun, Delta State, Nigeria

Corresponding Email: Obabire.akinleye@fepo.edu.ng, obabireakinleye@gmail.com, Corresponding

Email: olumi.toba@fepo.edu.ng

ABSTRACT

Power test of a binary hypothesis is the probability that the test rejects the null hypothesis when alternative hypothesis is true. The statistical power ranges between 0 and 1. As the statistical power increases, the probability of making type II error decreases. Heteroscedasticity occurs mostly because of beneath mistakes in variable, incorrect data transformation, incorrect functional form, omission of important variables, non-detailed model, outliers and skewness in the distribution of one or more independent variables in the model. All analysis were carried out in R statistical package. Five heteroscedasticity tests were selected, which are Park test, Glejser test, Breusch-Pagan test, White test and Goldfeld test, and used on simulated datasets ranging from 20,30,40,50,60,70,80,90 and 100 sample sizes respectively at different level of heteroscedasticity (that is at low level when $\sigma = 0.5$, mild level when $\sigma = 1.0$ and high level when $\sigma = 2.0$). Also, at the significance criterion $\alpha = 0.05$ and 0.10 respectively. However, each test was repeated 1000 times and the percentage of rejection was computed over 1000 trials. For Glejser test, the average empirical type I error rate for the test reject more than expected while Goldfeld has the least power value. Therefore, Glejser test has the highest capacity to detect heteroscedasticity in most especially on simulated datasets.

A003.25: A COMPARISON OF BAYESIAN REGRESSION AND CLASSICAL ORDINARY LEAST SQUARES METHODS FOR MULTIPLE LINEAR REGRESSION USING POST-UTME DATA

C. O. Odijie and N. Ekhosuehi

Department of Statistics, University of Benin, Benin City

Corresponding Email: cyril.odijie@uniben.edu

ABSTRACT

In this paper, the scores of post-UTME candidates of the University of Benin in the 2023/2024 academic session are modelled with ordinary least squares (OLS) method and its Bayesian multiple linear regression. The predictors envisaged include candidates' first choice of the course of study (course1), age at the time of the screening examination (age), male (an indicator variable for male), and utme (Unified Tertiary Matriculation Examination) score. The population involves 13730 students whose required details were completely available. The results were basically similar in the two methods, which should be expected since the whole population size was analyzed and is large enough for a robust inference leading to a convergence in results of the two methods, with JASP default non-informative beta prior ($\alpha = \beta = 1$, $r = 0.354$) used in the Bayesian analysis. All the regression coefficients were statistically significant, $\beta_{\text{course1}} = 0.0002516$, $t = 2.606$, $p = 0.009$; $\beta_{\text{age}} = 0.099$, $t = 3.366$, $p < 0.001$; $\beta_{\text{utme}} = 0.061$, $t = 22.312$, $p < 0.001$; $\beta_{\text{male}} = -0.456$, $t = -3.144$, $p = 0.002$), although they had poor explanatory power on post-UTME scores ($F = 154.952$, $p < 0.001$).

A004.25: ADVANCED MACHINE LEARNING ALGORITHMS FOR CARDIOVASCULAR DISEASE PREDICTION: COMPARATIVE EVALUATION WITH CLINICAL DATA INTEGRATION

Abubakar F. kabara, Ishaq O. Olawoyin, Auwalu Ibrahim, Yakubu Husaini, Hamisu A. Hamisu, Salma S. Suleiman, Nafisat Yusuf, Fatima U. Muhammad, Umar A. Huguma and James S. Ogheneovo

Aliko Dangote University of Science and Technology, Wudil, Nigeria

Corresponding Email: babington4u@gmail.com

ABSTRACT

Cardiovascular diseases (CVDs), particularly hypertension, are among the leading causes of mortality globally. This study explores the predictive accuracy of various machine learning (ML) algorithms in identifying hypertension risk using clinical and lifestyle data from 1,000 participants in Kano State, Nigeria. Fourteen ML models, including ensemble methods, deep learning, and classical classifiers, were evaluated using precision, recall, F1-score, and AUC. The stacking ensemble achieved the highest accuracy (87.5%) and balanced performance across metrics. Results emphasise the need for early, data-driven interventions and the relevance of ML in clinical decision support for cardiovascular risk prediction.



A005.25: COMPARATIVE ANALYSIS OF STOCHASTIC MODELS AND MACHINE LEARNING ALGORITHMS FOR INFLATION RATE PREDICTION IN NIGERIA

Edesiri Bridget Nkemnole and Abiodun Oyelami

Department of Statistics, University of Lagos, Akoka, Lagos, Nigeria.

Corresponding Email: abbeytech4uall@gmail.com

ABSTRACT

Inflation forecasting is a critical aspect of economic planning, particularly in developing economies like Nigeria, where inflation volatility significantly impacts policymaking, investment decisions, and overall economic stability. This study evaluates the predictive performance of traditional stochastic processes such as Vasicek, Cox-Ingersoll-Ross (CIR), and Geometric Brownian Motion (GBM) – against three machine learning algorithms: Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN), in modeling Nigeria's inflation trends. Our findings reveal that while stochastic models capture the theoretical dynamics of inflation, their predictive accuracy is relatively limited compared to machine learning techniques, particularly the Random Forest model which outperform stochastic models in accuracy and robustness and exhibited superior performance across key evaluation metrics. This research advocates for a paradigm shift in Nigeria's economic modeling strategies by emphasizing the integration of advanced machine learning methods into inflation forecasting.

A006.45: COMPARATIVE ANALYSIS ON EFFECTS OF EARLY TREATMENT OF LASSA FEVER AND SYMPTOMS IN NIGERIA

Kole¹, Emmanuel and Ilori², Adetunji K.

1.Department of Mathematics and Statistics, Kaduna Polytechnic, Kaduna

2.Statistics Programme, National Mathematical Centre, Abuja.

Corresponding Email: enkemnole@unilag.edu.ng

ABSTRACT

Lassa fever is an acute viral disease characterized by a broad range of clinical symptoms. This study presents a comprehensive analysis of Lassa fever cases in Nigeria, focusing on demographic patterns, symptomatology, geographic distribution, and factors influencing outcomes of patients. Using secondary data collected from the Nigeria Centre for Disease Control (NCDC) from 2018 to 2021, Geographically, Edo and Ondo States accounted for the majority of cases, contributing 49.1% and 23.6% of recoveries, respectively, highlighting the regional concentration of Lassa fever burden. Factor analysis of symptoms revealed two distinct clusters: Late-stage symptoms, such as bleeding, respiratory distress, and chest pain, which were strongly associated with fatal outcomes, and Early-stage symptoms, including fever, fatigue, and gastrointestinal disturbances, which facilitated initial disease detection. Logistic regression identified early-stage symptoms, late-stage symptoms and advanced age as significant predictors of mortality, while early-stage symptoms shows a less pronounced likelihood of death compared to late-stage symptoms when there is timely intervention. The findings emphasize the importance of early detection of symptoms and treatment as critical strategies to mitigate the impact of Lassa fever. Additionally, targeted interventions should focus on capturing Lassa fever cases in rural areas of other regions and prioritize regions with high disease prevalence and vulnerable populations, particularly older adults. Enhanced diagnostic accuracy, timely treatment, and symptom monitoring are crucial to reducing mortality and improving outcomes for Lassa fever patients.



A007.25: COUNTS DATA REGRESSION MODELING OF LASSA FEVER IN NIGERIA IN THE PRESENCE OF OVER-DISPERSION

Abiodun Obaromi,¹ Wadai Mutah,² Omoha Joy, Grace Abakpa³

¹Department of Biostatistics, Federal University of Health Sciences, Otukpo, Nigeria.

²Department of Mathematics, Federal University of Health Sciences, Otukpo, Nigeria.

³Department of Microbiology, Federal University of Health Sciences, Otukpo, Nigeria.

Corresponding Email: obaromi.davies@fuhso.edu.ng

ABSTRACT

To account for over-dispersion by the impact of confirmed laboratory cases, population, population density and average household size on Lassa Fever (LF) death cases in the thirty-seven (37) states of Nigeria and obtained from Nigeria Center for Disease Control (NCDC). This novel study compared robust count models. Various models, including the Poisson regression model; Negative Binomial and Generalized Negative Binomial models were used with the assumed presence of over-dispersion phenomenon in counts data. Modeling with Negative Binomial Regression (NBR) and Generalized Negative Binomial (GNB) models were used to solve the dispersion issues because both approaches accommodated the dispersion parameter. Based on both selection criteria, the NB and GNB performed better than the Poisson model. The findings also indicated that population density, average household size and number of confirmed laboratory cases have positive and significant impacts on number of deaths of Lassa Fever cases in all the states in Nigeria. The government of Nigeria, ought to intensify its efforts to combat the spread, incidence and prevalence of Lassa Fever by identifying, isolating, and treating Lassa Fever cases early in the population hot spots to reduce mortality. Also, more mathematical and statistical modeling of the transmitting dynamics of LF should be intensified.

A008.25: DETERMINANTS OF MINIMUM WAGE COMPLIANCE IN NIGERIAN STATES: A PANEL LOGISTIC REGRESSION ANALYSIS

Abiola T. Owolabi,¹ Zainab T. Abdullahi,¹ Olasunkanmi J. Oladapo,¹
Janet I. Idowu¹ and Kayode Ayinde²

¹Department of Statistics, Ladoke Akintola University of Technology, Ogbomosho, Nigeria.

²Department of Mathematics and Statistics, Northwest Missouri State University,
Maryville, Missouri, USA

Corresponding Email: atowolabi@lautech.edu.ng

ABSTRACT

The implementation of minimum wage laws in Nigeria has faced persistent challenges, often hindered by political, economic, and structural factors. This study investigates the predictors of minimum wage compliance across Nigerian states from 2019 to 2023, aiming to identify key determinants that influence adherence to wage policies. Using a sample of panel data from the National Bureau of Statistics and the "State of States" report, a fixed-effects logistic regression model was applied to analyse the relationship between compliance and explanatory variables, including the ruling party in each state, net allocations, internal and external debt, and internally generated revenue (IGR). The results indicate that political alignment with the federal government, higher internally generated revenue (IGR), and external debt significantly affect compliance, while internal debt and federal allocations do not. These findings suggest that political alignment and a strong internal revenue base play crucial roles in facilitating minimum wage adherence. In conclusion, to promote consistent wage compliance across all states, it is recommended that the federal government provide equitable support to states regardless of political affiliation and establish accountability measures to enforce wage policies. This study contributes to understanding the complex dynamics of wage policy implementation in Nigeria and offers insights for future policy reforms.

A009.25: DYNAMICS OF CRUDE OIL PRICE, PRODUCTION AND EXPORTATION IN NIGERIA (2006 – 2025): A TIME-SERIES ANALYSIS

G. C. Ibeh¹, J. C. Ajaraogu¹ and C. E. Onyenekwe²

¹Department of Mathematics/Statistics, Federal Polytechnic Nekede, Owerri, Imo State, Nigeria

²Department of Statistics, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria

Corresponding Email: gibeh@fpno.edu.ng

ABSTRACT

This study examines the dynamic relationships between crude oil prices, domestic production, and exportation in Nigeria from January 2006 to January 2025 using advanced time-series methodologies. Monthly data sourced from the Central Bank of Nigeria was analysed, encompassing crude oil prices (USD/barrel), production, and export volumes (million barrels/day). A missing data point for April 2023 was addressed using linear interpolation. Seasonal-Trend Decomposition using Loess (STL) revealed underlying trend structures, though statistical tests confirmed no significant seasonal effects across the variables. Stationarity was established through differencing, and a Vector Autoregressive (VAR) model with Granger causality testing found no significant lagged influence of price fluctuations on production. Volatility analysis using a Dynamic Conditional Correlation GARCH (DCC-GARCH) model identified strong persistence in crude oil price and production volatility but no short-term volatility spillovers. Comparative analysis across pre- and post-2014 and COVID-19 periods highlighted structural shifts in volatility patterns and a decline in output and export volumes. The study concludes with policy recommendations aimed at improving resilience in Nigeria's oil sector amidst external market shocks.

A010.25: EDUCATIONAL CURRICULUM DEVELOPMENT FOR DATA SCIENCE AND ANALYTICS: A STRATEGIC IMPERATIVE FOR NIGERIA'S DIGITAL FUTURE

H. R. Bukari¹, M. Kaka¹ and F. W. Usman^{2,1}

¹Department of Statistics, University of Maiduguri, Maiduguri

²Ramat Polytechnic Maiduguri

Corresponding Email: harunbakari@gmail.com

ABSTRACT

The global demand for data-literate professionals continues to surge, driven by advances in artificial intelligence, big data technologies, and digital transformation initiatives. Nigeria is at a critical moment where educational institutions must align their academic offerings with the growing need for data science and analytics skills. This paper examines the current state of data science education in Nigeria, identifies key curriculum gaps, and proposes a structured framework for the development of a comprehensive, industry-aligned curriculum for undergraduate and postgraduate programs. The paper further emphasizes the role of interdisciplinary collaboration, experiential learning, and partnerships with industry stakeholders in creating a sustainable talent pipeline. The goal is to ensure Nigeria is not left behind in the global data revolution.

A011.25: WASTE MANAGEMENT IN NIGERIA: A VECTOR ERROR CORRECTION MODEL ANALYSIS OF POPULATION AND EMPLOYMENT IMPACTS

Shehu Abdullahi

Energy and Environment Statistics Division, National Bureau of Statistics, Abuja.

Corresponding Email: osyshe@gmail.com

ABSTRACT

This study investigates the impact of population and employment on solid waste generation in Nigeria, using a Vector Error Correction Model (VECM) to analyse the short-run and long-run dynamics between these variables. The results confirm a significant long-run relationship between solid waste generation, employment, and population growth, highlighting population growth as a key driver of increased waste generation. Employment, particularly in waste management sectors, shows potential for mitigating some of these effects. The study underscores the urgent need for a holistic and forward-thinking approach to waste management in Nigeria, integrating economic, environmental, and demographic considerations. Policy recommendations include strengthening waste management infrastructure, expanding recycling initiatives, fostering employment-driven waste solutions, and integrating population dynamics into waste management strategies. Formalizing the role of informal waste workers and incentivizing private sector participation in waste recovery can further enhance efficiency and sustainability. By prioritizing waste minimization, economic incentives for recycling, and stronger regulatory enforcement, Nigeria can achieve a cleaner and more sustainable urban environment.

A012.25: FUZZY MODELLING ON MALARIA DISEASE DYNAMICS

Enwere Kelachi Precious and Ogoke Uchenna Petronilla

Department of Statistics, University of Port Harcourt, PortHarcourt

Corresponding Email: kelachi_enwere@uniport.edu.ng

ABSTRACT

Machine learning has led to innovations through advancements in artificial intelligence, and computational systems. However, most machine learning models suffers inaccuracy due to uncertainty, imprecise information or the effective use of extra information, which is often qualitative in its nature. Thus, fuzzy techniques have been integrated into machine learning to deal with those uncertainties. This research work presented some areas related to fuzzy set theory and its applications to fuzzy SEIR-SEI spatial temporal malaria disease model to forecast its dynamics. We consider the transmission rate and recovery rate of malaria as fuzzy numbers. We compute fuzzy basic reproduction number to study the stability of the equilibrium points. The results show that our basic reproduction number for malaria outbreak, $R_0(v)$ value is $0.0049 \geq 1$, this indicates that an epidemic is less likely to occur. However, since the amount of parasite in the blood $v > 4$, then fuzzy disease-free equilibrium cannot be attained. This means that malaria disease is present all year round in Nigeria. In conclusion, fuzzy machine learning techniques such as fuzzy SEIR-SEI model, are more robust to the many and varied types of uncertainty found in real-world problems.

A013.25: MULTIVARIATE APPROACH TO INVESTIGATE RISK FACTORS PREDISPOSING UNDER-FIVE TO SEVERE BRONCHO-PNEUMONIA

S. S. Ahmed, B. Maijama'a and M. O. Adenomom

Department of Statistics and Data Analytics, Nasarawa State University, Keffi Nasarawa State

Corresponding Email: ahmedsule710@gmail.com

ABSTRACT

Severe broncho-pneumonia is a leading cause of morbidity and mortality among under-five children worldwide. Identifying risk factors is crucial for targeted interventions. This work applies Principal Component Analysis to identify the risk factors that lead to prevalence of Broncho-Pneumonia in under-five children. The data used in this study were collected from the health institutions in Minna south, Niger state. Predictor variables which are well-recognized for characterizing broncho-pneumonia were categorized, these includes; Birth Demographic, Birth Milestones, Nutritional Status, Environmental Factor and Socioeconomic Status. The Principal Component Analysis was used to identify critical/risk factors that contribute significantly to the predisposing under-five to severe broncho-pneumonia infant. Our results unveil 6 most significant predictor variables. The critical/risk factors that contribute to the predisposing of this disease in under-five children identified, these are; Malnutrition, Weight of Baby at Birth, Father's Education, Age of Baby, Father's Occupation and Baby's gender. The second objective achieved is the contribution of each risk factor to the development of severe broncho-pneumonia in infants. This was identified in the Component Matrix, where each risk factors shows a significant contribution.

A014.25: PREDICTIVE ANALYTICS FOR EARLY DETECTION OF DISEASE OUTBREAKS IN URBAN SLUMS: A MACHINE LEARNING APPROACH

Joseph A. Akinyemi, Matthew I. Ekum, Olatunji T. Arowolo and Bolarinwa O. Ajala

Department of Mathematical Sciences, Lagos State University of Science and Technology,
Ikorodu. Nigeria

Corresponding Email: akinyemijoseph93@gmail.com

ABSTRACT

Urban slums present significant challenges for public health surveillance due to high population density, inadequate sanitation, and limited access to healthcare. These conditions create fertile ground for the rapid spread of infectious diseases while constraining the capacity for timely detection and response. In such environments, conventional surveillance systems often rely on delayed reporting, leading to reactive rather than preventive interventions. This study proposes a machine learning-based predictive analytics framework to anticipate disease outbreaks in urban slum settings. The framework integrates simulated datasets—reflecting health facility reports, meteorological patterns, socio-demographic indicators, and spatial attributes—constructed to represent the conditions of Makoko, a densely populated informal settlement in Lagos State, Nigeria. Three predictive models—Random Forest, Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM) networks—were developed and evaluated using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC) as performance metrics. XGBoost achieved the highest predictive performance, followed closely by Random Forest and LSTM. The results highlight the potential of predictive modelling, even with simulated data, to inform early-warning systems, guide targeted interventions, and support evidence-based public health policy in resource-constrained urban communities. The approach is adaptable to other vulnerable settings and can be further enhanced as real-time and empirical data become available.



A015.25: PREDICTIVE MODELLING OF HEALTH STATUS AWARENESS AMONG ACADEMIC STAFF IN TERTIARY INSTITUTIONS IN KOGI STATE USING RANDOM FOREST

Adenike Oluwafunmilola Olubiyi¹, Oluwamayowa Opeyimika Olajide², Michael Sunday Olayemi²

¹Department of Statistics, Ekiti State University, Ado-Ekiti, Nigeria

²Department of Statistics, Kogi State Polytechnic, Lokoja, Nigeria

Corresponding Email: oluwamayowaolajide2@gmail.com

ABSTRACT

This study investigates the level of health status awareness among academic staff in selected tertiary institutions in Kogi State, Nigeria. The research employed both descriptive statistical methods and machine learning techniques, specifically the RandomForest classification algorithm to identify key factors influencing health awareness. A structured questionnaire was administered to 316 academic staff, capturing variables such as age, gender, marital status, academic rank, frequency of health checks, access to health facilities, and presence of chronic conditions. Descriptive analysis revealed that a significant proportion of staff engage in regular health checks and physical activity, suggesting moderate to high health awareness levels. The Random Forest model demonstrated strong predictive performance, with an accuracy of 87.2%, precision of 84.5%, recall of 89.3%, and an F1 score of 86.8%. Feature importance analysis showed that frequency of health checks, age group, and academic rank were the most influential predictors of health awareness. The findings underscore the role of individual health behaviours and demographic characteristics in shaping awareness. The study recommends institutional health campaigns targeted at junior academic staff and enhanced access to health facilities. It concludes that machine learning models offer a reliable approach for profiling health awareness and guiding evidence-based interventions in tertiary institutions.

A016.25: SPATIAL ASSESSMENT OF FLOODING IMPACT ON FOOD SECURITY IN COASTAL NIGERIA USING A GEOGRAPHICALLY WEIGHTED REGRESSION APPROACH

Odior, K. A. and Elugwu, F.

Department of Statistics, Delta State Polytechnic, Otefe Oghara

Corresponding Email: odifullness@gmail.com

ABSTRACT

Flooding remains a persistent and widespread natural disaster in Nigeria's coastal areas, largely driven by the effects of climate change and rising sea levels. Its impact on food security, encompassing dimensions such as availability, accessibility, stability, and utilisation, is both significant and complex. This study adopts a spatial econometric framework, employing Multiple Regression Analysis (MRA) alongside Geographically Weighted Regression (GWR) to examine how flooding influences food security across coastal communities. Data were sourced from various flood-affected regions within the study area. Findings indicate that flooding has a considerable adverse effect on food security, with the GWR model revealing that this impact varies geographically and is more pronounced in areas with lower socioeconomic status and limited adaptive capacity. These spatial disparities highlight the necessity for localised interventions. The study further validates the statistical relevance of key influencing variables in both models. Ultimately, the results underscore the urgent need for customised, community-specific strategies to address the negative implications of flooding on food security and to strengthen the resilience and adaptive capacity of vulnerable populations in coastal Nigeria.

A017.25: COMPARISON OF ARIMA-E WITH ARIMA-N PERFORMANCE IN MODELING NIGERIA'S GDP, 1960 – 2024

¹Salisu Shehu Umar, ²Muhammed Adamu Obomeghie and ³Bello Andrew Ojutomori

Department of Statistics, Auchi Polytechnic, Auchi, Edo State, Nigeria

Corresponding Email: rector@auchipoly.edu.ng

ABSTRACT

This paper compared the performance of two variants of the Autoregressive Integrated Moving Average with Exponential Error (ARIMA-E) and that of conventional ARIMA model (ARIMA-N) in modeling Nigeria's Nominal Gross Domestic Product (GDP) 1960 to 2024. GDP is a central macroeconomic indicator used to measure economic performance. Accurately modeling and forecasting GDP is crucial for policymakers, researchers, and financial institutions. The time-plot of Nigeria GDP exhibited exponential growth from year 2005 to 2021 period marked by steady growth with notable fluctuation. The long term memory 1960 - 2000, GDP surged from about 4.2 billion in 1960 to over 500 billion USD 2014. The GDP declined in subsequent years, 2020 432 billion USD before moderately rising. The identified model: ARIMA (1, 1, 1)-N; evaluation metrics MEA (0.85), RMSE (1.12), AIC (210.4), Residual Skew (~ 0) while that of ARIMA (1, 1, 1)-E; MEA (0.78), RMSE (1.05), AIC (205.7), Residual Skew (> 0.5). ARIMA-E shows slightly better performance in terms of error metrics and model fit. However, exponential error assumption may not generalize across all economies time periods. We suggest future studies could explore other non-Gaussian error structures.

A018.25: MODELING INSURANCE CLAIM DATA: AN INFINITE-SUM REPRESENTATION OF AGGREGATE CLAIM APPROVAL DISTRIBUTION

Friday I. Agu

Department of Statistics, University of Calabar, Calabar, Cross River State

Corresponding Email: agu.friday@unical.edu.ng

ABSTRACT

This paper develops and analyses a new compound distribution for modelling aggregate claim approvals in the insurance industry. The claim count is modelled as a zero-truncated Poisson-Binomial distribution, while the claim severity follows the shifted exponential distribution. We derive closed-form expressions for the probability mass function of both approved and unapproved claims, and represent the probability density function of the aggregate approved claim amounts as an infinite sum. Analytical properties of the model are explored, highlighting its flexibility under parameter variation. A saddlepoint approximation is developed for the cumulative distribution and probability density functions, enabling accurate and computationally efficient estimation. Monte Carlo simulation results demonstrate the asymptotic unbiasedness, strong consistency, and stability of the saddlepoint estimator. The proposed method offers a robust and precise tool for modelling aggregate claims, particularly in large-sample or parameter-perturbed settings.

A019.25: THE APPLICATION OF ORDERED LOGIT MODEL ON BLOOD PRESSURE (BP) OF PATIENTS. (A CASE STUDY OF DALHATU ARAF SPECIALIST HOSPITAL LAFIA)

Ehada M. I¹, Rilwanu Y. M.² and Gona.S.I.³

¹Nasarawa Bureau of Statistics, Nasarawa State, Nigeria

²Department of Mathematics/Statistics, Isa Mustapha Agwai I Polytechnic, Lafia

³Department of Health Information Management, Plateau State College of Health Technology, Zawan

Corresponding Email: yakuburilwanu2017@gmail.com

ABSTRACT

This study investigates the application of an ordered logit model to analyse blood pressure (BP) dynamics among patients at Dalhatu Araf Specialist Hospital, Lafia. Utilising data from clinic attendees, the study examines the influence of predictors including gender, health conditions, medication, lifestyle habits, marital status, and employment on BP levels. Categorical variables, summarised in Table 1, identify key independent factors, while Table 2 provides statistical insights into central tendencies and variability for variables such as BMI, age, and smoking habits. Pseudo R-square values (Cox and Snell, Nagelkerke, McFadden) in Table 3 indicate that the model explains 14.6–16.6% of BP variation, with McFadden's measure suggesting a lower explanatory power at 7.5%. The Model Summary (Table 4) reveals discrepancies between expected and observed values, potentially due to model specifications and outliers. Estimated coefficients (Table 5) and odds ratios (Table 6) highlight significant predictors, including sex, smoking status, BP medication, and BMI, while diabetes, age, marital status, exercise, and education show weaker associations. These findings elucidate key factors influencing BP, offering insights for targeted interventions to enhance healthcare management strategies.

A020.25: EMPOWERING THE NEXT GENERATION OF DATA SCIENTISTS: A COMPREHENSIVE CURRICULUM FRAMEWORK FOR DATA SCIENCE AND ANALYTICS EDUCATION

Kamaldeen Olarewaju Adetunji

Department of Mathematics, Kwara State College of Education, Ilorin

Corresponding Email: adetunjikamoldeen@gmail.com

ABSTRACT

In today's data-driven world, the demand for skilled data scientists and analysts is skyrocketing. To meet this demand, educational institutions must develop curricula that equip students with the knowledge, skills, and competencies required to excel in data science and analytics. This paper proposes a comprehensive curriculum framework that integrates theoretical foundations, practical skills, and real-world applications. The framework covers essential topics, including data preprocessing, machine learning, statistical inference, data visualization, and ethics. It also emphasizes hands-on learning, project-based assessments, and collaboration with industry partners. By adopting this framework, educators can prepare students to tackle complex data challenges, drive business insights, and contribute to informed decision-making. This curriculum development initiative aims to bridge the gap between academia and industry, ensuring that graduates are job-ready and equipped to thrive in the rapidly evolving field of data science and analytics.

A021.25: LEVERAGING PREDICTIVE ANALYTICS FOR ENROLLMENT TREND AND CURRICULUM INNOVATION IN NIGERIAN STATISTICS PROGRAMS

Chidimma Florence Ejiofor¹, Ijeoma Eberechukwu Okechukwu²,
Emeka Henry Chukwueloka³

¹Department of Statistics, Federal Polytechnic Oko

²Department of Statistics, Federal Polytechnic Oko

³Mathematics Unit, Nigeria Maritime University, Okerenkoko, Warri, Delta State

Corresponding Email: florence.ejiofor@federalpolyoko.edu.ng

ABSTRACT

Statistics remains a foundation of the fast-growing field of Data Science. Yet, declining enrollment in Statistics programs across Nigerian tertiary institutions threatens the future workforce for data-driven sectors. This study applies predictive analytics to model enrollment trends and explore implications for curriculum development. Building on Ejiofor (2023), which used the Poisson distribution to reveal declining enrollment in the National Diploma Statistics program at Federal Polytechnic Oko, this paper extends the analysis with machine learning methods, particularly XGBoost, to capture national patterns. Due to the lack of comprehensive national data, simulated datasets were employed to approximate real-world trends and illustrate the potential of predictive modeling once complete data become available. Additionally, survey responses from Statistics Department staff across institutions were analyzed to provide insights into perceptions of Statistics education. Findings aim to guide strategies to strengthen programs and support the integration of Data Science into Nigeria's higher education system.

A022.25: A GENERALIZED BRIDGE REGRESSION ON INFECTIOUS DISEASE MODELLING

Enwere Kelachi Precious, Ogoke Uchenna Petronilla and Ethelbert C. Nduka

Department of Statistics, University of PortHarcourt

Corresponding Email: kelachi_enwere@uniport.edu.ng

ABSTRACT

Machine-learning techniques such as Regression have been developed to investigate associations between risk factor and disease in multivariable analysis. The use of regularization technique with regression models such as the Bridge regression can help overcome the problem of collinearity in a dataset thus preventing overfitting. However, with other given values for the shrinkage parameter γ

of the Bridge regression between the bounds $0 < \gamma \leq 2$ and beyond will produces other regularized models to solve collinearity problem in a data set. Therefore, this study proposes a Generalized Bridge regression model that can be used in analysis as well as biopsy on infectious disease modelling. Again, this research work try to model disease incidence based on compartment modelling. A set of secondary data on the incidence of tuberculosis, gotten from referred published research, was used to study the behavior of the generalized model and for comparison. Again, the study employs data gotten from 'World Malaria Report for 2023, by the World Health Organization for Nigeria', which were used to study the behavior of the compartmental SEIR-SEI model. The results show that generalizing the Bridge regression model to include other shrinkage parameters for $\gamma_n \geq 1$ produces superior performances than other competing regression models. Nevertheless, since our $R_0(v)$ value is $0.0049 < 1$, this indicates that an epidemic is less likely to occur. Since $R_0(v) < 1$ and $v \geq 4$, then disease-free equilibrium cannot be attained. This means that malaria disease is present all year round in Nigeria. Our herd immunity threshold decreases to about 0.3582 at $t = 4$ which means that the spread of Malaria infection can be curbed through effective sanitary and protective measures, and vaccination among the populace in other to achieve herd immunity.



A023.25: DEVELOPMENT OF MACHINE LEARNING ALGORITHMS FOR PREDICTING PROSTATE CANCER IN BENUE STATE, NIGERIA

Edwin Hart Ogwuche^{1,2} and Rotimi Felix Afolabi²

¹Department of Biostatistics, Federal University of Health Sciences, Otukpo P.M.B. 145, Nigeria

²Department of Epidemiology & Medical Statistics, Faculty of Public Health, University of Ibadan, Ibadan, Nigeria.

Corresponding Email: ogwucheedwinhart11@gmail.com

ABSTRACT

Prostate cancer is a leading cause of cancer-related morbidity and mortality among men globally, with an especially high burden in sub-Saharan Africa. Predictive models using machine learning algorithms (MLAs) offer great potential for improving early diagnosis and personalised care. However, most MLA-based models are derived from high-income settings, limiting their applicability in low-resource settings like Nigeria. The study aimed to develop supervised machine learning models for predicting the risk of prostate cancer in Benue State, Nigeria. Medical records of 834 patients aged ≥ 40 years were extracted from the three major cancer referral centres in Benue State. Patients' socio-demographic, lifestyle, haematological, and clinical factors were used to develop the prostate cancer predictive supervised MLA-based models: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbours (KNN) and Extreme Gradient Boosting (XGBoost). The dataset was split into 70:30 ratio for training and testing. Using training dataset, models were developed and tuned with a 10-fold cross-validation, and subsequently tested using the testing dataset. Model performance metrics, including Area Under the Curve (AUC), accuracy, F1-score and Matthew Correlation Coefficient (MCC). SHAP was used to explain feature contributions of the best-performing model. Mean age was 66.17 ± 9.73 years, and 65.7% had family history of prostate cancer. The highest proportion with prostate cancer was aged 70-79 years (20.6%). Performance metrics were consistent in both training and testing sets. The RF outperformed other MLAs, with AUC (95% CI) being 74.3% (69.3-79.3) for LR, 73.6% (68.6-78.6) for SVM, 89.2% (84.2-94.2) for RF, 77.5% (72.5-82.5) for KNN, 88.7% (83.7-93.7) for XGBoost. Accuracy metrics were 69.8% (LR), 78.3% (SVM), 85.6% (RF), 69.5% (KNN) and 83.9% (XGBoost). The F1-scores were 74.2% (LR), 78.1% (SVM), 85.9% (RF), 70.5% (KNN) and 85.3% (XGBoost). The MCCs were 38.4% (LR), 59.3% (SVM), 69.4% (RF), 40.6% (KNN) and 68.5% (XGBoost). Age and PSA explained 56.0% of the prostate cancer risk. the developed random forest model indicated age and PSA as the most influential features in predicting the risk of prostate cancer. The model lays a scalable foundation for prostate cancer clinical decision support and public health interventions in resource-limited settings.

A024.25: EVALUATING SPHERICITY TESTS UNDER MULTIVARIATE NORMAL AND HEAVY-TAILED DISTRIBUTIONS: A SIMULATION STUDY OF TYPE I ERROR AND POWER ACROSS DIMENSIONS

A.D. Akinboro and G.M. Oyeyemi

Department of Statistics, University of Ilorin, Ilorin, Nigeria

Corresponding Email: danielakinboro10@gmail.com

ABSTRACT

This study systematically evaluates the performance of four sphericity tests – Mauchly's Test, the Traditional Likelihood Ratio Test (LRT), John's Invariant Test, and the Quasi-LRT – across varying dimensionalities ($p = 2$ to 5), sample sizes, and underlying data distributions. Through extensive

simulation under both multivariate normal and multivariate t-distributions, we assess the empirical Type I error rates and statistical power of each test to provide comprehensive insights into their practical reliability and robustness. Under multivariate normal conditions, Mauchly's test and the Traditional LRT generally maintain nominal Type I error rates and achieve high power for moderate-to-large samples and low dimensions. However, both exhibit inflated Type I error and instability in small samples and higher dimensions, with the LRT particularly vulnerable when eigenvalues approach zero. John's Invariant Test consistently demonstrates strong power and controlled Type I error across most scenarios, outperforming others under deviations from normality. The Quasi-LRT shows promising power in large samples and high-dimensional contexts but suffers from substantial Type I error inflation in small samples, especially under heavy-tailed distributions. When applied to heavy-tailed multivariate t-distributed data, all tests experience degradation in Type I error control, with Mauchly's and the Traditional LRT exhibiting increased liberalness in small samples. In contrast, John's Test and the Quasi-LRT display relative robustness, though none fully maintain nominal error rates. Power analyses reveal that John's and the Quasi-LRT tests retain strong sensitivity across distributions, while the LRT's performance is notably erratic under non-normal conditions. Our findings highlight the nuanced trade-offs between Type I error control and power across testing procedures, emphasizing that no single test is universally optimal. Practitioners are advised to consider sample size, dimensionality, and distributional assumptions when selecting sphericity tests, favoring John's Invariant Test or the Quasi-LRT under non-normal or small-sample conditions. Future research should explore bootstrap and permutation methodologies to enhance reliability, particularly in challenging scenarios.

A025.25: IMPACT OF MONETARY POLICY USING MACHINE LEARNING-BASED COUNTERFACTUAL ANALYSIS

Itiveh F.E¹ and Adams J.M²

¹Department of Statistics, Southern Delta University, Ozoro, Delta State, Nigeria

²Department of Statistics, P.M.B.13, Auchi Polytechnic, Auchi, Edo State, Nigeria

Corresponding Email: fenoitiveh4@gmail.com

ABSTRACT

Nigeria grapples with significant economic challenges, including high inflation, currency volatility, and unemployment, despite various monetary policy measures implemented by the Central Bank of Nigeria (CBN). This study addresses the problem of understanding the actual impact of these policies, as traditional econometric models often fail to capture the complexities of the Nigerian economy. The objectives are to analyze historical policy impacts, develop a machine learning-based counterfactual analysis framework, and provide empirical evidence to guide future monetary policy decisions. Data will be collected from reputable institutions, including the CBN and the National Bureau of Statistics, covering key economic indicators from 1990 to 2024, using a purposive sampling technique to select representative data points. The analysis will integrate traditional econometric methods, such as Vector Autoregression and Error Correction Models, with machine learning techniques, including Regression Trees and Neural Networks. Preliminary findings indicate that the Regression Tree model outperforms the Neural Network in predicting GDP, especially during economic fluctuations. The VAR model highlights significant interdependencies among GDP, inflation, interest rates, and exchange rates, underscoring the complexities of economic forecasting. Recommendations include enhancing policy formulation, refining forecasting models, and establishing regular monitoring mechanisms. This research aims to provide actionable insights for policymakers to effectively navigate Nigeria's economic challenges.



A026.25: IMPROVING AGRICULTURAL EFFICIENCY WITH ARTIFICIAL INTELLIGENCE AND DATA ANALYTICS: AN ANALYSIS OF CORN YIELD PREDICTION

Kenenisa Abdisa Kuse

Laboratoire de Biomathématiques et d'Estimations Forestières, University of Abomey-Calavi, 04
BP 1525, Cotonou, Benin.

Corresponding Email: abdisakenenisa40@gmail.com

ABSTRACT

Agriculture is increasingly faced with climate uncertainty, soil loss, and global food demand pressure. Traditional yield forecasting methods, normally based on past averages, are not well equipped to deal with changing and uncertain environmental conditions. There is a pressing need for advanced, data-driven solutions for enhancing accuracy levels and the contribution to global food security. We examine the application of Artificial Intelligence (AI) and data analytics for improving corn yield forecast. Employing the USDA CropNet dataset integrating multi-modal data like satellite-remote sensing vegetation indices (NDVI), meteorological conditions, and soil properties, we developed and contrasted two predictive models: a comparative baseline multiple linear regression and a sophisticated Long Short-Term Memory (LSTM) neural network. The LSTM was designed to process sequential data with two-layer architecture (64 and 32 units) to learn temporal relationships at monthly time steps. The LSTM model demonstrated an impressive performance improvement over the traditional regression technique. The baseline linear regression model yielded a satisfactory coefficient of determination ($R^2 = 0.85$). Nevertheless, the LSTM model surpassed it by a significant margin, achieving an R^2 of 0.92 and lower Mean Squared Error (MSE of 1.85 compared to 2.35), confirming its superior ability to capture the complex, non-linear connections and time dependencies involved in crop growth cycles. These results confirm that deep learning models, specifically LSTMs, are extremely suitable for agricultural yield prediction because of their ability to learn from time-series data. The research concludes that the accuracy of predictions can be strongly improved by AI-based tools, which is crucial for resilience building in contemporary agricultural systems. Agricultural stakeholders should invest in pilot implementations of AI-based predictive models in precision agriculture platforms.

A027.25: ON THE PREDICTION VARIANCE PERFORMANCES OF SPACE-FILLING DESIGN IN SOME RESPONSE SURFACE MODELS UNDER CUBOIDAL SPACE

J.C. Nwanya¹, H. I. Mbachu², R. Efeizomor³

¹Department of Mathematics & Statistics, Federal University of Technology, Ikot Abasi,

²Department of Statistics, Imo State University, Owerri

³Department of Mathematics & Statistics, University of Delta, Agbor

Corresponding Email: rita.efeizomor@unidel.edu.ng

ABSTRACT

Prediction variance describes the error involved with making a prediction using a response surface model. This study examines the prediction variance performance of Latin Hypercube Designs (LHDs) within second- and third-order response surface models. G-optimality, I-optimality criteria, and Fraction of Design Space (FDS) plots were employed to assess the predictive capabilities and accuracy of LHDs. The findings reveal that LHDs perform better under third-order models when evaluated using the G-optimality criterion, while under the I-optimality criterion, LHDs perform better in second-order models. The FDS plots further indicate that as the number of factors increases, the prediction errors across models become approximately similar.



A028.25: UNEMPLOYMENT AND ITS ROLE IN CRIME RATES IN NIGERIA: AN EMPIRICAL INVESTIGATION USING ARDL

Adigun K. Abimbola¹ Adigun A. Oladele², and Olubiyi O. A³

¹Department of Mathematical and Physical Sciences, Afe Babalola University Ado-Ekiti, Nigeria.

²Department of Securities and Investment, Federal University of Technology, Akure, Nigeria

³Department of Statistics, Ekiti State University Ado-Ekiti, Nigeria.

Corresponding Email: adigunka@abuad.edu.ng

ABSTRACT

This study investigates the relationship between unemployment and crime rate in Nigeria over the period 2012–2022 using various econometric techniques. Descriptive statistics reveal fluctuations in both crime and unemployment rates, with crime rate showing a significant upward trend starting in 2017, coinciding with rising unemployment. The Pearson correlation coefficient indicates a strong, statistically significant positive relationship between the two variables. Stationarity tests, including the Augmented Dickey-Fuller (ADF) test, confirm that the differenced series for both crime rate and unemployment are stationary, making them suitable for further time series analysis. An Autoregressive Distributed Lag (ARDL) model is employed to explore the dynamic interaction between differenced crime and unemployment rates. The model yields significant results, with the lagged crime variable ($p=0.002385$) and differenced unemployment ($p=0.01581$) both significantly influencing current crime levels. Diagnostic tests, including the Breusch-Godfrey Serial Correlation LM Test, show no evidence of serial correlation, further validating the model. The findings suggest a meaningful link between unemployment and crime rate, with important implications for policy formulation aimed at reducing crime through economic interventions. The study recommends that, government at different levels should provide employment opportunities for Nigerians to ease and reduce the crime rate in the country.

A029.25: NIGERIA'S POPULATION GROWTH ANALYSIS USING RIDGE AND XGBOOST REGRESSION

Uchenna J. Okeke¹, Cornelius Ayantse², OlaOluwa S. Yaya² and Evelyn N. Okeke³

Department of Mathematics and Statistics, Taraba State University, Nigeria

Department of Statistics, University of Ibadan, Ibadan, Nigeria

Department of Statistics, Federal University Wukari, Nigeria

Corresponding Email: joseph.okeke@tsuniversity.edu.ng

ABSTRACT

This study aims to analyze the impact that certain socio-economic variables have on the population growth of Nigeria. The study tries to find out variables with negative and positive impacts on the country's population growth, using time series data of 53 years (1970 to 2023). Ridge and XGBoost regression are the machine learning algorithm methods used. From the results, death rate and infant urban population growth, GDP growth and population density had a negative influence on the population growth while the fertility rate, life expectancy, rural population growth, birth rate, net migration and natural change had positive influence on population growth in Nigeria. Our findings show that gaining control of the death rate number of deaths per 1000 people, and fertility rate, number of births per 1000 women in the age range 15 to 44 years, is the secret to managing Nigeria's population growth and recommend that the Nigerian government take advantage of it to either increase, decrease or maintain its population growth.

A030.25: STATISTICAL ANALYSIS OF LIFE EXPECTANCY DATA USING A NEW PARAMETRIC MODEL

F. C. Opone^{1*}, D. Enegesele¹, A. J. Otunomeruke¹, P.E. Omosioni¹, J. C. Ehiwario² and N. Ekhosuehi³

¹Department of Statistics, Southern Delta University, Ozoro, Nigeria

²Department of Statistics, University of Delta, Agbor, Nigeria

³Department of Statistics, University of Benin, Benin City, Nigeria

Corresponding Email: festus.opone@uniben.edu

ABSTRACT

The use of non-parametric models in statistical analysis of life expectancy data has become a tradition in modern literature. This results from the uncertainty of the monotonicity property of the hazard rate function associated to certain time-to-event data. However, if a parametric model can adequately provide a good fit for the data, it becomes a better choice. This paper introduces a novel parametric model based on the weighted harmonic mean transformation. The model's nomenclature is known as the Weighted Harmonic Inverted Exponential (WHIE) Distribution. Useful distributional properties of the distribution are derived, mathematical expressions for seven distinct methods of parameter estimation are presented for the proposed WHIE distribution. To examine a well-suited estimation method for the WHIE distribution, a Monte Carlo simulation experiment is conducted to investigate the level of precision of the unknown parameter estimates of the WHIE distribution for each method. In practice, two real-life data sets representing the life expectancy at birth (male and female) of Nigerian population from 1960 – 2022, are utilized to illustrate the potentials of the WHIE distribution in practical data fittings.

A031.25: LOGISTIC QUANTILE REGRESSION MODEL USING OUTLIER DISTRIBUTED DATA SET: A COMPARATIVE ROBUSTNESS ANALYSIS

Besta Okey Onyegbuchule¹, Chialuka Adline Onyegbuchule², and Joy Chioma Nwabueze³

¹Department of Maths/Statistics, Imo State Polytechnic, Omuma

²Department Mathematics, Alvan Ikoku University of Education Owerri

³Department of Statistics, Michael Okpara University of Agriculture Umudike

Corresponding Email: Author: bokey@imopoly.net

ABSTRACT

Traditional mean regression models may be inadequate when dealing with asymmetric response distributions. In such cases, quantile regression offers a more robust alternative, accommodating outliers and error distribution misspecification by characterizing the entire conditional distribution of the outcome variable. This paper proposes a robust logistic quantile regression model, which involves adding 0.5 and 0.05 differently to the numerator and denominator of the logit link model. The performance of these two models is evaluated and compared using goodness of fit, AIC, skewness, kurtosis, RMSE, and MSE. Results indicate that both models are robust, with the model adding 0.5 to both the numerator and denominator of the logit link function performing slightly better. This study underscores the importance of selecting an appropriate distribution in quantile regression analysis, particularly for modeling skewed and normal data.

A032.25: LEVERAGING AI AND BIG DATA ANALYTICS FOR ENHANCED NATIONAL SECURITY INTELLIGENCE

Taoheedat Alanamu

Department of Mathematics, Kwara State College of Education, Ilorin, Nigeria

Corresponding Email: taoheedatalanamu@yahoo.com

ABSTRACT

National security agencies face increasing challenges in detecting and preventing threats amidst vast amounts of data. This paper explores the application of Artificial Intelligence (AI) and Big Data Analytics in national security intelligence, enabling proactive threat detection, predictive analytics, and informed decision-making. By harnessing AI-driven insights from diverse data sources, security agencies can enhance situational awareness, identify potential risks, and respond effectively to emerging threats. This research discusses the opportunities, challenges, and ethical considerations of integrating AI and Big Data Analytics into national security operations, highlighting the potential for improved intelligence gathering, analysis, and strategic decision-making.

A033.25: MACROECONOMIC SHOCKS AND INCOME GROWTH IN AFRICA: EVIDENCE FROM A PANEL STRUCTURAL VAR ANALYSIS

Umeokwobi Richard, Dalhatu Bello, Adams Joshua, and Anih David

Central Bank of Nigeria

Corresponding Email: roumeokwobi@cbn.gov.ng

ABSTRACT

African economies are particularly vulnerable to macroeconomic volatility due to structural fragilities, external commodity dependence, and weak policy buffers. This study investigates the transmission of macroeconomic shocks and their implications for income growth in 15 selected African economies from 2000 to 2024, using a Panel Vector Auto regression (PVAR) and a Panel Structural VAR framework implemented in Python and R Studio. By integrating time-series dynamics with cross-sectional interdependencies, the PVAR approach provides a robust platform for disentangling the complex interactions among key variables such as GDP, inflation, exchange rates, and GDP per capita growth. Using the PSVAR, we estimated two scenarios of an increase and a decrease in growth in Africa. We found that a positive shock to GDP brings about an initial increase in inflation in Africa, and the exchange rate showed to decline from the inflation shock, which leads to depreciation of the currency in Africa against the dollar. GDP per capita showed to increase from the shock of the exchange rate. On the other hand, a negative shock to GDP in Africa brings about a decline in Africa's GDP, this shock tends to bring about a decline in inflation in Africa. The shock from inflation brought about an increase in the exchange rate, which might be due to the increase in exchange rates from the monetary policies. GDP per capita showed to decline from the exchange rate shock. Which means an increase in GDP in Africa leads to an increase in inflation, a fall in exchange rate, and an increase in GDP per capita, while a decline in GDP brings about a decline in inflation, a rise in exchange rate, and a fall in GDP per capita. This paper recommends that policymakers in African economies should strengthen exchange rate management and build policy buffers to mitigate the adverse effects of inflation and external shocks on income growth. Secondly, given the dominant influence of exchange rate and inflation on GDP per capita, coordinated monetary and fiscal policies are essential to ensure macroeconomic stability and sustained income growth.



A034.25: A RE-APPRAISAL OF THE SUPPLY SIDE DETERMINANTS OF BANKING CREDIT IN NIGERIA

G. U. Ebuh, I. B. Ezike, J. O. Obeta, B. Dalhatu, R. O. Umeokwobi, D. S. Yusuf and B. A. Umemezia

Central Bank of Nigeria

Corresponding Email: dbello2@cbn.gov.ng

ABSTRACT

This study uses monthly data from January 2011 to August 2024 on bank specific, domestic monetary/financial market, and international attributes to empirically re-access the supply side determinants of banking credit in Nigeria. Employing Stepwise Regression and Autoregressive Distributed Lag (ARDL) model, the results show that banks deposit, banks foreign liabilities and exchange rate significantly determine the credit to the private sector in the long and short run, while non-performing loans, budget deficit and inflation only determine credit in the short run.

A035.25: STATISTICAL METHOD FOR THE ANALYSIS OF RESULTS FROM CONTROLLED COMPARATIVE DIAGNOSTIC SCREENING TESTS AND CLINICAL TRIALS FOR A CONDITION IN A POPULATION

¹Matthew Chukwuma Michael; ²Oyeka, I. C. A.; ³Ajibade, Bright; ³Onyiorah, A. A.; ³Efobi, C. C.

¹Department of Mathematics and Statistics, Faculty of Science, University of Delta, Agbor, Delta State, Nigeria.

²Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Anambra State.

³Department of Ophthalmology, Nnamdi Azikiwe University Teaching Hospital, Nnewi, Anambra State, Nigeria

³Department of Haematology and Blood Transfusion, Nnamdi Azikiwe University Teaching Hospital, Nnewi, Anambra State, Nigeria

ABSTRACT

In controlled comparative prospective or retrospective studies involving matched pairs of sample data of subjects in a population, the responses of subjects to a predisposing factor or to a condition may be simply dichotomous, but a broader classification into more than two mutually exclusive groups is finer and more informative. More so, the classical methods of analysing paired or matched pairs (t-test, z-test and chi-square tests) are either based on the assumption of normality or they are unstable with variation in sample sizes. This work developed a new method for the analysis of results from controlled comparative clinical trials in which the outcomes are on mutually exclusive ordinal groups, such that the assumption of normality and stability with sample size cannot be an impediment. The proposed method was tested with data obtained through a retrospective approach from a random sample of 200 malaria patients at the Nnamdi Azikiwe Teaching Hospital, Nnewi. The new method, which is based on proportions, is capable of estimating the probabilities that the responses of subjects who were administered a particular treatment are higher or lower than the responses of subjects who were administered the other treatment. Compared with the traditional chi-square test of independence, the new method produced a higher chi-square statistic value, 36.57, against 16.279 produced by the traditional chi-square test, suggesting that the new method is more efficient and more powerful.



A036.25: Determinants of Financial Inclusion in Nigeria: Evidence from the 2023 Nationally Representative Household Survey on Access to Financial Inclusion

¹Adanegbe Peter, and ²Zakari Isiaka Osheku

¹National Bureau of Statistics (NBS); adapeter2000@yahoo.com

²Smartdata Research Centre Ltd/Gte; zakry003@gmail.com

ABSTRACT

In Nigeria, Financial inclusion is increasingly recognized as a cornerstone of poverty reduction, economic growth, and social development. It enables individuals and households to access and use affordable financial products such as savings, credit, insurance, and payment services, thereby enhancing welfare and resilience. The study employs a logistic regression model to examine the relationship between the dependent variable (financial inclusion) and a set of explanatory variables, including household wealth status, education, place of residence (urban/rural), gender of household head, English literacy, possession of a National Identification Number (NIN), access to a mobile phone, and region (South/North). Regression results from the 2023 A2F survey indicate that socioeconomic factors, identity, and digital access are the strongest predictors of household financial inclusion. Nonpoor households are more than twice as likely to be included (OR=2.29; AME $\approx$ +12.6 pp), while tertiary education and possession of a National Identification Number (NIN) are among the most influential factors (Tertiary: OR $\approx$ 20.3; AME $\approx$ +36 pp; NIN: OR $\approx$ 3.21; AME $\approx$ +17.7 pp). Access to a mobile phone and residing in the South also significantly raise inclusion probability. Policy actions with the highest short-term impact include scaling up NIN enrolment and simplifying KYC linkages, expanding mobile-money agent networks, and targeted outreach to northern regions. Longer-term strategies should include integrating financial literacy into education, women-targeted initiatives, proportionate KYC, and monitoring via experimental rollouts to establish causal impact.

A037.25: Impact of Fuel Subsidy Reforms on Nigerians' Purchasing Power

Nasir Bukola Rasheed

nasconomics@gmail.com; 08037818710

Statistics Intelligence Unit

National Bureau of Statistics, Headquarters, Abuja.

No 1 Wole Olanipekun Street, Off Constitution Avenue, Central Business District, FCT-Abuja.

ABSTRACT

The article investigates the impact of fuel subsidy reform (PMS pump price) on Peoples' Purchasing Power in Nigeria, using Consumer Price Index (CPI) to measure general price level. It relies on secondary data to conduct the analysis. This study uses parametric tools "correlation and simple linear regression analysis" to examine the association between the variables. The findings of the study reveal that both variables in the system of equation had positive relationship that is strong. Empirically, it reveals that changes in PMS pump price over time explains 98.5% variation in CPI, and a unit tweak in PMS price will cause about 1.32-unit change in CPI. Implying that consumer purchasing power is constantly threatening at any given time there is policy change in oil regime in Nigeria. Therefore, the article recommends among other things that policymakers should ensure price stability to preserve consumers' purchasing power, in order to help MSMEs businesses to remain in business.



A038.25: SOCIOECONOMIC STATUS AND PRIMARY HEALTH CARE UTILIZATION IN NASARAWA STATE, NIGERIA

Job Ozoga

ABSTRACT

Primary healthcare (PHC) serves as the foundation of equitable and accessible health systems, particularly in low- and middle-income countries. Despite its importance, utilisation of PHC services in Nigeria remains uneven and is often influenced by socioeconomic disparities. This study examined the relationship between socioeconomic status and the utilisation of primary healthcare services in Nasarawa State, Nigeria. A cross-sectional survey was conducted among 387 respondents drawn from both urban and rural areas using multistage sampling techniques. Data were analysed using descriptive statistics, chi-square tests, and logistic regression in Stata to determine associations between income, education, and residence, sources of health information, decision-making factors, and PHC utilisation. Findings revealed that income level, educational attainment, and place of residence significantly influenced the frequency and pattern of PHC use. Chi-square results showed strong associations between socioeconomic indicators and healthcare-seeking behaviour ($p < 0.05$). Logistic regression further indicated that higher education and urban residence increased the likelihood of utilising PHC facilities, while lower income was associated with reduced access and underutilization. Additionally, sources of information such as radio and community health workers were key determinants of awareness and service uptake. The study concludes that socioeconomic inequalities strongly shape the utilisation of PHC services in Nasarawa State. To enhance equitable access, policymakers should strengthen health education, subsidise services for low-income households, and improve the distribution of facilities in rural areas. Addressing these disparities is critical for achieving universal health coverage and improving population health outcomes in Nigeria.

Keywords: Primary healthcare, socioeconomic status, healthcare utilisation, income disparities, educational attainment, rural-urban differences, health information sources.

A039.25: Statistical Analysis of Retail Prices of Staple Foods in Nigeria Based on National Bureau of Statistics (NBS) Data

¹Nuga Kehinde Adetayo and Òmotola Adeniyi Andrew

¹Federal School of Statistics, Manchok; tayo.ken.nuga@gmail.com

²Federal School of Statistics; ibadan2omotoadeni@gmail.com

ABSTRACT

This paper presents a simple statistical analysis of retail prices of staple foods in Nigeria, utilizing data from the National Bureau of Statistics (NBS) "Selected Food Prices Watch (October 2024)" report. Data collected from 10,534 informants across all 774 local governments in Nigeria. The study examines price trends, regional variations, and key factors influencing food price dynamics in the Nigerian market. The findings provide insights into food security challenges and offer data-driven recommendations for policy interventions.

Keywords: NBS; Nigeria Market; Regional Variations; Retail Prices; Staple Foods

A040.25: STATISTICAL ASSESSMENT OF TRENDS, PREVALENCE AND SOCIO-DEMOGRAPHIC CORRELATES OF SOME SELECTED CANCER TYPES AT UNIVERSITY OF NIGERIA TEACHING HOSPITAL (UNTH), ITUKU-OZALLA, ENUGU

Okanchi, Louis Chukwunweike

Ezeani, Chinyere Maria

Igwenagu, Chinelo Mercy

DEPARTMENT OF STATISTICS, FACULTY OF PHYSICAL SCIENCES, ENUGU STATE
UNIVERSITY OF SCIENCE AND TECHNOLOGY (ESUT), AGBANI, ENUGU STATE

ABSTRACT

Cancer, otherwise known as malignant disease, is a terminal illness spreading across regions and various countries of the world. The most prevalent forms of cancer differ significantly between nations and regions. As a result, the search for sustainable strategies to coping with the increasing burden of cancer has been a growing concern among researchers. This study empirically assessed the behaviour of reported cases of some selected cancer types at the University of Nigeria Teaching Hospital (UNTH), Ituku-Ozalla, Enugu. Specific objectives of the study were to estimate the prevalence rates, show the trends of cancer diseases for the period, determine the major stage at diagnosis for the various cancer diseases, and determine the socio-demographic correlates of the selected cancer types in the study area. The period covered was from 2009-2022. Relevant data were extracted from folders of 471 cancer patients in the hospital. Statistical techniques such as descriptive statistics and binary logistic regression were employed in the analyses of the data set. The descriptive results revealed increasing trends and prevalence of cancer types, while findings from the binary logistic regression identified age at diagnosis as the commonest socio-demographic determinant of cancer burdens. Based on the findings of the study, early presentation of cancer for diagnosis and treatment was recommended as an important step to reducing cancer burdens in the area. It was also recommended that the Government should show more commitment towards fighting this terminal illness by providing facilities for cancer diagnosis and treatment to be made available in the region, while the population should be sensitised on the essence of regular cancer check-ups.

KEY WORDS: Cancer types; Prevalence; Trends; Binary logistic regression

A041.25: STATISTICAL INVESTIGATION OF SOME DETERMINANTS OF HIGH BLOOD PRESSURE (HBP) AMONG WOMEN OF REPRODUCTIVE-AGE IN CHUKWUEMEKA ODUMEGWU OJUKWU UNIVERSITY TEACHING HOSPITAL (COOUTH), AMAKU, ANAMBRA STATE, NIGERIA

Ezeani, Chinyere Maria

Okanchi, Louis Chukwunweike

Igwenagu, Chinelo Mercy

DEPARTMENT OF STATISTICS, FACULTY OF PHYSICAL SCIENCES, ENUGU STATE
UNIVERSITY OF SCIENCE AND TECHNOLOGY (ESUT), AGBANI, ENUGU STATE

ABSTRACT

This study statistically investigated some determinants of High Blood Pressure (HBP) among women of reproductive-age in Chukwuemeka Odumegwu Ojukwu University Teaching Hospital (COOUTH), Amaku, Awka, Anambra State, Nigeria. The research successfully explored the prevalence and determinants of High



Blood Pressure (HBP) among women of reproductive-age in the study area. The specific objectives were to: (i) assess prevalence of HBP, (ii) ascertain the trend of reported cases of HBP; and (iii) identify the key factor(s) responsible for HBP among women of reproductive-age at COOUTH, Amaku in Anambra State, Nigeria. Monthly secondary data from hospital records of COOUTH were used, covering 2003 to 2020. The data includes variables such as age, weight, genetic factors, family history, lifestyle, and dietary habits. Using time plot analysis and Multiple Regression Analysis, the study found an increasing prevalence of HBP, with significant contributing factors including Age, Family History of HBP, Race/Ethnicity, Family History of Cardiovascular Disease, Obesity, and Kidney Problems. Recommendations emphasized continuous monitoring, collaborative health efforts, and targeted interventions to address these determinants and manage HBP effectively.

Keywords: High Blood Pressure, Reproductive-age women, Prevalence, Trend, Multivariate Regression

A042.25: THE IMPLEMENTATION OF MACHINE LEARNING IN THE PREDICTION AND DEVELOPMENT OF A SYSTEM FOR IMPROVED DIET RECOMMENDATION FOR CARDIOVASCULAR DISEASES

¹Turman Henry; ¹Christiana J. Bakut; ¹Vincent Kanwai; and ²Turman Felicia

¹Federal School of Statistics, Manchok

²University of Jos

Corresponding Author's email: turman112h@gmail.com

ABSTRACT

Cardiovascular diseases (CVDs) are the leading cause of death globally, accounting for an estimated 17.9 million deaths in 2019, with heart disease being a major contributor. The aim of this study is to make prediction and develop a system for improved diet recommendation for chronic cardiovascular diseases (heart disease and hypertension) using machine learning. The implementation machine learning is carried out using Python, leveraging popular libraries for data manipulation, visualization, and machine learning. The dataset was obtained from Kaggle and it contains health-related features. The findings from this study underscore the effectiveness and practicality of a machine learning-enhanced prediction and diet recommendation system in managing chronic diseases. By integrating artificial intelligence (AI) and rule-based logic, the system demonstrates the potential to revolutionize prediction, dietary planning and improve health outcomes for individuals at risk of conditions such as diabetes, hypertension, heart disease. Among the three models used for this research are Logistic Regression, Random Forest, and XGBoost they were implemented and evaluated for their ability to predict cardiovascular disease. Among the models XGBoost demonstrated the best performance, achieving an accuracy of 73.00% and an ROC-AUC score of 79.04%, outperforming both Logistic Regression and Random Forest.

Keywords: Prediction, Machine Learning, Diet Recommendation