



PROFESSIONAL STATISTICIANS' SOCIETY OF NIGERIA

THE BOOK OF ABSTRACTS

10TH INTERNATIONAL CONFERENCE

**THEME: STATISTICAL INTELLIGENCE FOR NATIONAL SECURITY
AND GOOD GOVERNANCE IN A DATA-DRIVEN ECONOMY**

VENUE: UNIVERSITY OF ILORIN, ILORIN, NIGERIA

JULY, 2026

Table of Contents

1. Classification of Diabetes Mellitus Incidences for Kano State, by Age, Gender, Diabetic Conditions and Residential Status: Data Mining Approach	6
*Musa Uba Muhammad; Sani Salihu Abubakar; Ibrahim Abdullahi; Salihu Gambo; Yakubu Hussein; Adepoju A. Akeem; Shamsuddeen Ahmad Sabo; Zakariyya Ibrahim Musa and Rabbilu Ado	
2. Study on Exponentiated Chen Mixture Long -Term Survival Model with Application to Medical Data	6
M.K. Suleiman, A. Yakubu, A. Usman, U.K. Abdullahi and I.K. Adamu	
3. Generalised Additive Models for Interpretable Pipeline Degradation Modelling: Regulatory Auditability and Decision Support in Nigerias Oil and Gas Sector	7
Muhammad Bala Maradun, Yakubu Musa, Umar Usman and Abdulkarim Bello	
4. Why Traditional Statistical Models Fail on Nigerian Pipeline Degradation Data: Empirical Evidence and Theoretical Explanation	8
Muhammad Bala Maradun, Yakubu Musa, Umar Usman and Abdulkarim Bello	
5. Artificial Intelligence Models vs. Traditional Statistical Methods: Exploring Linear Relationships	8
Fanen G. Ibi ;Mercy Ene Unogwu;Shakirudeen Adewale Yusuf	
6. The Transmuted Exponentialnew Weibull Pareto Distribution: a Novel Model for Skewed and Heavy-Tailed Lifetime Data.	9
Onwuamaeze Charity Uchenna; Abubakar Muhammad Auwal; Ugah Tobias Ejiofor; Onwuamaeze Robinson Ugwu; Chinonso Michael Eze	
7. Reimagining National Survey, Census Design, and Sampling Frameworks for Credible, Inclusive, and Timely Data	9
K.O. Adetunji; T. Alanamu; E.J. Adefila	
8. Daily Rainfall Prediction in Lagos Using Logistic Regression, Random Forest and Support Vector Machine with Web Development	10
L. A. Oladimeji; S. A. Oke; T. O. Oguntola; O. O. Akinade; A. A. Adebayo	
9. Performance Analysis of an ATM Queuing System Using G/G/1 and G/g/c Queueing Models	10
L. A. Oladimeji; O. O. Akinade; A. A. Muritala; A. A. Adebayo; V. O. Akintunde	
10. A Monte Carlo Comparison of Regression, Propensity Score, and Causal Machine Learning Methods for Estimating the Average Treatment Effect Under Confounding.	11
S. Muhammad; Y. V. Mbaga	
11. Adaptive Bayesian Updating Probabilistic Framework Under Model Misspecifications	11
Afolabi Bilikis Taiye; Garba Mohammed Kabir	
12. Spatial Analysis of Multidimensional Poverty and Its Determinants Across Local Government Areas in Nigeria	12
Isaac Oluwaseyi Ajao; Saheed Lekan Mumini; and Oyedeki I. Osowole	
13. Development of a Desktop-Based Statistical Tool as a Mode of Teaching AIDS in Secondary Schools	12
A.O Adeoye;O.D.Ojo;F.A.Taofeek-Ibrahim;W.K.Abdulraheem;A.T.Oyebode	
14. Effect of Taxation on Economic Growth of Nigeria	13
A.O. Adeoye; O.D. Ojo; O.A. Sanusi; K.O. Badmus; K.W. Abdulraheem; S.A. Azeez	
15. Polynomial Regression Analysis of Reproductive Age Patterns: a Hospital-Based Study in Wudil, Kano State, Nigeria	13
Musa Yahaya; Razaq B. Afolayan; Usman Aliyu Goga; Abdullahi Muhammad Rashad; Iliyasu Abubakar Salisu	
16. Correlation-Filtered Iterative Bayesian Model Averaging for High-Dimensional Survival Prediction and Biomarker Validation in Colorectal Cancer	14
R.K Lamidi; Sanni Ishola Bello	
17. A Modification of Quality Control Chart for a Hyper-Geometrically Distributed Process Data	14
O. I. Ikpan, J. T. Ofem, J. E. Usen, P. A. Azor	
18. Generalized Cook's Distance for Detecting Influential Observations in Two-Way Random Effects Panel Data Models	15
Ridwanullahi, A. T.; Garba, M. K.	
19. Eigen-Weighted Pairwise Influence Score for Influential Diagnostics	15
E. I. Joel; T. E. Ugah	
20. Modeling Between-Study Heterogeneity in Meta-Analyses of Insomnia Treatment Interventions Using Advanced Random-Effects and Meta-Regression Approaches	16
T. Abdulmalik-Jawondo; M.K. Garba	

THE BOOK OF ABSTRACT, 10TH INTERNATIONAL CONFERENCE

21.Assessment of Public Health Awareness and Early Screening Behaviour in the Prevention of Chronic Kidney Disease	16
Oladotun, A.O.; Oladotun, M. O.; Garba, M. K.; Adeoye, A. O.	
22.Frchet Weibull Distribution: Properties and Application	17
M. Saad; Prof. K. E. Lasisi; Dr. A. Bishir; Dr. U. F. Abbas and Rapheal R. Nenlat	
23.Properties and Applications of an Extended Weibul Distribution	17
M. Saad; Prof. K. E. Lasisi; Dr. A. Bishir; Dr. U. F. Abbas and Rapheal R. Nenlat	
24.Cases of Examination Malpractice in Federal University Wukari	18
M. Nuhu, A. Bello	
25.Cox Versus Parametric Survival Models for Diabetes Patient Survival: Evidence from a Nigerian Tertiary Hospital	18
R. O. Dauda; M. K. Garba and W. B. Yahya	
26.Cure Fraction Models Based on Frchet-Weibull Distribution with Application to Cancer Survival Data	19
M. Saad; K. E. Lasisi; A. Bishir; U. F. Abbas and Rapheal R. Nenlat	
27.Modeling Customer Satisfaction in Short-Term Rentals: Evidence from Airbnb Listings in Lagos, Nigeria	19
Y. O. ajiboyee; G. I. Bamidele; G. B. Karokatose; A. A. ayodeji.	
28.Development of a Bayesian Estimation Framework for Skew-Normal Canonical Correlation Analysis with Application to Terrorism Burden and Governance Indicators	20
I. O. Ayodeji; O. V. Osunro	
29.Parametric Survival Models for Estimating Endotracheal Intubation Rates in Infants with Respiratory Failure.	21
Lami Edibo	
30.Logistic Regression Model for Prognosis of Severe Malaria Among Patients in Lagos State, Nigeria: a Retrospective Cross-Sectional Analysis of Patient Records	21
Lami Edibo	
31.The Transmission Nexus of External Debt, Exchange Rate Volatility, and Inflation on Economic Growth: Evidence from Nigeria (19902024).	22
Afolabi N. O., Ayinde Y.O., Hassan F. A.	
32.Machine Learning Algorithms for Predicting Non-Marital Fertility in Nigeria: an Evaluation of the Second Demographic Transition Theory	23
I.D. John; I.O. Ayodeji	
33.Interpretable Statistical Learning for Breakpoint Identification in Nonlinear Compositionalfunctional Relationships: a Hybrid Marsrandom Forest Framework	23
O.J. Ibidoja; L.A. Oladimeji; W.I.W. Rosli; M.K.M. Ali; N. Hamzah	
34.A Study on Weibull-Exponential Power Distribution with Multi-Layer Artificial Neural Network Modeling	24
Ogunde A. A., Adeleye N. A.,Omosigho D. O	
35.Model Adequacy of Smooth Transition Autoregressive Models for Inflation Rate Forecasting	24
Akintunde Mutairu Oyewale	
36.A New Logarithmic Lomaxinverted Weibull Model: Properties, and Application to Neck Cancer Data	25
Ogunde A. A., Omosigho D. O., Adeleye N. A.	
37.A Modified Partitioned Spline-Based Aft Model with Spatial and Interval Frailties: a Simulation Study	25
Peter Enesi Omaku; Musa Ganaka Kubi ; Fabio Mathias Correa	
38.The Development of a Bayesian Estimation Framework for Skew-Normal Canonical Correlation Analysis with Application to Terrorism Burden and Governance Indicators	26
I.O. Ayodeji; O.V. Osunro	
39.Determination of Optimal Window Length in Singular Spectrum Analysis	27
O.K. Agunloye ; T.K. Ajetunmobi	
40.Determination of Optimum Window Length in Singular Spectrum Analysis	27
O. K. Agunloye; T. K. Ajetunmobi	
41.Hierarchical Bayesian Logistic Regression with Adaptive Exclusive Lasso Prior for High-Dimensional Data	28
James Mayaki; Waheed Babatunde Yahya	
42.Robust Hierarchical Systematic Review-Derived Priors for Bayesian Causal Inference	28
Dongnaan G. A.; Adehi M.U.; Oguntade E.S	
43.Comparative Performance of Ensemble and Traditional Machine Learning Models for Breast Cancer Diagnosis	29
Memunat Manzuma; Ahmed Oloduwa Ameen; Mohammed Kabir Garba	

THE BOOK OF ABSTRACT, 10TH INTERNATIONAL CONFERENCE

44. Predicting Inter-Communal Conflict and Banditry Hotspots in Nigeria: a Machine Learning Approach Using Acled Conflict Data	30
I.C Ezeilo ; C. K. Onyekwere	
45. A Bayesian Hierarchical Meta-Regression Framework Integrating Distributed Lag, Nonlinear Exposure-response, and Climate-policy Interaction Effects in Malaria Evidence Synthesis.	31
T. F. Aderemi; M. U. Adehi	
46. Survival Analysis of Volatility of Life Expectancy: a Truncated Two-Parameter Pranav Distribution Approach	31
C.K. Nwankwo; M.U. Adehi; G.A. Dongnaan	
47. Modeling the Sporadic Behavior of Rainfall Time Series Using Ets State Space and Sarima Models in the Limpopo Province, South Africa	32
T.A. Lasisi; O.O. Akanni; S.V. Molautsi; M. Lesaoana; D.K. Shangodoyin	
48. Bayesian Modelling of Family Planning Discontinuation Under Socioeconomic and Environmental Uncertainty in Rural Communities in South East Nigeria	33
E. M. Ikegwu; G. B. Karokatose; M. A. Idris	
49. Comparative Analysis of Arima and Machine Learning Models for Forecasting Crude Oil Prices in Nigeria	34
Afolayan Tawakalt, S. T. Raji, S. O. Jabaru	
50. A New Robust Estimator for Modeling Multiple Linear Regression Model in Presence of Outliers	34
W.B. Yahya; M. Animashahun	
51. Analysis of Covariance (ancova) in Rcbd: a Non-Gaussian Residual	35
W.B. Yahya; O.J. Sholademi; L.O. Mashood; K. Jimoh	
52. Performance Comparison of Count Regression Models Under Simultaneous Overdispersion, Multicollinearity and Outlier Contamination	35
G. M. Oyinloye and A. Bolakale	
53. Firth Penalized Logistic Regression for Predicting Abnormal Neuroimaging Findings Among Headache Patients in North-Eastern Nigeria	36
S.M. Oguiche; F.Y. Sayuti; M.B. Imam; A.Y.F. Jibrin; H.U. Okemmiri	
54. Bayesian Tournament-Level Forecasting Under Structural Expansion: Validation, Calibration, and Monte Carlo Stability Using Fifa World Cup Data	37
Taiwo M. Adegokea ; Oladapo M. Oladoja ; Kazeem B. Kasali ; Adefunke B. Familua ; Waheed B. Yahya	
55. When Can Informative Priors Be Trusted in Bayesian Gaussian Change-Point Analysis? Operating Characteristics, Prior-data Conflict, and Posterior Robustness	38
Taiwo M. Adegoke ; Waheed B. Yahya	
56. Integrated Stochastic Modeling of Parity, Spatial Fertility, Heterogeneity, and Intergenerational Kinship Dynamics in Rural South-East Nigeria	39
E. M. Ikegwu; G. B. Karokatose; M. A. Idris	
57. Development and Evaluation of a Robust Lad-Based New Two-Parameter Estimator for Linear Regression Models	39
T. J. Adejumo; A. T. Owolabi; O. J. Oladapo, J. I. Idowu	
58. Some Improved Generalised Liu Estimators for Linear Regression Models in the Presence of Multicollinearity	40
Okegbade A.I.; 2Job O.; 1Adejumo T.J.; 1Adesina O.A.; 1Okunlola O.O and 1Egbewumi A.O	
59. An Empirical Analysis of Machine Learning Algorithms for Predicting Poverty Vulnerability and Informing Social Protection Policy in Nigeria	40
Musa, Y.Y.; Suleiman, M.K.; Aliyu U; and Musa, I.	
60. An Integrated Security Risk Index for Crime Risk Mapping: a Spatial Statistical Framework with Application to Kidnapping in Nigeria	41
K. O. Olorede; J. S. Olumoh; W. B. Yahya	
61. A Feasible Generalized Ridge Estimator for Linear Regression Under Multicollinearity and Non-Spherical Errors	42
Y. H. Aliyu; M. K. Garba	
62. Agricultural Credit and Sustainable Food Systems in Nigeria: Implications for Agricultural Productivity, Food Security and Public Health	43
M. A. Umar; O. A. Oyegoke; O. O. Oladapo	
63. A Nonlinear Scale Mixture Framework for Statistical Intelligence in Infectious Disease Surveillance	43
M. A. Umar; N. R. Zakariyau; W. B. Yahya	

THE BOOK OF ABSTRACT, 10TH INTERNATIONAL CONFERENCE

64. Predicting Naira Stress: Inflation and External Reserve Adequacy as Early-Warning Indicators in Nigeria	44
G. T. ZEBULUN; L. A. ADEYEYE	
65. A New Weibull-Gamma Distribution: Theory, Estimation and Application's	45
G. O. Suleiman ;W. B. Yahya	
66. Adaptive Bayesian Polynomial Regression via Structured Gaussian Priors	46
Akinlolu Adeseye Olosunde; Oluwafemi Lawal Bisiriyu	
67. Hybrid Robust Estimation Methods for Non-Linear Regression Models	47
Bassa S. Yakura; Abraham Okolo; Emmanuel Torsen; Akinrefon A. Adesupo	
68. Statistical Evaluation of Equity in Access to the Nigerian Education Loan Fund (nelfund) Among Students in Nigerian Tertiary Institutions	48
Saminu Umar, Paul Sunday, Okwuobodo Monday	

Classification of Diabetes Mellitus Incidences for Kano State, by Age, Gender, Diabetic Conditions and Residential Status: Data Mining Approach

*Musa Uba Muhammad; Sani Salihu Abubakar; Ibrahim Abdullahi; Salihu Gambo; Yakubu Hussein; Adepoju A. Akeem; Shamsuddeen Ahmad Sabo; Zakariyya Ibrahim Musa and Rabbilu Ado

Department of Statistics Aliko Dangote University of Science and Technology Wudil, Kano State

**Corresponding Author: Musa Uba Muhammad ; musaubamuhammad@gmail.com*

A precise classification for diabetes mellitus (DM) decides the appropriate treatment and management of its nuisance, particularly in emerging countries like Nigeria. This study intends to employ data mining procedures for classifying the identified diabetes cases by age, gender, diabetic conditions, and residential area in Kano state of Nigeria, based on real-life data collected from public hospitals in the State. A K-mean assessment was employed to cluster the instances, the events classified out of 3062: 2692 (87.9%) non-insulin dependent (NID), 181 (5.91%) insulin-dependent (IND), and 189 (6.17%) gestational diabetes (GTD). The total number of identified diabetes cases was 3062: 1395 males (45.56%) and 1667 females (54.44%). The greater incidence was found in females compared to males, and mostly in cities and towns, rather than in villages (36.5%, 34.2%, and 29.3%, respectively). The highest incidence among the age groups was in the age group 50-69 years, which conveys about 43.9% of the total identified cases. Moreover, the NID situation had the highest cases (87.9%). These details have been of utmost importance to the healthcare authorities and non-governmental establishments for the proper planning and managing the diabetes.

Keywords: *diabetes mellitus; age; gender; diabetics-condition; classification; K-means;*

Study on Exponentiated Chen Mixture Long -Term Survival Model with Application to Medical Data

M.K. Suleiman, A. Yakubu, A. Usman, U.K. Abdullahi and I.K. Adamu

M.K. Suleiman & I.K. Adamu (Waziri Umaru Federal Polytechnic, Birnin Kebbi), A. Yakubu, A. Usman, U.K. Abdullahi (Ahmadu Bello University, Zaria)

**Corresponding Author: Mansur Suleiman Kardi ; mansursuleimankardi@gmail.com*

Abstract Usually in the analysis of lifetime data, it is often assumed that all individuals/Subjects in a study population will eventually experience the event of interest if observed for a sufficiently long period. However, advances in medicine have led to situations where some individuals survive far longer than expected. These individuals are referred to as long-term survivors. This study therefore introduces a long-term survivor model based on the Exponentiated Chen distribution. The model parameters were estimated using Maximum Likelihood estimation technique. To demonstrate the usefulness of the model, a medical dataset was analyzed, and the results were compared with those obtained from a model that does not account for a proportion of long-term survivors.

Keywords: *Survival model, Long-term survivor model, mixture long-term survivor model, Exponentiated Chen Distribution*

Generalised Additive Models for Interpretable Pipeline Degradation Modelling: Regulatory Auditability and Decision Support in Nigerias Oil and Gas Sector

Muhammad Bala Maradun, Yakubu Musa, Umar Usman and Abdulkarim Bello

Usman Danfodiyo University, Sokoto; Federal University, Gusau

**Corresponding Author: Bala Maradun Muhammad ; balamaradun@gmail.com*

Pipeline Remaining Useful Life (RUL) prediction is increasingly addressed through deep neural network architectures that achieve high predictive accuracy. Yet, their inherent opacity creates significant barriers to adoption in safety-critical regulatory environments. This opacity is particularly problematic in Nigerias oil and gas sector, where maintenance scheduling decisions must be defensible under the Petroleum Industry Act (Federal Republic of Nigeria, 2021) and Nigerian Upstream Petroleum Regulatory Commission (NUPRC) enforcement regulations. This paper proposes and evaluates Generalised Additive Models (GAMs) as the interpretable statistical layer within a hybrid neural-statistical framework for pipeline degradation prediction. The GAM receives concatenated neural network embeddings from Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Graph Neural Network (GNN) components, producing the final RUL prediction as a sum of smooth, individually auditable functions of each embedding dimension. Applied to a dataset of 500 Nigerian pipeline segments observed over 21 years, the proposed framework yields strong predictive performance, with the hybrid Neural + Bayesian GAM achieving a Root Mean Square Error (RMSE) of 10.8 days, Mean Absolute Error (MAE) of 7.6 days, and R of 0.93, outperforming the neural-only baseline (RMSE = 18.6 days, R = 0.81). The GAM layer produces partial-effect plots revealing: a strongly nonlinear U-shaped effect of temperature on degradation (estimated degrees of freedom = 4.8, $F = 23.7$, $p < 0.001$); a threshold effect of soil conductivity at approximately 2.3 mS/cm (95% CI: 2.12.5 mS/cm); and a dominant monotone decline in RUL with pipeline age (edf = 2.1, $F = 187.4$, $p < 0.001$). Dynamic GAMs (DGAMs) incorporating hierarchical temporal smooth functions are evaluated for pipeline monitoring data exhibiting structural breaks and temporal non-stationarity. Bayesian GAMs with spike-and-slab priors are applied for uncertainty-aware variable selection and probabilistic prediction interval calibration, achieving calibration errors within 2% compared with 6.66.9 percentage-point under-coverage for neural Monte Carlo dropout. These findings collectively establish GAMs as an essential layer of interpretability, calibration, and auditability within the hybrid framework, enabling compliance with Nigerias evolving transparency requirements for pipeline integrity management.

Keywords: *Generalised Additive Models; Interpretability; Pipeline Degradation; Penalised Splines; Regulatory Compliance; Dynamic GAM; Bayesian GAM; Nigeria; Oil and Gas; Remaining Useful Life.*

Why Traditional Statistical Models Fail on Nigerian Pipeline Degradation Data: Empirical Evidence and Theoretical Explanation

Muhammad Bala Maradun, Yakubu Musa, Umar Usman and Abdulkarim Bello

Usman Danfodiyo University, Sokoto; Federal University, Gusau

*Corresponding Author: Bala Maradun Muhammad ; balamaradun@gmail.com

Parametric regression and survival models have long served as the primary quantitative tools for pipeline integrity management. Yet, their application to Nigerian oil and gas pipeline data consistently produces results that are statistically inadequate and operationally misleading. This paper provides a systematic, empirical, and theoretical analysis of why four widely used traditional approaches linear regression, Cox proportional hazards, ARIMA time-series models, and two-parameter Weibull survival models fail to characterise Nigerian pipeline degradation data adequately. Empirical evidence is drawn from a longitudinal dataset of 500 Nigerian pipeline segments observed over 21 years (2003-2023). Anderson-Darling tests decisively reject normality of failure times ($A = 15.34$, $p < 0.001$), and Kolmogorov-Smirnov tests reject standard distributional assumptions for corrosion rate data ($D = 0.287$, $p = 0.002$). Linear regression explains only 43% of the variance in failure times ($R = 0.43$), with residual patterns indicating systematic model misspecification. Cox proportional hazards models achieve a concordance index of $C = 0.67$, with Schoenfeld residual tests rejecting the proportional hazards assumption ($\chi^2 = 41.2$, $p < 0.001$). ARIMA models exhibit significant residual autocorrelation (Ljung-Box $Q = 23.7$, $p < 0.001$) and explain less than 50% of the variability in degradation ($R^2 < 0.50$). The traditional two-parameter Weibull model produces $RMSE = 21.6$ days, $MAE = 17.3$ days, and $R = 0.82$ for remaining useful life (RUL) prediction substantially inferior to the hybrid neural-statistical benchmark ($R = 0.96$, $RMSE = 8.3$ days, $MAE = 6.1$ days). Temperature demonstrates a U-shaped nonlinear effect on degradation (GAM estimated degrees of freedom = 4.8, $F = 23.7$, $p < 0.001$), whilst soil conductivity exhibits threshold-like regime shifts at 2.3 mS/cm that are invisible to linear models. Geographic failure clustering (Moran's $I = 0.34$, $p = 0.003$) violates the spatial independence assumptions of all four traditional model classes. Together, these findings provide the empirical and theoretical justification for replacing traditional approaches with flexible, hybrid neural-statistical frameworks for Nigerian pipeline integrity management.

Keywords: Traditional Models; Linear Regression; Cox Proportional Hazards; ARIMA; Weibull; Model Failure; Pipeline Degradation; Nigeria; Nonlinearity; Spatial Autocorrelation.

Artificial Intelligence Models vs. Traditional Statistical Methods: Exploring Linear Relationships

Fanen G. Ibi ;Mercy Ene Unogwu;Shakirudeen Adewale Yusuf

Federal University of Lafia;Federal University of Lafia;Federal University of Lafia

*Corresponding Author: Fanen Goodness Ibi ; fanen.ibi@science.fulafia.edu.ng

A notable methodological divide between artificial intelligence (AI) models and traditional statistical techniques has been highlighted by the changing field of predictive analytics. With an emphasis on basic linear regression, this paper investigates each of their capacities for modelling linear relationships. Traditional models use pre-established equations to estimate parameters and strictly follow the rules and assumptions of probability theory. AI models, on the other hand, take a more inquisitive stance, using adaptive algorithms that dynamically identify intricate patterns in datasets. This research uses a consistent dataset to thoroughly assess the performance indicators of both AI and traditional statistical models in the setting of linearity. The study results show that while AI models show equivalent predictive power, and scalability for big and unstructured datasets, traditional regression models perform marginally better with respect to the performance measures employed, especially the Mean Squared Error and Root Mean Squared Error. The results of this study point out the fact that the best methodological option depends on the kind of dataset and the particular goals of the research. Despite their improved ability to navigate the complexities of huge datasets, Artificial intelligence (AI) models performed slightly worse than traditional models, especially in terms of simplicity and clarity. Furthermore, this study adds to the ongoing discussion on the combination of statistical and artificial intelligence techniques in improving predictive analytics.

Keywords: Predictive Analytics; Linear Regression; Artificial Intelligence (AI); Traditional Statistical Models; Performance Indicators

The Transmuted Exponentialnew Weibull Pareto Distribution: a Novel Model for Skewed and Heavy-Tailed Lifetime Data.

Onwuamaeze Charity Uchenna; Abubakar Muhammad Auwal; Ugah Tobias Ejiofor; Onwuamaeze Robinson Ugwu; Chinonso Michael Eze
Department of Statistics, University of Nigeria, Nsukka Nigeria; Department of Statistics & Data Analytics, Nasarawa State University, Keffi Nigeria; Department of Statistics, Enugu State Polytechnic, Iwolo Nigeria.

**Corresponding Author: ONWUAMAEZE UCHENNA CHARITY ; uchenna.onwuamaeze@unn.edu.ng*

Lifetime data arising in reliability engineering, survival analysis, and related applied fields often display features such as skewness, heavy tails, elongation, and complex hazard rate structures that limit the effectiveness of classical probability distributions. To address this challenge, this study introduces a new flexible lifetime model known as the Transmuted ExponentialNew Weibull Pareto (TENWP) distribution, obtained by applying the Transmuted Exponential-G (TEG) family to the New Weibull Pareto distribution. The proposed model is a five-parameter distribution defined on the positive real line and is capable of capturing a wide range of data behaviours, including skewed and heavy-tailed patterns. An important feature of the model is its ability to exhibit a bathtub-shaped hazard rate function, making it suitable for modelling systems characterized by early failures, a period of stability, and eventual wear-out. Several statistical properties of the distribution are derived, including the moments, moment generating function, quantile function, survival function, and hazard function. Parameter estimation is carried out using the maximum likelihood estimation (MLE) method, while a Monte Carlo simulation study is conducted to evaluate the performance of the estimators in terms of bias, variance, and mean squared error under varying sample sizes. The usefulness of the TENWP distribution is further illustrated through applications to two real-life datasets. Comparative analysis using information criteria and goodness-of-fit tests demonstrates that the proposed model outperforms several existing distributions in fitting complex lifetime data.

Keywords: *Transmuted ExponentialNew Weibull Pareto (TE-NWP) Distribution; Bathtub Hazard Rate; Maximum Likelihood Estimation; Lifetime Modelling*

Reimagining National Survey, Census Design, and Sampling Frameworks for Credible, Inclusive, and Timely Data

K.O. Adetunji; T. Alanamu; E.J. Adefila

Kwara State College of Education, Ilorin , P.M.B 1527. Nigeria.7

**Corresponding Author: KAMALDEEN OLANREWAJU ADETUNJI ; karlmarx4@yahoo.com*

*Accurate, inclusive, and timely data underpin evidence-based policy and sustainable development. Yet many low- and middle-income countries rely on outdated census and survey designs with coverage gaps, high costs, and long production cycles. This study applies survey methodology, design-based and model-assisted inference, and adaptive sampling theory to evaluate modern approaches for strengthening national data systems. Through comparative analysis of recent census and survey redesigns in Sub-Saharan Africa and Asia, and simulations of hybrid sampling schemes, we find that adaptive listing and mobile data collection improve coverage of hard-to-reach populations. Rolling surveys and administrative data linkages cut dissemination lag from years to months. Transparent weighting, post-stratification, and uncertainty quantification enhance credibility, while geospatial stratification optimizes cost and precision. The proposed framework shifts from rigid decadal censuses to flexible, integrated systems that combine traditional methods with modern data sources, enabling credible, inclusive, and timely statistics for diverse and dynamic policy needs.

Keywords: *National Survey, Census Design, Sampling Framework, Time Data.*

Daily Rainfall Prediction in Lagos Using Logistic Regression, Random Forest and Support Vector Machine with Web Development

L. A. Oladimeji; S. A. Oke; T. O. Oguntola; O. O. Akinade; A. A. Adebayo

Ladoke Akintola University of Technology, Ogbomoso, Oyo State.

*Corresponding Author: Lukman Abiola Oladimeji ; laoladimeji24@lautech.edu.ng

Accurate rainfall prediction is critical for agricultural planning, flood risk management, and urban infrastructure resilience, particularly in tropical coastal cities like Lagos, Nigeria, where unpredictable rainfall events cause significant socioeconomic disruption. This study evaluates three machine learning classification models Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM) for daily binary rainfall prediction using a 22-year meteorological dataset of 8,314 observations sourced from Visual Crossing. Two features were engineered from the raw data: daily temperature range and a seasonal indicator. To address the asymmetric cost of false negatives in a tropical rainfall context, Youden's J statistic was applied to optimise the decision threshold of each model, prioritising sensitivity over the conventional 0.5 default. RF achieved the strongest overall performance, recording an accuracy of 76.28%, sensitivity of 79.27%, and AUC-ROC of 0.8454, outperforming SVM (AUC-ROC: 0.8215) and LR (AUC-ROC: 0.8001). Humidity, cloud cover, dew point, and visibility emerged as the most consistent predictors across models, while moon phase and wind speed showed negligible importance in all three classifiers. All three trained models were deployed in an interactive R Shiny web application, enabling non-technical users, including farmers, planners, and policymakers, to obtain real-time rainfall predictions from meteorological inputs. This study provides an end-to-end rainfall prediction pipeline tailored to the Lagos environment and demonstrates the practical value of coupling machine learning with accessible deployment frameworks for climate decision-making in developing countries.

Keywords: Rainfall Prediction; Machine Learning; Random Forest; Logistic regression; Support vector machine; Shiny

Performance Analysis of an ATM Queuing System Using G/G/1 and G/g/c Queueing Models

L. A. Oladimeji; O. O. Akinade; A. A. Muritala; A. A. Adebayo; V. O. Akintunde

Ladoke Akintola University of Technology, Ogbomoso, Oyo State.

*Corresponding Author: Lukman Abiola Oladimeji ; laoladimeji24@lautech.edu.ng

Automated Teller Machines (ATMs) are widely used to provide banking services; however, customers still experience congestion and long waiting times, particularly during peak periods. Many previous studies have evaluated ATM performance using classical queueing models without first verifying their underlying assumptions. This study assessed the performance of an ATM queuing system using general queueing models after testing the suitability of conventional models. Data on customer arrivals, service times, and the number of operational ATMs were collected over seven consecutive days at a selected commercial bank. Chi-square goodness-of-fit tests were conducted to examine whether arrival and service processes followed Poisson and exponential distributions, respectively. The results showed that the assumptions of the M/M/1 and M/M/c models were violated on most days; consequently, G/G/1 and G/G/c queueing models were adopted. The estimated performance measures revealed substantial variation throughout the week. Saturday recorded the highest utilization factor ($\rho = 0.93$), average waiting time in queue ($W_q = 15.96$ minutes), and average queue length ($L_q = 6.37$ customers), indicating high congestion. In contrast, Wednesday and Friday exhibited negligible waiting times and queue lengths due to the availability of multiple operational ATMs. The findings demonstrate that ATM availability and variability in customer arrivals significantly influence system performance. The study recommends increasing the number of active ATMs during peak periods to reduce waiting times, improve service efficiency, and enhance customer satisfaction.

Keywords: Queueing theory; ATM system; waiting time; service capacity; system utilization

A Monte Carlo Comparison of Regression, Propensity Score, and Causal Machine Learning Methods for Estimating the Average Treatment Effect Under Confounding.

S. Muhammad; Y. V. Mbagha

Modibbo Adama University Yola

**Corresponding Author: Suleiman N/A Muhammad ; muhdsule90@gmail.com*

Confounding remains a central challenge in observational studies, as failure to account for relevant covariates can lead to biased and misleading estimates of treatment effects. In particular, naive regression models that omit confounding variables often produce distorted conclusions, underlining the need for robust adjustment strategies. While a range of methods including adjusted regression, propensity score techniques, and causal machine learning approaches, have been developed to address confounding, there is limited systematic evidence comparing their performance within a unified framework. This study presents a comparative evaluation of naive regression, adjusted regression, propensity score matching, and Double Machine Learning (DML) using a Monte Carlo simulation design. Synthetic data were generated from a predefined data-generating process with known confounding structure. All methods were evaluated based on their ability to estimate the average treatment effect (ATE), with performance measured using bias, standard deviation, and root mean squared error (RMSE). The results demonstrate that naive regression substantially overestimates the treatment effect due to unaccounted confounding. In contrast, both adjusted regression and DML yield estimates with minimal bias and low RMSE, while propensity score matching reduces bias but exhibits higher variability. Overall, the findings highlight the importance of properly addressing confounding and provide practical insights into the relative performance of traditional and machine learningbased causal inference methods.

Keywords: *Causal inference; Monte Carlo simulation; Average treatment effect; Confounding; Double Machine Learning.*

Adaptive Bayesian Updating Probabilistic Framework Under Model Misspecifications

Afolabi Bilikis Taiye; Garba Mohammed Kabir

Al-Hikmah University, Ilorin, Nigeria; University of Ilorin, Nigeria

**Corresponding Author: Bilikis Taiye Afolabi ; bilikisolaafolabi@gmail.com*

Bayesian inference is widely used for uncertainty quantification, probabilistic learning, and predictive modelling. However, conventional Bayesian updating assumes correct model specification, an assumption often violated in complex real-world applications. Under model misspecification, posterior estimates may become biased, overconfident, and poorly calibrated. This study developed an adaptive divergence-aware Bayesian updating framework to improve inference under misspecified models. A simulation-based methodology using Monte Carlo experiments examined heavy-tailed distributions, contaminated data, nonlinear relationships, omitted latent variables, and distributional mismatches. The proposed approach extends generalized Bayesian inference by incorporating divergence-based discrepancy measures and adaptive learning-rate calibration, allowing posterior updates to adjust to the degree of misspecification. Simulations were implemented in Python, PyMC, and Stan using 1,000 replications with Markov Chain Monte Carlo estimation. Performance was evaluated using posterior bias, root mean squared error, coverage probability, predictive log score, and Wasserstein distance. Compared with conventional Bayesian inference, the proposed framework consistently reduced posterior bias, improved predictive calibration, increased coverage probability, and demonstrated greater robustness to outliers and structural misspecification. The study shows that adaptive divergence-aware updating provides a flexible and computationally feasible approach for robust Bayesian inference in imperfect modelling environments.

Keywords: *Bayesian inference; Model misspecification; Divergence-based learning; Adaptive Bayesian updating; Posterior calibration*

Spatial Analysis of Multidimensional Poverty and Its Determinants Across Local Government Areas in Nigeria

Isaac Oluwaseyi Ajao; Saheed Lekan Mumini; and Oyedele I. Osowole

Federal Polytechnic, Ado-Ekiti; Federal Polytechnic, Ado-Ekiti; University of Ibadan, Ibadan

*Corresponding Author: ISAAC OLUWASEYI AJAO ; softdataconsult@gmail.com

This study examines the spatial distribution and determinants of multidimensional poverty across Nigeria's Local Government Areas (LGAs) using the 2018 Nigeria Demographic and Health Survey. A Multidimensional Poverty Index (MPI) was constructed at the household level following the Alkire-Foster methodology and aggregated to 646 LGAs via spatial linkage of DHS clusters to administrative boundaries. Exploratory spatial analysis revealed strong positive spatial autocorrelation in MPI (Global Moran's $I = 0.694$, $p < 0.001$), with Local Indicators of Spatial Association identifying significant poverty hotspots (18.58% of LGAs) in the Northwest and Northeast, and prosperity coldspots (15.48% of LGAs) in the Southwest. Lagrange Multiplier diagnostics favored the Spatial Error Model (SEM) over the Spatial Lag Model, and SEM substantially outperformed Ordinary Least Squares ($R = 0.916$ versus 0.817), confirming that ignoring spatial dependence distorts both coefficient estimates and the relative importance of determinants notably sanitation, whose effect nearly tripled once spatial structure was accounted for. Geographically Weighted Regression further improved fit ($R = 0.924$) and revealed substantial spatial heterogeneity: literacy's poverty-reducing effect varied more than thirty-fold across LGAs, while water access showed a counterintuitive sign reversal in parts of the Niger Delta, plausibly reflecting water quality issues uncaptured by standard DHS access measures. These findings demonstrate that multidimensional poverty in Nigeria follows a persistent geographic logic that uniform national policies overlook, providing an empirical basis for spatially differentiated poverty-reduction strategies, with priority attention warranted for the Northwest and Northeast.

Keywords: Multidimensional Poverty Index; spatial autocorrelation; Moran's I , LISA; Geographically Weighted Regression; Nigeria

Development of a Desktop-Based Statistical Tool as a Mode of Teaching AIDS in Secondary Schools

A.O Adeoye; O.D.Ojo; F.A.Taofeek-Ibrahim; W.K.Abdulraheem; A.T.Oyebode

Department of Statistics Federal Polytechnic Offa; Department of Statistics Federal Polytechnic Offa; Department of Cyber Security Federal Polytechnic Offa; Department of Statistics Federal Polytechnic Offa; Department of Statistics Federal Polytechnic Offa

*Corresponding Author: AKEEM OLANREWAJU ADEOYE ; akeemwaju@gmail.com

We discovered that many educators lack the training to use complex statistical tools like SPSS, R, or Excel macros. Therefore, there is need for simple tools for teaching of statistics in secondary schools which necessitate the development of a Desktop Based Statistical Tools as Mode of Teaching Aids in Secondary Schools. This program was aimed to simplified teaching of statistics in secondary schools. Python and EEL was used to developed the program and the result is accurate when using this program to solve measure of central tendency and measure of variation, we discovered that this program is accurate, very easy and simple to use compared to other statistical software.

Keywords: Statistical Tools, Teaching Aids, Desktop, Python

Effect of Taxation on Economic Growth of Nigeria

A.O. Adeoye; O.D. Ojo; O.A. Sanusi; K.O. Badmus; K.W. Abdulraheem; S.A. Azeez

Department of Statistics Federal Polytechnic Offa; Department of Statistics Federal Polytechnic Offa; Department of Mathematics Federal Polytechnic Offa ,

*Corresponding Author: Kajola Waliyu Abdulraheem ; abdulraheemwaliyu@gmail.com

This research work focuses on the effect of Taxation on the economic growth of Nigeria was aim to identifying the impact of taxation on economic development in Nigeria. The specific objectives is to fit a regression model and to predict the future occurrence through the fitted model. Data used for this study was gathered from journals of economics and various publication of the Central bank of Nigeria comprising Economic and Financial review, Annual report and Accounts as well as from National Bureau of Statistics. Data was analyzed with the help of regression as a statistical tool, we analyzed the postulated hypothesis which dwelt on the relationship that exist between taxation and Economic development in Nigeria. It was discovered that taxation has a positive relationship with the economic development in Nigeria. The multiple regression and correlation coefficient showing the association between the predictor variables (unemployment rate, inflation and GDP) and tax revenue was 0.872. The adjusted R squared is 0.760, this means that the predictor variables accounted for 76% variation in the taxation which shows that there was high degree of relationship between the variables involved. We thereby recommend a well-defined policy for inter-governmental collaboration, co-operation and co-ordination between different tiers and agencies of government. Efforts should also be made by the government to ensure they channel revenue from taxation towards socio-economic activities that will benefit the tax payers.

Keywords: Taxation, Economic growth, Revenue Regression and Correlation

Polynomial Regression Analysis of Reproductive Age Patterns: a Hospital-Based Study in Wudil, Kano State, Nigeria

Musa Yahaya; Razaq B. Afolayan; Usman Aliyu Goga; Abdullahi Muhammad Rashad; Iliyasu Abubakar Salisu

Federal University Dutsin-Ma; University of Ilorin, Ilorin; Federal Polytechnic Kaura-Namoda

*Corresponding Author: Musa Yahaya ; musayahaya4185@gmail.com

AbstractBackground: The mean reproductive age is an important indicator of fertility patterns and provides valuable evidence for reproductive health planning. This study estimated the mean reproductive age of women in Wudil Local Government Area (LGA), Kano State, Nigeria, using hospital-based birth data. **Methodology:** Data were obtained from maternity records at Wudil General Hospital, while population figures were projected from the 2006 census baseline reported in the Kano State Bureau of Statistics (KSBS) Statistical Yearbook (2014) using an exponential growth rate of 3.6% per annum. Age-specific birth rates (ASBR) were computed for seven five-year age groups (15-49). A quadratic (second-degree) polynomial regression model was fitted using Ordinary Least Squares (OLS), and model adequacy was assessed using Analysis of Variance (ANOVA) at the 5% significance level. The mean reproductive age was estimated from the fitted polynomial regression model. **Results:** The fitted model explained approximately 63% of the variation in fertility rates ($R = 0.630$). Although the overall regression was not statistically significant ($F = 3.406$, $p = 0.137$), the limited sample size (seven age groups; error df = 4) likely reduced statistical power. The fitted model estimated the mean reproductive age at 30.23 years, indicating that childbearing is concentrated around the early thirties. **Conclusion:** Polynomial regression provides a practical and reproducible approach for modelling age-specific fertility patterns and estimating the mean reproductive age from routine health data. The estimated mean reproductive age of 30.23 years is consistent with fertility patterns in high-fertility populations and offers useful evidence for maternal health planning, family planning programmes, and population policy in Kano State and similar settings.

Keywords: Keywords: age-specific birth rate, fertility, mean reproductive age, Nigeria, polynomial regression, Wudil LGA

Correlation-Filtered Iterative Bayesian Model Averaging for High-Dimensional Survival Prediction and Biomarker Validation in Colorectal Cancer

R.K Lamidi; Sanni Ishola Bello

Kwara State University Malete

**Corresponding Author: Rasheed Kehinde Lamidi ; rasheed.lamidi15@kwasu.edu.ng*

High-dimensional genomic data present major challenges for survival analysis due to the large number of predictors relative to the available sample size, leading to model instability and difficulties in identifying reliable prognostic biomarkers. This study builds upon the Correlation-Filtered Iterative Bayesian Model Averaging (Correlation-IBMA) framework and extends its application to colorectal cancer survival prediction while incorporating bootstrap stability assessment and functional enrichment analysis for enhanced biomarker validation. Using the EMTAB colorectal cancer microarray dataset consisting of 10,357 genes from 129 patients, correlation-based filtering reduced the feature space to 175 candidate genes, from which Iterative Bayesian Model Averaging identified a parsimonious 10-gene prognostic signature. Survival prediction performance was evaluated with KaplanMeier analysis, Cox proportional hazards modelling, parametric survival models, Harrell's concordance index and classification-based metrics. The proposed framework demonstrated better separation between high-risk and low-risk patient groups ($\chi^2 = 16.6$, $p < 0.0001$), prediction accuracy (66.2% vs. 62.5%) and concordance index (0.725 vs. 0.663) compared to conventional Iterative BMA. Bootstrap resampling showed high stability in the selection of the most important biomarkers such as FAM65B, RAB7A, and PTRH2, and functional enrichment analysis showed that the key biomarkers were associated with cancer progression in the following pathways: mitochondrial metabolism, intracellular trafficking, endocytic recycling, and nucleotide-binding processes. The study results show that the expanded Correlation-IBMA framework can be used to develop a stable, interpretable and biologically meaningful method to discover biomarkers and predict survival in high-dimensional colorectal cancer genomics and has potential applications in precision oncology.

Keywords: *High-Dimensional Data; Bayesian Model Averaging; Gene Expression; KaplanMeier; Functional Enrichment*

A Modification of Quality Control Chart for a Hyper-Geometrically Distributed Process Data

O. I. Ikpan, J. T. Ofem, J. E. Usen, P. A. Azor

University of Cross River State; Imo State university

**Corresponding Author: Okim Inyang Ikpan ; ikpanlouis@gmail.com*

In this study, a hypergeometric control chart (hg-chart), developed to detect unusual patterns in process data that follow a hypergeometric distribution and are obtained through sampling without replacement, was revised. The original chart incorporated a parameter, a , representing the minimum number of events, into the calculation of its control limits. The inclusion of this parameter can potentially widen the limits, thereby affecting the symmetry of their placement around the process mean. Furthermore, there is a risk that the control limits may be positioned disproportionately on one side of the centerline, resulting in unintended chart patterns and possible misinterpretation of process behavior. Incorrect interpretation of the hg-chart may lead to either Type I or Type II errors, making accurate assessment of the hypergeometric process a critical concern. To address these issues, the control limits of the original chart were modified by setting $a=0$, and zone boundaries were introduced. This modification facilitates the identification of specific chart patterns that may signal process variation or other underlying problems. The resulting modified hypergeometric control chart (mhg-chart) provides organizations and industries with a more effective tool for detecting variations that could indicate process risks or opportunities for improvement. Additionally, the development of dedicated software incorporating the mhg-chart is recommended to support the efficient monitoring and analysis of process data that follows a hypergeometric distribution.

Keywords: *Control chart, hypergeometric process, count data, zone boundaries*

Generalized Cook's Distance for Detecting Influential Observations in Two-Way Random Effects Panel Data Models

Ridwanullahi, A. T.; Garba, M. K.

Department of Statistics, University of Ilorin, Nigeria

*Corresponding Author: Mohammed Kabir Garba ; garba.mk@unilorin.edu.ng

Cook's distance remains one of the most widely used case-deletion diagnostics for locating influential observations in regression analysis, yet its ordinary least squares formulation is not well suited to panel data in which observations are correlated within units and across time. This article develops and evaluates a generalized Cook's distance for two-way random effects panel data models estimated by generalized least squares, extending the classical case-deletion argument to accommodate hierarchical dependence, serial correlation, and heteroscedastic disturbances. The generalized statistic is derived analytically from the block-deletion normal equations using the Sherman-Morrison-Woodbury identity and is expressed in closed form through a generalized hat matrix and whitened residuals. Its performance is assessed through an extensive Monte Carlo experiment spanning balanced and unbalanced panels of varying dimension, AR(1) and ARMA(1,1) error processes, and unit-, time-, and observation-level contamination at 5%, 10%, and 15%. Across every design point examined, the generalized Cook's distance achieved higher true positive rates, negligible false positive rates, and superior F1-scores relative to the classical statistic. The findings support the routine use of covariance-adjusted Cook's distance whenever regression diagnostics are applied to correlated panel data.

Keywords: Cook's distance; Influence diagnostics; Panel data; Generalized least squares; Two-way random effects; Outlier detection

Eigen-Weighted Pairwise Influence Score for Influential Diagnostics

E. I. Joel; T. E. Ugah

Dennis Osadabay University Asaba; University of Nigeria, Nsukka

*Corresponding Author: Israel Eviano Joel ; israel.joel@dou.edu.ng

High-quality information cannot be achieved solely through sophisticated modelling of data. To obtain high-quality information, the data set requires thorough diagnostic sieving capable of identifying anomalous observations prior to analysis. Current influence diagnostics are often inadequate when data sets contain a large number of influential subsets of observations. This paper proposes a spectral framework for the detection of influential subsets in linear regression through the construction of a pairwise influence operator that simultaneously incorporates residual magnitude and leverage structure among observations. The proposed diagnostic is derived from a symmetric influence matrix whose spectral decomposition yields a set of influence modes representing distinct directions of regression perturbation. Unlike classical influence diagnostics, which primarily assess the marginal effect of individual observations, the proposed methodology explicitly models relational influence structures through pairwise interaction components derived from the regression projection space. The resulting pairwise spectral influence components quantify the degree to which observations jointly participate in dominant influence directions, thereby providing a rigorous mechanism for identifying influential pairs. The methodology reveals that influential observations frequently act collectively rather than independently, and that such cooperative perturbations generate coherent spectral signatures that are not adequately captured by conventional single-deletion diagnostics. The major importance of the proposed method is that it transforms influence analysis from marginal diagnostics to relational influence diagnostics. This is crucial because influential observations frequently act jointly rather than individually. Its alignment with previously validated findings glaringly shows that the procedure is effective and efficient enough for identification of influential subsets.

Keywords: Diagnostics; influential subset; outlier; eigenvalues; eigenvectors

Modeling Between-Study Heterogeneity in Meta-Analyses of Insomnia Treatment Interventions Using Advanced Random-Effects and Meta-Regression Approaches

T. Abdulmalik-Jawondo; M.K. Garba

University of Ilorin

**Corresponding Author: Toyibah Abdulmalik-Jawondo ; toyibahmalik@gmail.com*

Meta-analysis is widely used to synthesize evidence from clinical studies, yet substantial between-study heterogeneity can undermine the validity and interpretability of pooled treatment estimates. This challenge is particularly evident in meta-analyses of insomnia interventions, where differences in study design, participant characteristics, intervention protocols, and outcome measures introduce considerable variability. This study investigates heterogeneity in meta-analyses of insomnia treatment interventions using advanced random-effects and meta-regression models to improve the reliability of effect size estimation. A systematic review of studies published between 2000 and 2024 was conducted in accordance with PRISMA guidelines. Standardized effect sizes (Hedges g) were estimated, and heterogeneity was assessed using Cochran's Q , Higgins I^2 , and between-study variance (τ^2). Meta-regression was employed to identify study-level moderators, while robustness was evaluated using sensitivity analysis, influence diagnostics, and publication bias assessment. The results revealed substantial between-study heterogeneity ($Q=1071.13$, $I^2=98.8\%$, and $\tau^2=317.947$). The fixed-effect model produced a pooled treatment estimate of 19.9330.493, whereas the random-effects model yielded a lower estimate of 14.8605.208, highlighting the impact of heterogeneity on treatment effect estimation. Meta-regression explained part of the observed variability, while influence diagnostics identified a small number of studies with disproportionate effects on the pooled estimates. Evidence of publication bias further underscored the need for robust analytical procedures. The findings demonstrate that integrating advanced random-effects modeling with meta-regression and diagnostic analyses provides a more reliable framework for synthesizing heterogeneous clinical evidence, thereby improving the accuracy and interpretability of treatment effect estimates for insomnia interventions.

Keywords: *Meta-analysis; Insomnia; Heterogeneity; Random-effects model; Meta-regression; Treatment effectiveness*

Assessment of Public Health Awareness and Early Screening Behaviour in the Prevention of Chronic Kidney Disease

Oladotun, A.O.; Oladotun, M. O.; Garba, M. K.; Adeoye, A. O.

Department of Chemistry Education, Usmanu Danfodiyo University, Sokoto; Department of Physical Sciences, Al-Hikmah University, Ilorin; Department of Statistics, University of Ilorin; Department of Statistics, Federal Polytechnic Offa

**Corresponding Author: Mohammed Kabir Garba ; garba.mk@unilorin.edu.ng*

Chronic kidney disease (CKD) is an emerging public health challenge associated with late diagnosis, high treatment costs, and poor preventive health-seeking behaviour. This study assessed public health awareness and early screening practices for the prevention of chronic kidney disease among adults attending the University of Ilorin Teaching Hospital (UITH), Ilorin, Nigeria. Specifically, the study evaluated respondents' level of awareness of CKD prevention, knowledge of its risk factors and preventive. A cross-sectional survey research design was adopted for this study. The population comprised adult patients attending selected outpatient clinics at University of Ilorin Teaching Hospital, Ilorin, while a sample size of 300 respondents was selected using multistage sampling technique. It was concluded that there is fairly awareness and knowledge of chronic kidney disease prevention, this had sufficiently translated into low routine preventive screening behaviour. The study recommended intensified structured patient education, improved accessibility and affordability of routine renal screening services, and stronger hospital-based behavioural health interventions.

Keywords: *Health Awareness; Early Screening Practices; Kidney Disease; Preventive Healthcare; Prevention of Chronic Kidney Disease*

Frchet Weibull Distribution: Properties and Application

M. Saad; Prof. K. E. Lasisi; Dr. A. Bishir; Dr. U. F. Abbas and Rapheal R. Nenlat

Federal Polytechnic Bauchi; Abubakar Tafawa Balewa University, Bauchi; Abubakar Tafawa Balewa University, Bauchi; Abubakar Tafawa Balewa University, Bauchi; Federal Polytechnic Bauchi

*Corresponding Author: Mohammed Sa'ad ; mohammedsaadtoro@gmail.com

Recently, significant effort has been dedicated to developing new probability distributions that offer greater flexibility in modeling real-world data. To provide a robust alternative for these applications, this paper introduces an extension of the Weibull distribution using the Frchet Generalized (Fr-G) family of distributions. The resulting compound model, the Frchet Weibull distribution, offers enhanced skewness and flexibility. We derive and analyze its key mathematical and statistical properties, and estimate its unknown parameters using the maximum likelihood estimation method. Finally, the model is applied to real-world dataset and compared against related distributions. The empirical results demonstrate that the proposed extension outperforms the competing models in flexibility and fit for the datasets examined.

Keywords: Frchet G, Weibull Distribution, Maximum Likelihood, Real Life Data

Properties and Applications of an Extended Weibul Distribution

M. Saad; Prof. K. E. Lasisi; Dr. A. Bishir; Dr. U. F. Abbas and Rapheal R. Nenlat

Federal Polytechnic Bauchi; Abubakar Tafawa Balewa University, Bauchi; Abubakar Tafawa Balewa University, Bauchi; Abubakar Tafawa Balewa University, Bauchi; Federal Polytechnic Bauchi

*Corresponding Author: Mohammed Sa'ad ; mohammedsaadtoro@gmail.com

Recently, significant effort has been dedicated to developing new probability distributions that offer greater flexibility in modeling real-world data. To provide a robust alternative for these applications, this paper introduces an extension of the Weibull distribution using the Frchet Generalized (Fr-G) family of distributions. The resulting compound model, the Frchet Weibull distribution, offers enhanced skewness and flexibility. We derive and analyze its key mathematical and statistical properties, and estimate its unknown parameters using the maximum likelihood estimation method. Finally, the model is applied to real-world dataset and compared against related distributions. The empirical results demonstrate that the proposed extension outperforms the competing models in flexibility and fit for the datasets examined.

Keywords: Frchet G, Weibull Distribution, Maximum Likelihood, Real Life Data

Cases of Examination Malpractice in Federal University Wukari

M. Nuhu, A. Bello

Federal University Wukari

*Corresponding Author: Mathias Nuhu ; mathiasnuhu2019@gmail.com

Examination malpractice is one of the biggest challenges that face the education system in Nigeria. The rate at which students in institutions of higher learning commit examination malpractice, despite stiffer penalties, is alarming. This research was conducted to analyze the cases of examination malpractice at Federal University Wukari. The research work aimed at determining whether the reported cases of examination malpractice depend on gender, levels of study, semester, faculty of students involved and to find out whether the number of reported cases of examination malpractice is uniformly distributed over the sessions. Secondary data collected from the examination and record unit, Federal University Wukari was analyzed using statistical software. The findings from this study revealed that students involvement in examination malpractice is not gender-based and that there is no significant difference between the levels of students that are involved in examination malpractice. Also, there is a significant difference between the semesters and faculties of students involved in examination malpractice. Results of the analysis also show that the number of reported cases of examination malpractice is uniformly distributed over the sessions. Base on the findings of the study, it is recommended that strict supervision should be giving to first semesters examinations and that Faculty of Pure and Applied Sciences and Social Sciences should be given more attention during examinations.

Keywords: *Kruskal-Wallis, Mann-Whitney test, Chi-square goodness of fitness test*

Cox Versus Parametric Survival Models for Diabetes Patient Survival: Evidence from a Nigerian Tertiary Hospital

R. O. Dauda; M. K. Garba and W. B. Yahya

University of Ilorin; University of Ilorin Teaching Hospital

*Corresponding Author: ROQEEB OLAMILEKAN DAUDA ; roqeeb.o.dahud@gmail.com

Diabetes mellitus is an increasing public health burden in sub-Saharan Africa, yet robust survival analyses based on real-world clinical data from Nigerian tertiary hospitals remain limited. Although the semi-parametric Cox proportional hazards model is widely used, the suitability of fully parametric survival models is rarely evaluated. This study analysed retrospective records of 219 diabetes patients treated at the University of Ilorin Teaching Hospital (UITH), Nigeria. Of the sampled patients, 43 died while 176 observations were censored. Kaplan-Meier estimation and log-rank tests assessed survival differences by gender, age group, and diabetes type. The proportional hazards assumption was evaluated using Schoenfeld residuals. A Cox proportional hazards model and two accelerated failure time (AFT) models (log-logistic and log-normal) were fitted, with model performance compared using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Log-rank tests revealed significant survival differences across age groups ($p = 0.0421$) and diabetes type ($p = 0.0316$), but not gender ($p = 0.7421$). The proportional hazards assumption was satisfied (global $\chi^2(6) = 6.20$, $p = 0.5162$). Cox regression identified age above 45 years ($HR = 1.932$, $p = 0.002$) and Type 2 diabetes ($HR = 0.344$, $p = 0.020$) as significant predictors of survival. Among the competing models, the log-normal AFT model provided the best fit ($AIC = 426.23$, $BIC = 449.95$), outperforming both the log-logistic and Cox models. Both AFT models indicated a non-monotonic hazard pattern consistent with chronic disease progression. Age and diabetes type were the principal determinants of survival, whereas gender showed no independent prognostic effect. The log-normal AFT model best represented the survival experience of diabetes patients at UITH, highlighting the importance of evaluating alternative parametric models rather than relying solely on the Cox proportional model. Heightened clinical monitoring is recommended for patients older than 45 years and those with Type 1 diabetes.

Keywords: *Diabetes mellitus; Survival analysis; Kaplan-Meier; Cox proportional hazards; Log-normal Model; Accelerated failure time*

Cure Fraction Models Based on Frchet-Weibull Distribution with Application to Cancer Survival Data

M. Saad; K. E. Lasisi; A. Bishir; U. F. Abbas and Rapheal R. Nenlat

Federal Polytechnic Bauchi; Abubakar Tafawa Balewa University, Bauchi; Abubakar Tafawa Balewa University, Bauchi; Abubakar Tafawa Balewa University, Bauchi; Federal Polytechnic Bauchi

**Corresponding Author: Mohammed Sa'ad ; mohammedsaadtoro@gmail.com*

Standard survival models fail when clinical populations contain a cured fraction of individuals who will never experience the event (Kumar, et al., 2016). To address this, we formulate mixture and non-mixture cure models based on the Frchet-Weibull distribution (FWD) and derive their maximum likelihood estimators. Evaluating these models on clinical data shows that both cure formulations significantly outperform the baseline FWD. Specifically, the Frchet-Weibull Distribution, mixture Cure model (FWD-MCM) yields higher cure estimates, indicating superior suitability for modeling cured proportions. Clinically, Treatment Group I exhibited a higher cured fraction than Group II, a finding validated by Kaplan-Meier survival plots. Incorporating cure parameters into the FWD framework dramatically improves statistical fit and clinical insights for cohorts with long-term survivors.

Keywords: *Frchet-Weibull distribution; Mixture cure model; Non-mixture cure model; Maximum likelihood estimation; Kaplan-Meier.*

Modeling Customer Satisfaction in Short-Term Rentals: Evidence from Airbnb Listings in Lagos, Nigeria

Y. O. ajiboye; G. I. Bamidele; G. B. Karokatose; A. A. ayodeji.

Department of Statistics, Yaba College of Technology, Yaba Lagos

**Corresponding Author: Yemisi Olamide Ajiboye ; olamideyemisi06@gmail.com*

This study examines the determinants of customer satisfaction in peer-to-peer accommodation by addressing methodological limitations in existing Airbnb research. Using Airbnb listings from Lagos, Nigeria, three econometric models are compared: a conventional model including service-quality variables, a reduced model excluding endogenous predictors, and an enhanced model incorporating log-transformed pricing and interaction effects. The results show that the high explanatory power of conventional models ($R = 0.783$) is largely driven by mechanical correlations arising from the inclusion of rating components as independent variables. When these variables are excluded, explanatory power declines significantly ($R = 0.041$), indicating that observable listing attributes explain only a limited proportion of customer satisfaction. The enhanced model further reveals diminishing marginal returns of amenities at higher price levels. This study contributes methodologically by demonstrating how model specification can distort empirical findings and provides context-specific insights from an underexplored African market.

Keywords: *Customer Satisfaction; Airbnb; Short-term Rentals; Regression; Heteroskedasticity; Model Specification.*

Development of a Bayesian Estimation Framework for Skew-Normal Canonical Correlation Analysis with Application to Terrorism Burden and Governance Indicators

I. O. Ayodeji; O. V. Osunro

Obafemi Awolowo University, Ile-Ife, Nigeria

*Corresponding Author: Osarieme Vanessa Osunro ; vosarieme@gmail.com

Skew-normal Canonical Correlation Analysis (CCA) has been developed to accommodate asymmetric distributions in the relationship between two sets of variables. However, existing estimation procedures are based on the frequentist approach and rely on likelihood-based optimization. Consequently, they provide only point estimates with limited inference on parameter uncertainty. This study addresses this limitation by developing a Bayesian estimation framework for skew-normal CCA that incorporates prior information and provides full posterior inference. The proposed framework was implemented using the No-U-Turn Sampler (NUTS), an adaptive Hamiltonian Monte Carlo algorithm for efficient posterior sampling. The method jointly estimates the canonical correlation coefficient, canonical vectors, and skewness parameter while providing full posterior inference through posterior distributions and credible intervals. It was illustrated using terrorism burden and governance indicators to investigate the multivariate association between governance and terrorism burden. The maximum a posteriori estimate of the canonical correlation coefficient was 0.780, while the posterior mean was 0.770, indicating a strong association between the two variable sets. The posterior mean of the skewness parameter was estimated as -1.843 , with a 95% credible interval of $(-3.1885, -0.6958)$, providing evidence of significant negative skewness in the data. Among the governance indicators, Political Stability exhibited the highest canonical loading (0.880), whereas Terrorist Incidents (-0.980), Terrorism-related Deaths (-0.954), and Kidnapping Attacks (-0.963) exhibited the largest absolute canonical loadings in the second set, indicating that they contributed most strongly to the first canonical variate. The proposed Bayesian NUTS framework extends existing skew-normal CCA methodology to a fully Bayesian approach that accommodates skewed multivariate data while yielding coherent posterior inference. The framework offers a practical alternative for multivariate analyses in applications where departures from normality are expected. The empirical findings highlight the strong association between governance and terrorism burden, suggesting that improvements in governance, particularly political stability, may contribute to reducing terrorism burden.

Keywords: Bayesian inference; Canonical Correlation Analysis; Skew-normal distribution; No-U-Turn Sampler; Governance indicators; National security

Parametric Survival Models for Estimating Endotracheal Intubation Rates in Infants with Respiratory Failure.

Lami Edibo

Department of Mathematics and Statistics; Federal Polytechnic Idah, Kogi State

**Corresponding Author: LAMI EDIBO ; lamiedibo@gmail.com*

Conventional survival analysis methods such as the Kaplan-Meier estimator, exponential models, and similar techniques assume independent censoring. When this assumption is violated, survival probabilities are overstated, hazard rates are underestimated, and median survival time become biased. Furthermore, the Kaplan-Meier estimator often used in survival analysis have some drawbacks including inability to estimate survivals beyond follow-up time, inability to produce smooth curves because it is a step-function, and instability with sparse data. This study evaluated the performance of parametric survival models such as the Exponential distribution, Weibull distribution and Bernoulli-Exponential (BE) model proposed by Alih (2025) through simulation experiment. Performance metrics such as Integrated Squared Error (ISE) and Akaike Information Criterion (AIC) were estimated for each model at each simulation scenarios. The simulation results showed that the Weibull and BE models are both preferable to the Exponential model for practical survival analysis because they are flexibility and robust under varying hazard structures. Even though the Exponential model performed better when the true hazard was constant, it became poor in performance when the data exhibited more complex hazard behaviour. Between the Weibull and BE models, neither model completely dominated the other across all simulation schemes; rather, each performed better when aligned with the true underlying distribution. Nevertheless, the Weibull model demonstrated slightly better overall robustness due to its consistently competitive performance across all scenarios and its ability to approximate the Exponential distribution as a special case. Real life application was performed in estimation of intubation rates in infants with respiratory failure.

Keywords: *Survival Analysis, Integrated Squared Error, Akaike Information Criterion, Respiratory Failure.*

Logistic Regression Model for Prognosis of Severe Malaria Among Patients in Lagos State, Nigeria: a Retrospective Cross-Sectional Analysis of Patient Records

Lami Edibo

Department of Mathematics and Statistics; Federal Polytechnic Idah, Kogi State

**Corresponding Author: LAMI EDIBO ; lamiedibo@gmail.com*

Severe malaria is major public health burden in Nigeria and accurate identification of clinical symptoms associated with severe malaria is necessary for early diagnosis, timely treatment, and improved patient outcomes. This study aimed to describe the clinical symptom profile of severe malaria and identify symptoms significantly associated with severe malaria among patients managed in Lagos State, Nigeria. A retrospective cross-sectional study was conducted using anonymised patient records obtained from the Lagos State Ministry of Health between 2019 and 2025. A total of 3,566,952 records were analysed. Severe malaria diagnosis, defined using laboratory confirmation and World Health Organization (WHO) severity criteria, served as the outcome variable. Descriptive statistics were used to summarise symptom prevalence. Binary logistic regression analysis was employed to identify symptoms associated with severe malaria, and adjusted odds ratios (ORs) with 95% confidence intervals (CIs) were reported. Among the 3,566,952 patients analysed, 231,566 were diagnosed with severe malaria, corresponding to a prevalence of 6.49%. Fever (69.98%), headache (65.02%), and fatigue (55.00%) were the most frequently reported symptoms, whereas convulsion (4.98%) and Coca-Cola urine (3.00%) were relatively uncommon. All fifteen symptoms demonstrated statistically significant associations with severe malaria ($p < 0.001$). The findings highlight the importance of maintaining a high index of suspicion among clinicians and reinforce existing WHO recommendations regarding the prompt recognition of severe clinical features suggestive of severe malaria.

Keywords: *Severe malaria, Odd ratio, symptom profile, logistic regression*

The Transmission Nexus of External Debt, Exchange Rate Volatility, and Inflation on Economic Growth: Evidence from Nigeria (1990-2024).

Afolabi N. O., Ayinde Y.O., Hassan F. A.

Ogun State Institute of Technology, Igbesa.

**Corresponding Author: Yusuf Olarewaju Ayinde ; ayindeyusuf1@gmail.com*

This study investigates the transmission nexus among external debt, exchange rate, and inflation on economic growth in Nigeria using annual time series data spanning 1990 to 2024. The Autoregressive Distributed Lag (ARDL) bounds testing approach and the Toda-Yamamoto Granger non-causality framework are employed to capture long-run equilibrium, short-run dynamics, and directional causal pathways within an integrated system. The bounds test confirms a long-run cointegrating relationship among the variables (F-statistic = 5.252, $p < 0.10$), with an error correction term indicating that approximately 8.2 percent of disequilibrium is corrected annually. The short-run ARDL estimates reveal a striking temporal asymmetry in the exchange rate-growth relationship: contemporaneous depreciation exerts a strong contractionary effect on GDP (-1.04 percent, $p < 0.01$), while lagged depreciation generates an equally powerful expansionary effect (+1.05 percent, $p < 0.01$), demonstrating a complete reversal within one period. External debt exhibits no significant direct effect on GDP in either the short or long run. The Toda-Yamamoto causality results map a clear transmission hierarchy in which external debt emerges as the exogenous driver, Granger-causing both inflation ($p = 0.040$) and the exchange rate ($p = 0.018$), while inflation acts as the central transmission belt, Granger-causing both GDP ($p = 0.008$) and the exchange rate ($p = 0.006$). GDP and the exchange rate are locked in a bidirectional feedback loop. The causal architecture reveals that debt affects growth entirely through indirect inflationary and exchange rate channels. The model demonstrates strong explanatory power ($R = 0.986$) and passes critical diagnostic tests for serial correlation and heteroskedasticity. These findings imply that exchange rate adjustments require sequenced safety nets to bridge the contractionary phase, debt accumulation demands mandatory transmission impact assessment, and inflation control must be elevated from a welfare objective to a systemic stability mandate.

Keywords: *External debt, exchange rate, inflation, economic growth, ARDL bounds testing*

Machine Learning Algorithms for Predicting Non-Marital Fertility in Nigeria: an Evaluation of the Second Demographic Transition Theory

I.D. John; I.O. Ayodeji

Obafemi Awolowo University, Ile Ife; Obafemi Awolowo University, Ile Ife

*Corresponding Author: Imoleayo Deborah John ; ioogunnubi@oauife.edu.ng

Non-marital fertility, defined as childbearing outside marriage, is an emerging demographic phenomenon in Nigeria with important implications for social policy and public sector planning. The Second Demographic Transition (SDT) theory links rising non-marital fertility to increasing individualism, changing family formation, and weakening traditional norms. Although SDT is frequently cited as the principal framework for explaining non-marital fertility, its empirical validity has not been tested in the Sub-Saharan African context. This study evaluates the predictive relevance of SDT-related factors for non-marital fertility in Nigeria relative to cultural, behavioural, and socio-demographic characteristics. Data were obtained from the 2024 Nigeria Demographic and Health Survey comprising 13,163 women aged 20-49 years. The outcome variable was non-marital fertility, defined as having the first birth before the first marriage. Machine learning methods such as Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machine were fitted. Model performance was evaluated using Accuracy and the Area Under the Receiver Operating Characteristic Curve (AUROC), while SHAP was used to quantify predictor importance and enhance model interpretability. Results indicated that non-marital fertility accounted for 12% of first births in Nigeria, with the highest prevalence observed in Cross River, Bayelsa, and Rivers States. Gradient Boosting achieved the best predictive performance (Accuracy = 0.87; AUROC = 0.727). SHAP analysis identified ethnicity and religion as the most influential predictors, followed by age at first sex, contraceptive use and age. In contrast, variables associated with SDT exhibited limited predictive importance. The findings demonstrate the value of explainable machine learning for evaluating demographic theories and predicting fertility outcomes. The superior performance of Gradient Boosting highlights their potential for demographic forecasting and evidence-based public sector planning. Furthermore, the limited predictive contribution of SDT-related variables suggests that cultural and behavioural factors should receive greater attention in policies aimed at addressing non-marital fertility in Nigeria.

Keywords: Machine learning; Gradient Boosting; SHAP; Non-marital fertility; Second Demographic Transition; Nigeria.

Interpretable Statistical Learning for Breakpoint Identification in Nonlinear Compositionalfunctional Relationships: a Hybrid Marsrandom Forest Framework

O.J. Ibidoja; L.A. Oladimeji; W.I.W. Rosli; M.K.M. Ali; N. Hamzah

Department of Mathematical Sciences, Faculty of Science, Federal University Gusau, Gusau, Nigeria; School of Health Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia; Department of Statistics, Ladoké Akintola University of Technology Ogbomoso Nigeria; School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia

*Corresponding Author: Olayemi Joshua Ibidoja ; ojibidoja@fugusau.edu.ng

This study proposes a hybrid nonlinear regression framework combining Multivariate Adaptive Regression Splines (MARS) and Random Forest (RF) to identify data-driven compositional breakpoints and model complex nonlinear relationships in multivariate systems. MARS is employed to construct piecewise linear hinge functions that capture structural change points, which are subsequently transformed into derived features and integrated with the original predictor space. RF regression is then used for predictive modelling and nonlinear function approximation. Model performance is evaluated using train-test validation and leave-one-group-out cross-validation. The hybrid model achieves strong predictive accuracy, with $R = 0.8565$ (RMSE = 8.3567) on training data and $R = 0.8977$ (RMSE = 7.7610) on testing data, while LOCOCV yields stable performance ($R = 0.82980.8634$). The MARS component identifies stable nonlinear breakpoints across predictors, including carbohydrate (64.71), crude fibre (0.42), moisture (27.68), fat (0.13), protein (0.13), and total dietary fibre (7.46), indicating regime-switching behaviour in the system. Overall, the proposed framework provides an interpretable and accurate approach for nonlinear breakpoint detection and prediction in complex multivariate regression problems.

Keywords: MARS; Random Forest; breakpoint; Statistical learning; banana sweeteners

A Study on Weibull-Exponential Power Distribution with Multi-Layer Artificial Neural Network Modeling

Ogunde A. A., Adeleye N. A., Omosigbo D. O

Department of Statistics, University of Ibadan; Department of Statistics, Ogun State Institute of Technology, Igbesa Ogun State; Department of Mathematics, Federal Polytechnic Ado-Ekiti.

**Corresponding Author: Adebisi Ogunde ; debiz95@yahoo.com*

We present a novel approach in this study by incorporating the Weibull-Exponential Power distribution (WEPD) into the framework of artificial neural network (ANN) modelling. In order to increase prediction accuracy and interpretability, we incorporate the WEPD into our ANN modeling framework. We intend to gain new insights, improve forecast accuracy, and gain a deeper comprehension of the mechanics underlying our data through this integration. After performing computations for a number of situations using the Hazard Rate Function (HRF), Cumulative Density Function (CDF), Probability Density Function (PDF), and Reliability (R) functions, a procedure for data collection has been developed. With this WEPD and ANN modelling combo, it is expected that the ability to analyze and predict these functions would greatly improve. Two separate artificial neural network models have been created using a total of 76 data sets gathered. The generated multi-layer perceptron network models utilized 15% for model validation, 70% of data for model training, and 15% for model testing. The findings demonstrate that ANNs are highly accurate at predicting the PDF, CDF, HRF, and R functions of the Weibull-Exponential Power model.

Keywords: *Weibull-Exponential Power model; Artificial Neural Network Modeling; Reliability; Multi-layer perceptron.*

Model Adequacy of Smooth Transition Autoregressive Models for Inflation Rate Forecasting

Akintunde Mutairu Oyewale

Department of Statistics, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria

**Corresponding Author: Mutairu Oyewale AKINTUNDE ; waleakintunde2004@gmail.com*

This research investigates the ability of STAR models (Logistic Smooth Transition Autoregressive (LSTAR) and Exponential Smooth Transition Autoregressive (ESTAR) models to analyze and forecast Nigerian monthly inflation from January 2000 to March 2026. Tarsvirta linearity, Tsay nonlinearity, Augmented Dickey-Fuller, Phillips-Perron, Jarque-Bera, Ljung-Box, ARCH-LM and residual nonlinearity tests were employed to determine the non-linear characteristics of Nigerian monthly inflation. The forecast power of the STAR models was assessed using in-sample and out-of-sample measures, such as root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and Theil's U statistic. The analyses revealed strong evidence of significant non-linearity in the Nigerian inflation processes, which indicates that the STAR models (LSTAR and ESTAR) appear to be appropriate for modeling Nigerian inflation. All primary diagnostic and adequacy tests were met for both LSTAR and ESTAR models, providing statistically credible and properly specified approaches for capturing financial and macroeconomic time series phenomena. In general, both models produced reasonably accurate out-of-sample forecast, but the LSTAR model exhibited consistently better fit than the ESTAR, in terms of lower information criteria, and forecast errors for both estimation and evaluation samples. The outcomes show that STAR models provide credible econometric tools for describing regime-dependence, threshold effects, asymmetric and symmetric adjustment, structural changes, etc. It is therefore suggested that STAR modeling approach provides reasonable alternative to linear AR models in forecasting inflation and many other non-linear financial and macroeconomic time series, in an environment undergoing persistent policy changes, structural break and volatility like Nigeria's economy.

Keywords: *LSTAR; ESTAR; nonlinear time series; inflation forecasting; financial time series; economic time series*

A New Logarithmic LomaxInverted Weibull Model: Properties, and Application to Neck Cancer Data

Ogunde A. A., Omosigho D. O., Adeleye N. A.

Department of Statistics, University of Ibadan; Department of Mathematics, Federal Polytechnic Ado-Ekiti; Department of Statistics, Ogun State Institute of Technology, Igbesa Ogun State.

**Corresponding Author: Adebisi Ogunde ; debiz95@yahoo.com*

Here, we study the logarithmic LomaxInverted Weibull (Log-LIW) distribution, a two three-parameter lifetime model within the LomaxG family, developed to provide good flexibility for modeling skewed and heavy-tailed data encountered in reliability and survival analysis. The model used the Exponential distribution as the baseline and incorporates two additional shape parameters. The resulting hazard rate function (HRF) is highly flexible and can exhibit non-monotonic decreasing/increasing shapes based on the parameter values, making the model suitable for a wide range of lifetime data. Fundamental statistical properties are investigated, including the probability density function, cumulative distribution function (CDF), reliability and hazard functions, cumulative and reverse hazard rates, the quantile function, Renyi entropy, and the mean residual life function (MRL). Additional characteristics, including the moment generating function (MGF), skewness, kurtosis, and order statistics, are also investigated. The model parameters are estimated via maximum likelihood. The practical usefulness of the Log-LIW distribution is illustrated using a real dataset. Goodness-of-fit comparisons, based on log-likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Kolmogorov-Smirnov (KS), Anderson-Darling (AD), and Cramer-von Mises (CvM) statistics together with their corresponding p values, are performed against nine competing models. The results consistently show that the proposed Log-LIW distribution outperforms the competing models, confirming its suitability as a flexible and reliable alternative for modeling lifetime data.

Keywords: *Logarithmic Lomax-Inverted Weibull distribution; Renyi entropy; moment generating function; hazard rate function.*

A Modified Partitioned Spline-Based Aft Model with Spatial and Interval Frailties: a Simulation Study

Peter Enesi Omaku; Musa Ganaka Kubi ; Fabio Mathias Correa

University of the Free State and The Federal Polytechnic Nasarawa; Abubakar Tafawa Balewa University; University of the Free State

**Corresponding Author: Peter Enesi Omaku ; omakupete@gmail.com*

Abstract This study proposes and evaluates a semi-parametric Bayesian Accelerated Failure Time (AFT) survival modeling framework that integrates spline-based temporal effects with spatial and interval random effects to address several demographic and health studies across the globe. Using simulated data at increasing sample sizes ($n = 250, 500, 750, 1000$), and a survey data set on Age to the use of contraceptives in Nigeria, where an in sampling of 5000, 10,000 and 50,000 samples were randomly collected for scenarios of 25%, 50%, and 75% censoring to evaluate the proposed model stability for the life data. We compared multiple model specifications, ranging from classical Weibull AFT to Piecewise Exponential Models (PEM) with spatial (ICAR) and temporal flexibility. Model comparisons based on Deviance Information Criterion (DIC), Watanabe Akaike Information Criterion (WAIC), and Conditional Predictive Ordinate (CPO) demonstrated that the proposed flexible spline-based AFT models with interval and spatial frailties (M4 and M5) consistently outperform all alternatives, particularly in larger samples where Spatial risk maps show increasing precision and clearer regional clustering patterns. These findings highlight the advantage of flexible Bayesian survival models in capturing complex temporal and spatial heterogeneity in public health research. We recommend the adoption of spline-enhanced flexible AFT models with spatial random effects in survival analysis where geographic disparities and non-linear time effects are prominent.

Keywords: *Bayesian survival analysis; Accelerated Failure Time; spatial frailty; piecewise exponential model; Nigeria.*

The Development of a Bayesian Estimation Framework for Skew-Normal Canonical Correlation Analysis with Application to Terrorism Burden and Governance Indicators

I.O. Ayodeji; O.V. Osunro

Department of Statistics, Obafemi Awolowo University

**Corresponding Author: Idowu Oluwasayo Ayodeji ; ioayodeji@oauife.edu.ng*

Skew-normal Canonical Correlation Analysis (CCA) has been developed to accommodate asymmetric distributions in the relationship between two sets of variables. However, existing estimation procedures are based on the frequentist approach and rely on likelihood-based optimization. Consequently, they provide only point estimates with limited inference on parameter uncertainty. This study addresses this limitation by developing a Bayesian estimation framework for skew-normal CCA that incorporates prior information and provides full posterior inference. The proposed framework was implemented using the No-U-Turn Sampler (NUTS), an adaptive Hamiltonian Monte Carlo algorithm for efficient posterior sampling. The method jointly estimates the canonical correlation coefficient, canonical vectors, and skewness parameter while providing full posterior inference through posterior distributions and credible intervals. It was illustrated using terrorism burden and governance indicators to investigate the multivariate association between governance and terrorism burden. The maximum a posteriori estimate of the canonical correlation coefficient was 0.780, while the posterior mean was 0.770, indicating a strong association between the two variable sets. The posterior mean of the skewness parameter was estimated as -1.843 , with a 95% credible interval of $(-3.1885, 0.6958)$, providing evidence of significant negative skewness in the data. Among the governance indicators, Political Stability exhibited the highest canonical loading (0.880), whereas Terrorist Incidents (-0.980), Terrorism-related Deaths (-0.954), and Kidnapping Attacks (-0.963) exhibited the largest absolute canonical loadings in the second set, indicating that they contributed most strongly to the first canonical variate. The proposed Bayesian NUTS framework extends existing skew-normal CCA methodology to a fully Bayesian approach that accommodates skewed multivariate data while yielding coherent posterior inference. The framework offers a practical alternative for multivariate analyses in applications where departures from normality are expected. The empirical findings highlight the strong association between governance and terrorism burden, suggesting that improvements in governance, particularly political stability, may contribute to reducing terrorism burden.

Keywords: *Bayesian inference; Canonical Correlation Analysis; Skew-normal distribution; No-U-Turn Sampler; Terrorism burden; Governance indicators*

Determination of Optimal Window Length in Singular Spectrum Analysis

O.K. Agunloye ; T.K. Ajetunmobi

Obafemi Awolowo University, Ile-Ife, Nigeria

*Corresponding Author: Oluokun Kasali Agunloye ; kagunloye@oauife.edu.ng

Determination of optimal window length required for the implementation of singular spectrum analysis has remained a challenging task for researchers and empirical analysts given the array of candidate window lengths available for consideration. This window length is the most crucial parameter that determines the structure of the trajectory matrix on which the entire singular value decomposition is performed. An inappropriate choice of the window length may result in poor signal separation, unstable decomposition, information loss, and inaccurate reconstruction. Existing methods for selecting the window length are largely based on heuristic rules, visual inspection of eigentriples, or repeated trial-and-error procedures using different forecast accuracy measures such as the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). These approaches are subjective, computationally expensive, and often require a validation dataset. In this paper, we propose optimal window-length index (OWLI) as a criterion to determine the optimal window length among a set of candidate window lengths. The proposed criterion is an optimization approach that evaluates the intrinsic quality of the trajectory matrix produced by each candidate window length. We use real-life data to illustrate the application of our proposed criterion. We demonstrate that our proposed criterion out-performed the existing method used for the selection of window length in the literature

Keywords: *Singular Spectrum Analysis (SSA); Singular Value Decomposition (SVD); Grouping; Reconstruction; Embedding dimension or window length; Trajectory Matrix*

Determination of Optimum Window Length in Singular Spectrum Analysis

O. K. Agunloye; T. K. Ajetunmobi

Obafemi Awolowo University

*Corresponding Author: Toyin Kayode AJETUNMOBI ; tkajetunmobi@oauife.edu.ng

Determination of optimal window length required for the implementation of singular spectrum analysis has remained a challenging task for researchers and empirical analysts, given the array of candidate window lengths available for consideration. This window length is the most crucial parameter that determines the structure of the trajectory matrix on which the entire singular value decomposition is performed. An inappropriate choice of the window length may result in poor signal separation, unstable decomposition, information loss, and inaccurate reconstruction. Existing methods for selecting the window length are largely based on heuristic rules, visual inspection of eigentriples, or repeated trial-and-error procedures using different forecast accuracy measures such as the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). These approaches are subjective, computationally expensive, and often require a validation dataset. In this paper, we propose optimal window-length index (OWLI) as a criterion to determine the optimal window length among a set of candidate window lengths. The proposed criterion is an optimization approach that evaluates the intrinsic quality of the trajectory matrix produced by each candidate window length. We use real-life data to illustrate the application of our proposed criterion. We demonstrate that our proposed criterion out-performed the existing methods used for the selection of window length in the literature.

Keywords: *Singular Spectrum Analysis (SSA); Singular Value Decomposition (SVD); Grouping; Reconstruction; Embedding dimension or window length; Trajectory Matrix*

Hierarchical Bayesian Logistic Regression with Adaptive Exclusive Lasso Prior for High-Dimensional Data

James Mayaki; Waheed Babatunde Yahya

Department of Statistics, University of Ilorin, Ilorin, Nigeria, Department of Statistics, Federal University of Technology, Minna, Nigeria

**Corresponding Author: James Mayaki ; j.mayaki@futminna.edu.ng*

High-dimensional classification is common in genomics, yet conventional methods such as LASSO, Group LASSO, and Sparse Group LASSO enforce uniform shrinkage with fixed penalties and offer no uncertainty quantification. Even frequentist Exclusive LASSO cannot adapt when signal strength varies across biologically meaningful groups. We propose Hierarchical Bayesian Exclusive LASSO Logistic Regression (BEXL-H), which learns group-specific shrinkage through a hierarchical prior on the Exclusive LASSO penalty. Posterior inference uses Hamiltonian Monte Carlo with the No-U-Turn Sampler. We establish posterior properties, consistency, adaptive contraction, and support recovery. Across three simulation regimes, BEXL-H outperforms other competitors where they collapse. At the 20-signal stress point under high-moderate correlation, LASSO's MCC falls to nearly zero and Group LASSO's AUC drops to 0.42, while BEXL-H maintains AUC 0.78 and well-calibrated probabilities. Under the most demanding regime at 40 signals, BEXL-H achieves the best MCC (0.58), AUC (0.84), and calibration (Brier 0.16). On three genomic datasets (Colon, Breast Cancer, Lymphoma), BEXL-H corroborates these gains. Where Group LASSO collapses to AUC 0.54 on Breast Cancer, BEXL-H achieves the highest F1 (0.67) and MCC (0.52). On Lymphoma, it records the lowest FDR (0.04), highest precision (0.96), and best-calibrated probabilities (Brier 0.07). Adaptive hierarchical regularization thus offers a principled framework for structured variable selection, combining within-group competition with Bayesian uncertainty quantification.

Keywords: *Bayesian logistic regression; Exclusive LASSO; hierarchical shrinkage; structured sparsity; high-dimensional classification.*

Robust Hierarchical Systematic Review-Derived Priors for Bayesian Causal Inference

Dongnaan G. A.; Adehi M.U.; Oguntade E.S

Department of Statistics; Faculty of Physical Sciences; University of Abuja; Nigeria

**Corresponding Author: Apameh Godswill Dongnaan ; apameh.dongnaan@uniabuja.edu.ng*

Bayesian causal inference allows us to use our prior knowledge along with current data to make inferences about the causal effect of a certain treatment on an outcome variable. There are various methods for constructing priors used within Bayesian causal inference, but most of them employ weakly informative priors or traditional meta-analytic predictive (MAP) priors, which either fail to incorporate external information adequately or presume full transportability of historical information. Thus, in case the current data is in conflict with historical data, traditional MAP priors tend to have too much weight in such situations. In this paper we introduce a framework for transforming evidence extracted from systematic reviews and random-effects meta-analysis into robust adaptive hierarchical priors for Bayesian causal inference called Robust Hierarchical Systematic Review-Derived Prior (RHSRDP). Our method allows to calculate three new measures - Prior Discordance Index (PDI), Evidence Compatibility Score (ECS), and Adaptive Evidence Weighting Factor (AEWF), which assess the compatibility of historical and current evidence and control the weight of prior information. Posterior inference is done using Bayesian logistic regression with adjustment based on propensity scores. Performance of the methods has been compared with frequentist inverse probability weighting, weakly informative Bayesian modeling with prior specification, Bayesian modeling with conventional MAP priors, and the original Hierarchical Systematic Review-Derived Prior (HSRDP). Conventional MAP priors resulted in the lowest RMSE (0.017), yet had excessive shrinkage problems, while weakly informative Bayesian and frequentist methods incurred higher RMSE (0.208). The proposed RHSRDP has shown good compromise between robustness and efficiency, providing RMSE 0.165, 97% coverage probability, stable posterior inference, and high convergence (Rhat 1.00). The proposed approach offers a reproducible framework for incorporation of systematic review information into Bayesian causal inference for comparative effectiveness and real world evidence research.

Keywords: *Bayesian causal inference; Systematic review; Informative priors; Propensity score weighting; Observational studies.*

Comparative Performance of Ensemble and Traditional Machine Learning Models for Breast Cancer Diagnosis

Memunat Manzuma; Ahmed Oloduwo Ameen; Mohammed Kabir Garba

Department of Computer Science, University of Ilorin, Nigeria; Department of Computer Science, University of Ilorin, Nigeria; Department of Statistics, University of Ilorin, Nigeria

**Corresponding Author: Mohammed Kabir Garba ; garba.mk@unilorin.edu.ng*

Breast cancer is one of the most prevalent and life-threatening diseases affecting women worldwide and remains a major public health concern. It is characterized by the uncontrolled growth of abnormal breast cells that can invade surrounding tissues and metastasize to distant organs if not detected and treated early. Despite the development of numerous machine learning models for breast cancer classification, no single algorithm has been universally accepted as the most effective because performance varies across datasets and evaluation metrics. Moreover, many studies focus on individual algorithms without providing comprehensive comparative analyses, making model selection for clinical application challenging. This study compares the performance of four machine learning algorithms: Nave Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Logistic Regression (LR) for breast cancer classification. In addition, a stacked ensemble model was developed by combining these algorithms, and its performance was evaluated against that of the individual classifiers using accuracy, precision, recall, and F1-score. The results show that the stacked ensemble achieved 99% accuracy, precision, recall, and F1-score, outperforming Logistic Regression (98%), K-Nearest Neighbors (95%), and Nave Bayes (93%), while matching the performance of Support Vector Machine (99%). Although stacking did not improve upon SVM's already optimal performance, it provided the robustness of ensemble learning by leveraging the complementary strengths of multiple classifiers. Logistic Regression also demonstrated excellent performance as a standalone classifier, whereas KNN and Nave Bayes produced comparatively lower classification accuracy. The study recommends the stacked ensemble model as the preferred computer-aided diagnostic approach for breast cancer classification because of its consistently superior performance across all evaluation metrics. Nevertheless, SVM and Logistic Regression remain strong standalone alternatives and should be further validated using larger and more diverse real-world datasets.

Keywords: *Breast Cancer Diagnosis; Machine Learning; Stacked Ensemble Learning; Support Vector Machine (SVM); Classification; Computer-Aided Diagnosis (CAD)*

Predicting Inter-Communal Conflict and Banditry Hotspots in Nigeria: a Machine Learning Approach Using Acled Conflict Data

I.C Ezeilo ; C. K. Onyekwere

Department of Statistics, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria ; Department of Statistics, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria

**Corresponding Author: Ifeyinwa Christiana Ezeilo ; ifenyinwaezeilo@gmail.com*

Nigeria continues to experience escalating inter-communal violence and armed banditry, particularly across the North-West, North-Central and North-East regions, resulting in more than 11,000 conflict-related deaths in 2025 alone. This study presents a machine learning framework for predicting high-risk Local Government Areas (LGAs) prone to inter-communal conflict and banditry using data from the Armed Conflict Location and Event Data Project (ACLED). Three supervised classification algorithms; Logistic Regression, Random Forest and Extreme Gradient Boosting (XGBoost), were developed to forecast conflict hotspots at the LGA level. Predictor variables included lagged fatality counts, event frequencies, rolling average fatalities and seasonal temporal features. The dataset comprised 53,703 LGA-month observations, exhibiting severe class imbalance, with only 532 high-risk observations compared to 53,171 low-risk observations. Among the models evaluated, XGBoost demonstrated the strongest predictive performance, achieving a Receiver Operating Characteristic Area under the Curve (ROC-AUC) of 0.7898. Rolling mean fatalities emerged as the most influential predictor, accounting for 69.3% of the model's feature importance. The model identified 22 high-risk LGA-month predictions, with Borgu Local Government Area in Niger State recording the highest predicted risk probability (0.878). Although all models achieved overall classification accuracies exceeding 98%, their ability to correctly identify the minority high-risk class remained limited, underscoring the persistent challenge of predicting rare conflict events. The findings highlight the implications of class imbalance for operational conflict early warning systems and suggest threshold optimization and synthetic oversampling techniques as promising strategies for improving predictive performance. This study contributes to the growing application of artificial intelligence in security intelligence and provides a data-driven framework to support evidence-based security planning and public sector decision-making in Nigeria.

Keywords: *Machine Learning; Conflict Prediction; ACLED; Early Warning Systems; XGBoost; Security Intelligence.*

A Bayesian Hierarchical Meta-Regression Framework Integrating Distributed Lag, Nonlinear Exposureresponse, and Climatepolicy Interaction Effects in Malaria Evidence Synthesis.

T. F. Aderemi; M. U. Adehi

1National Bureau of Statistics;2University of Abuja, Faculty of Physical Sciences, Department of Statistics, Abuja, Nigeria.

*Corresponding Author: Taiwo Aderemi ; taiwoaderemi41@gmail.com

Climate variability has become an important driver of malaria transmission in Africa, particularly among children under five years of age. Numerous time-series studies have investigated the associations between climatic variables and malaria incidence however, classical meta-analytic methods generally assume linear and instantaneous exposureresponse relationships and seldom account simultaneously for delayed climatic effects, nonlinear associations, intervention effect modification, and multiple sources of heterogeneity. Therefore, pooled estimates may inadequately characterize the complex, delayed, nonlinear, and interaction effects linking climatic factors, malaria control interventions, and malaria incidence. This study proposes a Bayesian Hierarchical Distributed lag -Interaction Meta-Regression model integrating distributed lag, nonlinear exposureresponse, and interaction effects for synthesizing climat malaria evidence. A systematic review and meta-analysis of time-series and ecological studies published between 2014 and 2024 will be conducted in accordance with the PRISMA 2020 guideline. Random-effects meta-analysis will first estimate pooled effect sizes and quantify between-study heterogeneity using the between-study variance (τ^2). The proposed framework will then incorporate distributed lag functions to model delayed climatic effects, penalized smoothing functions to capture nonlinear exposureresponse relationships, and interaction terms to evaluate how malaria control interventions modify climatic effects. Bayesian hierarchical modelling will simultaneously account for within-study uncertainty, between-study heterogeneity, and spatial and temporal variability. Posterior inference will be performed using the Integrated Nested Laplace Approximation (INLA). The proposed framework extends classical meta-regression by unifying distributed lag modelling, nonlinear exposureresponse analysis, climate-policy interaction and Bayesian hierarchical inference within a single evidence-synthesis framework. It is expected to provide more reliable estimation of climate-sensitive malaria risk, improve uncertainty quantification, and offer a general statistical methodology applicable to complex environmental and public health evidence synthesis.

Keywords: Bayesian hierarchical meta-regression; Distributed lag; Nonlinear exposureresponse; Climate change; Malaria.

Survival Analysis of Volatility of Life Expectancy: a Truncated Two-Parameter Pranav Distribution Approach

C.K. Nwankwo; M.U. Adehi; G.A. Dongnaan

University of Abuja

*Corresponding Author: Chinedu Kingsley Nwankwo ; kcnwankwo01@gmail.com

This study used the Truncated Two-Parameter Pranav distribution to model the national life expectancy data for Nigeria. The research aims to bridge the gap in the application of the truncated two-parameter Pranav distribution for modelling life expectancy data, which was not previously used for such data. The research used the life expectancy data for the years 1950- 2025. The parameters were estimated using the maximum likelihood method, and the Truncated Pranav distribution was compared with the Exponential, Lindley, Akash, Ishita, and Pranav distributions. The proposed distribution was found to perform better compared to the other distributions using the Akaike Information Criterion and the Bayesian Information Criterion. The empirical results showed that the Truncated Two-Parameter Pranav Distribution achieved the lowest negative log-likelihood ($LL = -240.20$), Akaike Information Criterion ($AIC = 484.40$), and Bayesian Information Criterion ($BIC = 489.06$), outperforming the competing distributions. The proposed distribution, having filled the gap in the application of the truncated two-parameter Pranav distribution for modelling the life expectancy data, was found to provide useful insights into the longevity of the people of Nigeria.

Keywords: Truncated Two-Parameter Pranav Distribution; Life Expectancy Modeling; Lifetime Distribution; Maximum Likelihood Estimation; Survival Analysis

Modeling the Sporadic Behavior of Rainfall Time Series Using Ets State Space and Sarima Models in the Limpopo Province, South Africa

T.A. Lasisi; O.O. Akanni; S.V. Molautsi; M. Lesaoana; D.K. Shangodoyin

Department of Statistics, Ladoke Akintola University of Technology, Nigeria; Department of Statistics, University of Ibadan, Nigeria; Department of Statistics and Operation Research, University of Limpopo, South Africa; Department of Statistics and Operation Research, University of Limpopo, South Africa; Department of Statistics, Ladoke Akintola University of Technology, Nigeria

**Corresponding Author: OLUKUNMI OLATUNJI AKANNI ; tjstatistician@yahoo.com*

The effects of ozone depletion on climate change have, in recent years, become a reality, impacting changes in rainfall patterns and the intensity of extreme floods or extreme droughts. The majority of people across the entire African continent live in semi-arid and drought-prone areas. Extreme droughts are prevalent in Somalia and eastern Africa, while life-threatening floods are common in Mozambique and some parts of other Southern African Development Community countries. Research has signaled that climate change in South Africa might lead to increased temperatures and reduced amounts of rainfall, thereby altering their timing and putting more pressure on the country's scarce water resources, with implications for agriculture, employment, and food security. The average annual rainfall for South Africa is about 460mm, which falls far below the global average rainfall of 860mm. The Limpopo Province, one of the nine provinces in South Africa and of interest to this study, is predominantly agrarian, basically relying on availability of water, with rainfall being the major source for water supply. It is, therefore, pertinent that the rainfall pattern in the province be monitored effectively to ascertain the rainy period for farming activities and other uses. This study employed the SARIMA and ETS State Space models to capture the sporadic behavior of rainfall data. These two models have been widely applied to climatic data by many scholars and adjudged to perform creditably well. In an attempt to find a suitable prediction model for rainfall patterns in Limpopo Province, data from Macuville and Marnits Stations, ranging from January 1999 to December 2025, were analyzed. SARIMA(1,0,1)(1,1,1)₁₂ and ETS(A,N,A) models based on exponential smoothing were built, and the results showed that the two models were adequate.

Keywords: *Time Series, Rainfall, SARIMA model, ETS State Space model, Forecasting*

Bayesian Modelling of Family Planning Discontinuation Under Socioeconomic and Environmental Uncertainty in Rural Communities in South East Nigeria

E. M. Ikegwu; G. B. Karokatose; M. A. Idris

Department of Statistics, Yaba College of Technology, Yaba Lagos

**Corresponding Author: Emmanuel Mmaduabuchi Ikegwu ; macioe2002@yahoo.com*

Despite increasing contraceptive awareness, discontinuation rates continue to contribute to unintended pregnancies, maternal health risks, and population pressure. This study investigated the determinants of family planning discontinuation among women of reproductive age in South-East Nigeria using a hierarchical Bayesian Stochastic modelling framework. The study utilised the IRPGS dataset collected from the five states of South-East Nigeria, comprising household-level reproductive-health, socioeconomic, migration, healthcare accessibility, and environmental variables. A hierarchical Bayesian logistic regression model was employed to estimate the probability of contraceptive discontinuation while accounting for latent state-level heterogeneity. Fixed effects included age, healthcare accessibility, food price instability, flood exposure, migration experience, drought exposure, educational attainment, and family planning knowledge, while state-level random effects captured contextual variability. Posterior estimation, convergence diagnostics, sensitivity analysis and predictive information criteria, were used for model validation. The findings revealed that environmental shocks significantly influenced contraceptive discontinuation behaviour. Drought exposure substantially increased discontinuation risk, while flood exposure also exhibited a positive association with discontinuation. Migration experience emerged as a major protective factor, significantly reducing discontinuation probability. Increased travel time to healthcare facilities slightly elevated discontinuation risk. State-level random effects indicated substantial geographic heterogeneity, with Ebonyi, Enugu and Imo States exhibiting higher latent discontinuation vulnerability compared with Abia and Anambra States. Convergence diagnostics showed satisfactory posterior convergence with Gelman-Rubin statistics below 1.1, while sensitivity analysis confirmed robustness of posterior estimates. The study concludes that family planning discontinuation in South-East Nigeria is shaped by complex interactions among environmental uncertainty, healthcare accessibility, migration dynamics, and latent geographic factors.

Keywords: *Family planning discontinuation; Hierarchical Bayesian modelling; Reproductive health; Environmental uncertainty; Stochastic demographic modelling.*

Comparative Analysis of Arima and Machine Learning Models for Forecasting Crude Oil Prices in Nigeria

Afolayan Tawakalt, S. T. Raji, S. O. Jabaru

Ilorin College of Health Technology; Federal Polytechnic Offa; Al-Hikmah University.

**Corresponding Author: Tawakalt Afolayan ; afolayanjummy@gmail.com*

Comparative Analysis of ARIMA and Machine Learning Models for Forecasting Crude Oil Prices in Nigeria Afolayan Tawakalt, S. T. Raji, S. O. Jabaru Ilorin College of Health Technology, Ilorin, Kwara State, Nigeria Department of Statistics, Federal Polytechnic Offa, Offa, Kwara State, Nigeria Department of Physical Science, Al-Hikmah University, Ilorin, Kwara State, Nigeria Corresponding Author: Afolayan Tawakalt E-mail: afolayanjummy@gmail.com

ABSTRACT Forecasting crude oil prices is essential for effective economic planning, fiscal management, and evidence-based policy formulation in oil-dependent economies such as Nigeria. Persistent fluctuations in global crude oil prices pose significant challenges to government revenue generation, exchange rate stability, inflation management, and national development planning. This study proposes a comparative analysis of the AutoRegressive Integrated Moving Average (ARIMA) model and selected machine learning models, namely Random Forest (RF) and Support Vector Regression (SVR), for forecasting crude oil prices in Nigeria. Monthly Brent crude oil price data covering the period from January 2000 to December 2025 will be obtained from the U.S. Energy Information Administration (EIA). Data preprocessing and descriptive statistical analysis will be conducted, while stationarity will be examined using the Augmented DickeyFuller (ADF) unit root test. The ARIMA model will be developed using the BoxJenkins methodology, whereas the Random Forest and Support Vector Regression models will be implemented using machine learning techniques. Forecasting performance will be evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The closer these three measures to zero, the better their prediction power. The study is expected to identify the forecasting model that provides the highest predictive accuracy and demonstrate the value of integrating statistical intelligence with advanced forecasting techniques for evidence-based economic planning in Nigeria. The expected findings will contribute to improved fiscal planning, policy formulation, and sustainable economic development in a data-driven economy.

Keywords: *Crude oil price forecasting; ARIMA; Machine learning; Random Forest; Nigeria.*

A New Robust Estimator for Modeling Multiple Linear Regression Model in Presence of Outliers

W.B. Yahya; M. Animashahun

University of Ilorin, Ilorin; University of Ilorin, Ilorin

**Corresponding Author: Musaphau - Animashahun ; musaphauanimashahun@gmail.com*

In this work, the comparison of methods of modelling multiple linear regression model in presence outliers was done. Hubber and Thein Sen methods were fused together so as to get the proposed method. The proposed method performed relatively well while subject to several conditions of different outlier percentages and sample sizes. The Root Mean Square of Error (RMSE) was used to determine relative efficient of the proposed method. Equally, the average ranks of the RMSE equally showed that the performances of the proposed was the best.

Keywords: *outliers, modelling, Hubber, Thein-Sen, proposed Estimator, RMSE*

Analysis of Covariance (ancova) in Rcbd: a Non-Gaussian Residual

W.B. Yahya; O.J. Sholademi; L.O. Mashood; K. Jimoh

Department of Statistics, University of Ilorin, Kwara State, Nigeria; Department of Mathematics and Statistics, Air Force Institute of Technology, Kaduna State, Nigeria; Department of Mathematics and Statistics, Air Force Institute of Technology, Kaduna State, Nigeria; Department of Mathematics and Statistics, Air Force Institute of Technology, Kaduna State, Nigeria

**Corresponding Author: KAMORU JIMOH ; jimkaminsha@alhikmah.edu.ng*

This study presents an application of Analysis of Covariance (ANCOVA) in a Randomized Complete Block Design (RCBD) with a Non-Gaussian residual. The study highlights the common occurrence of concomitant variables confounding the relationship between independent factors and dependent variables in experimental research. The study aimed to perform analysis of covariance involving two factors with a concomitant variable, presenting the impact of independent variables, the concomitant variable, and understanding their individual impacts. The methodology involves simulating data for the independent factors, dependent variable, and concomitant variable from normal random variables with an exponential random error. An ANCOVA model is specified, incorporating the relevant variables, and assumptions were checked. The interpretation of the results considers the significance of the independent variables, interaction effects, and the concomitant variable. The conclusion emphasizes the importance of covariance analysis with a non-Gaussian residual in obtaining a more accurate understanding of the relationship between independent factors and the dependent variable, contributing to evidence-based decision making in various fields of study.

Keywords: ANCOVA; ANOVA; RCBD; Concomitant Variable; Non-Gaussian Residual; Centered Model

Performance Comparison of Count Regression Models Under Simultaneous Overdispersion, Multicollinearity and Outlier Contamination

G. M. Oyinloye and A. Bolakale

University of Ilorin, Ilorin

**Corresponding Author: Glory Moyinoluwa Oyinloye ; shekinahglory1025@gmail.com*

Count data encountered in public health, transportation, criminology, and related fields often exhibit overdispersion, multicollinearity, and outliers, violating the assumptions of the classical Poisson regression model. While these challenges have been widely studied individually, their combined effect on the performance of count regression models has received limited attention. This study compares the performance of Poisson Regression, Generalized Poisson (GP), Negative Binomial (NB), ConwayMaxwell Poisson (CMP), ConwayMaxwell Poisson Ridge (CMP Ridge), and Zero-Inflated Negative Binomial (ZINB) models under varying levels of overdispersion, multicollinearity, and outlier contamination. A Monte Carlo simulation based on a Poisson log-normal data-generating process was conducted using a full factorial design comprising 135 simulation scenarios with 1,000 replications per scenario. Model performance was assessed using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), log-likelihood, and convergence rate. The results showed that the GP model consistently achieved the best overall performance across the simulation scenarios, whereas the Poisson model performed poorly under increasing overdispersion. The NB model ranked second overall, while the CMP Ridge implementation failed to converge. These findings provide empirical guidance for selecting appropriate count regression models when multiple data complexities occur simultaneously.

Keywords: Count data regression; Generalized Poisson regression; Overdispersion; Multicollinearity; Outliers; Monte Carlo simulation.

Firth Penalized Logistic Regression for Predicting Abnormal Neuroimaging Findings Among Headache Patients in North-Eastern Nigeria

S.M. Oguche; F.Y. Sayuti; M.B. Imam; A.Y.F. Jibrin; H.U. Okemmiri

Federal University of Technology, Minna; Federal University of Technology, Minna; Niger state Polytechnic Zungeru; Federal Polytechnic, Bida, Niger State, El-Amin University, Eastern Bypass Minna

**Corresponding Author: SAMUEL MICHAEL OGUCHE ; samueloguche690@gmail.com*

Neuroimaging is broadly employed to assess patients with headache, but this is not cost-effective, as the yield from the indiscriminate use is significantly lower. The use of clinical red-flag frameworks to direct imaging is standard practice, but few of these frameworks have been rigorously validated using predictive modelling, especially in North Eastern Nigeria and sub-Saharan Africa. The prevalence of clinically significant (abnormal) neuroimaging findings in headache patients, independent predictors of these findings and the use of Firth penalized logistic regression to overcome small-sample bias and quasi-separation in the data were determined. 348 patients who had neuroimaging for headache in two tertiary centres (University of Maiduguri Teaching Hospital and State Specialist Hospital) in North-Eastern Nigeria were analysed. There were 188 (54.02%) abnormal neuroimaging results, which were comparable between both centres. Imaging type ($\chi^2 = 6.18$, $p = 0.013$), headache subtype ($\chi^2 = 115.07$, $p < 0.001$), and acute presentation ($\chi^2 = 10.92$, $p = 0.001$) were significantly associated with imaging diagnosis. The headache-subtype variable was very strongly associated with outcome, suggesting almost perfect separation, which was the rationale for using Firth penalized regression. In the Firth model, headache subtype (AOR = 121.33; 95% CI: 23.64622.80; $p < 0.001$) and imaging type (AOR = 2.09; 95% CI: 1.213.62; $p = 0.008$) were independent predictors of abnormal findings, while age and sex were not. The model had an AUC of 0.8174, moderate sensitivity (53.72%), 99.38% specificity and 99.02% PPV. Red-flag headache subtype is the most important independent predictor of abnormal neuroimaging in this population. In a quasi-separation setting, regression using Firth's penalized estimator gave more stable estimation than standard logistic regression and this finding was robust to the quasi-separation setting and could be applied to other clinical settings with similar structures.

Keywords: *Firth; Penalized Logistic; Neuroimaging; Headache, North-Eastern*

Bayesian Tournament-Level Forecasting Under Structural Expansion: Validation, Calibration, and Monte Carlo Stability Using Fifa World Cup Data

Taiwo M. Adegokea ; Oladapo M. Oladoja ; Kazeem B. Kasali ; Adefunke B. Familua ; Waheed B. Yahya

*a*Department of Mathematics and Statistics, Abiola Ajimobi Technical University, Ibadan, Oyo state, Nigeria; Department of Statistics, University of Iloin, Kwara State, Nigeria; Centre for Statistical Genomics and Computational Biology, University of Iloin, Kwara State, Nigeria

*Corresponding Author: Taiwo Mobolaji Adegoke ; adegoketaiwom@gmail.com

Tournament expansion creates a challenging statistical transport problem because predictive models developed under one competition structure must be applied to a target format containing a different number of teams and advancement stages. This study develops a Bayesian probabilistic framework for forecasting multistage tournament outcomes under structural expansion, using the FIFA World Cup as an empirical application. The model was trained on seven complete 32-team tournaments held between 1998 and 2022, comprising 224 team tournament observations and 434 eligible stage-transition records. Pre-tournament World Football Elo ratings were used as standardized measures of team strength. A hierarchical logistic model estimated stage-specific Elo effects and between-tournament heterogeneity. Posterior parameters from the five historical advancement transitions were transported to the six-stage 48-team format using normalized tournament-depth interpolation. Format-constrained simulation with weighted sampling without replacement ensured that every simulated tournament followed the exact progression (48->32->16->8->4->2->1). The primary analysis combined 8,000 posterior draws with 50 nested tournament simulations per draw, producing 400,000 complete tournament realizations. Locked out-of-sample validation yielded a mean Brier score of 0.0787, mean logarithmic score of 0.2410, mean area under the ROC curve of 0.9162, mean average precision of 0.8942, and mean expected calibration error of 0.0692. The highest-ranked team was the observed champion, while the two highest-ranked teams were the observed finalists. All semifinalists, finalists, and the champion were contained in the corresponding predicted top-(k) sets. Computational robustness was assessed across 20 combinations of posterior draws and tournament simulations. The maximum probability deviation was below one percentage point, champion-ranking correlations exceeded 0.95, and the top-10 set remained unchanged across all scenarios. The proposed framework provides a reproducible computational approach for transporting Bayesian tournament forecasts across structural changes while preserving uncertainty, logical coherence, exact competition constraints, and numerical stability.

Keywords: Bayesian forecasting; tournament simulation; probabilistic prediction; calibration; proper scoring rules; Monte Carlo stability; Elo ratings; computational statistics

When Can Informative Priors Be Trusted in Bayesian Gaussian Change-Point Analysis? Operating Characteristics, Priordata Conflict, and Posterior Robustness

Taiwo M. Adegoke ; Waheed B. Yahya

Department of Mathematics and Statistics, Abiola Ajimobi Technical University, Ibadan, Oyo state, Nigeria ; Department of Statistics, University of Iloin, Kwara State, Nigeria ; Centre for Statistical Genomics and Computational Biology, University of Iloin, Kwara State, Nigeria

**Corresponding Author: Taiwo Mobolaji Adegoke ; adegoketaiwom@gmail.com*

Bayesian change-point models are widely used for detecting structural changes in ordered data across numerous scientific disciplines. Although informative priors are frequently employed to incorporate substantive prior knowledge, relatively little is known about the finite-sample conditions under which they improve inference or become sensitive to prior misspecification. In particular, practical guidance regarding when informative priors should be preferred, avoided, or regarded as practically negligible remains limited. This study investigates the operating characteristics of informative conjugate priors for segment means and variances conditional on an unknown single change point in Gaussian models. We develop a comprehensive Bayesian framework that examines the joint effects of sample size, signal strength, prior strength, and prior-data conflict on change-point localization and posterior uncertainty. The methodology employs Normal-Inverse-Gamma informative priors together with posterior summaries based on bias, root mean squared error, exact and near-exact change-point recovery, credible interval coverage, posterior entropy, and relative performance measures. Informative priors are classified into beneficial, practically negligible, and harmful operating regimes using relative root mean squared error thresholds. Robustness is further investigated under prior misspecification, non-central change points, Student-t errors, unequal variances, serial dependence, and no-change scenarios. Results from the validated simulation framework demonstrate that informative priors exhibit identifiable operating regimes rather than uniformly improving change-point estimation. Moderate informative priors are remarkably robust to moderate prior misspecification, whereas severe prior-data conflict can induce substantial deterioration in localization performance for highly concentrated priors. Increasing sample size and signal strength substantially reduce prior sensitivity, with informative priors frequently becoming practically negligible relative to weakly informative alternatives in highly informative settings. Under no-change scenarios, the proposed methodology exhibits low false-positive rates and increasing resistance to spurious change-point detection as information accumulates. A sensitivity analysis of the classical Nile River annual flow data further suggests that informative priors primarily influence posterior concentration rather than substantive posterior location when the observed data strongly support a change point. These findings suggest that informative conjugate priors should be viewed not as mechanisms for imposing change points, but as principled devices for incorporating prior scientific knowledge while preserving likelihood-driven inference. The proposed operating-characteristic framework provides practical guidance regarding when informative priors may be beneficial, practically negligible, or harmful in Bayesian Gaussian change-point analysis.

Keywords: *Bayesian change-point analysis; informative priors; prior-data conflict; posterior robustness; operating characteristics; uncertainty quantification; Nile River data.*

Integrated Stochastic Modeling of Parity, Spatial Fertility, Heterogeneity, and Intergenerational Kinship Dynamics in Rural South-East Nigeria

E. M. Ikegwu; G. B. Karokatose; M. A. Idris

Department of Statistics, Yaba College of Technology, Yaba Lagos

**Corresponding Author: Emmanuel Mmaduabuchi Ikegwu ; macioe2002@yahoo.com*

Understanding fertility dynamics in subnational populations requires integration of count regression modelling, hierarchical small-area estimation (SAE), spatial dependence structures, and stochastic branching processes governing intergenerational expansion. Using microdata ($n = 1,092$) from the IRPGS survey across five South-East Nigeria states (Abia, Anambra, Ebonyi, Enugu, Imo), we developed an integrated framework combining Negative Binomial (NB) model for parity outcomes; Bayesian hierarchical small-area shrinkage; conditional autoregressive (CAR) spatial smoothing; migration-modified Galton-Watson branching processes; longitudinal fertility projection; and replacement threshold stability analysis via Lyapunov exponents. Results show mean children ever born (CEB) = 3.58 (SD = 1.86), with moderate inter-state variation but negligible overdispersion (0). Bayesian shrinkage modestly stabilises district estimates while spatial smoothing reveals moderate positive clustering. Branching simulations indicate supercritical kinship expansion (mean 225 descendants in four generations), reduced substantially under migration scenarios. Longitudinal projections suggest potential transition below replacement fertility (2.1) by mid century under modest decline assumptions, though demographic momentum persists. Findings provide a unified stochastic-demographic framework for rural fertility control policy.

Keywords: *Bayesian Hierarchical modelling; Branching process; Fertility heterogeneity; Kinship expansion; Negative Binomial regression; Spatial Conditional Autoregressive model*

Development and Evaluation of a Robust Lad-Based New Two-Parameter Estimator for Linear Regression Models

T. J. Adejumo; A. T. Owolabi; O. J. Oladapo, J. I. Idowu

Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.

**Corresponding Author: Abiola Timothy Owolabi ; atowolabi@lautech.edu.ng*

This study proposed a Robust Least Absolute Deviation New Two-Parameter (LAD-NTP) estimator for linear regression models in the presence of multicollinearity and extreme observations. The estimator was developed by integrating the robustness of Least Absolute Deviation estimation with the shrinkage property of the New Two-Parameter estimator. Its theoretical properties were established through Mean Squared Error (MSE) derivation and mathematical comparisons with existing estimators. A Monte Carlo simulation involving sample sizes of 10, 20, 30, 40, 50, 100, 250 and 500, correlation levels ranging from 0, 0.8, 0.9, 0.95 and 0.99, error variances of 1, 5, and 10, and different proportions and magnitudes of outliers was conducted to evaluate performance. The results showed that MSE increased with increasing multicollinearity, error variance, and outlier contamination but decreased as sample size increased. The proposed LAD-NTP estimator consistently outperformed all competing estimators considered in the study, attaining the highest cumulative top-ranking frequency of 2,572 across all simulation scenarios. Application to the Portland cement dataset, characterized by severe multicollinearity (VIFs between 19.87 and 160.78), further confirmed its superiority by producing the lowest MSE (21.51186) among all estimators considered. The study concludes that the LAD-NTP estimator provides a highly efficient and robust alternative for regression analysis involving simultaneously correlated predictors and extreme observations.

Keywords: *LAD-NTP; Multicollinearity; Outliers; Regression; Estimator.*

Some Improved Generalised Liu Estimators for Linear Regression Models in the Presence of Multicollinearity

Okegbade A.I.; 2Job O.; 1Adejumo T.J.; 1Adesina O.A.; 1Okunlola O.O and 1Egbewumi A.O

Ladoke Akintola University of Technology, Ogbomoso; Department of Statistics University of Ilorin

*Corresponding Author: AYOBAMI IBUKUN OKEGBADE ; aiokegbade@lautech.edu.ng

The Generalised Liu regression estimator is one of the notable biased estimation techniques in linear regression analysis. It relies on a shrinkage parameter d , which is a function of the ridge parameter k . Several forms of k have been proposed in the literature, including those by Kibria (2003) and Muniz et al. (2009), while alternative formulations for d have been suggested by authors such as Shakallioglu and Kasciranlar (2006), demonstrating reduced mean square error (MSE) compared to existing approaches. This study addresses the need to further improve the generalised Liu estimator for linear regression models affected by multicollinearity. A Monte Carlo simulation with 1000 replications was conducted, considering three exogenous variables ($p=3$) and six levels of sample size ($n=10,20,30,50,75,100$). The error term was generated from a normal distribution with mean zero and variance 2, Various error variance considered are 1, 5, 10 and heteroscedasticity was set at 1. Correlation levels were varied at $\rho=0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.99$. For each scenario defined by sample size, correlation level, and error variance, the performance of both existing and improved generalised Liu estimators was evaluated using the MSE criterion and subsequently ranked. The results showed that enhancing the shrinkage parameter d using the inverse of the maximum variance inflation factor consistently improved the Liu estimator's performance across all conditions, except when 0.95 and $n=75$, where no improvement was preferable. The study therefore recommends adopting this improved formulation of the shrinkage parameter d in generalised Liu estimation for better estimation accuracy in most scenarios. Keywords: Biased, Estimation, Liu, Ridge, and Multicollinearity.

Keywords: Keywords: Biased; Estimation; Liu; Ridge; and Multicollinearity.

An Empirical Analysis of Machine Learning Algorithms for Predicting Poverty Vulnerability and Informing Social Protection Policy in Nigeria

Musa, Y.Y.; Suleiman, M.K.; Aliyu U; and Musa, I.

Waziri Umaru Federal Polytechnic, Birnin kebbi

*Corresponding Author: Yakubu Yeldu Musa ; yeldouyakubu@gmail.com

AbstractPoverty vulnerability remains one of Nigeria's most pressing socioeconomic challenges, with millions of households exposed to welfare shocks that existing social protection systems inadequately address. Traditional statistical approaches to poverty targeting relying primarily on income thresholds and limited predictor variables frequently produce significant exclusion errors that undermine policy effectiveness and fiscal efficiency. This study empirically evaluates the predictive performance of three machine learning (ML) algorithms: Random Forest (RF), Extreme Gradient Boosting (XGBoost) and Logistic Regression (LR) in identifying poverty-vulnerable households in Nigeria. Using the National Bureau of Statistics (NBS) Nigeria Living Standards Survey (NLSS) 2018/2019 dataset, the study constructs a multidimensional vulnerability index incorporating socioeconomic, demographic, geographic and asset-based predictors. Model performance is assessed using accuracy, precision, recall, F1-score and Area under the Receiver Operating Characteristic Curve (AUC-ROC) metrics. SHAP (Shapley Additive Explanations) values are employed to interpret the key drivers of vulnerability at both household and regional levels. Empirical results indicate that ensemble methods, particularly XGBoost significantly outperform conventional logistic regression in predictive accuracy (AUC-ROC: 0.91 vs. 0.74) and generalization. Key vulnerability drivers identified include household size, educational attainment of the household head, access to clean water, geographic location (rural/urban) and asset ownership. The findings offer actionable insights for improving the targeting efficiency of Nigeria's National Social Investment Programme (NSIP) and demonstrate the value of integrating machine learning tools into national poverty monitoring frameworks. Keywords: Poverty Vulnerability, Machine Learning, XGBoost, Random Forest, Social Protection Policy, Nigeria, SHAP, NLSS

Keywords: Poverty Vulnerability, Machine Learning, XGBoost, Random Forest, Social Protection Policy, Nigeria, SHAP, NLSS

An Integrated Security Risk Index for Crime Risk Mapping: a Spatial Statistical Framework with Application to Kidnapping in Nigeria

K. O. Olorede; J. S. Olumoh; W. B. Yahya

Kwara State University, Kwara State University, 7 P.M.B. 1530, Malete, 241104, Kwara State, Nigeria; Department of Statistics, American University of Nigeria, 98, Lamido 9 Zubairu Way, Yola, 640211, Adamawa State, Nigeria; Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin, 11 240003, Kwara State, Nigeria

**Corresponding Author: Kabir Opeyemi Olorede ; kabir.olorede@kwasu.edu.ng*

Kidnapping is a persistent security challenge in Nigeria, yet existing assessments often prioritise observed incident counts without simultaneously accounting for structural conditions associated with underlying risk. This study develops and validates an Integrated Security Risk Index (ISRI) that combines a Structural Vulnerability Index based on multidimensional poverty, unemployment, population density and literacy-related vulnerability with an Operational Burden Index constructed from recorded kidnapping incidents. The analysis used 3,742 unique events recorded across Nigeria's 36 states and the Federal Capital Territory between April 2000 and November 2025. The index was evaluated using Global Moran's I, Local Indicators of Spatial Association, Getis-Ord G_i^* hotspot analysis, time-series forecasting, external validation against January-June 2026 ACLED records and continuous weighting sensitivity analysis. The ISRI showed statistically significant positive spatial autocorrelation, with persistent high-risk concentrations across north-western Nigeria and Zamfara, Katsina and Kaduna consistently occupying the highest-risk positions. External validation yielded 97.3% observed agreement and Cohen's $\kappa=0.916$, while the principal rankings remained stable across plausible weighting specifications. Integrating realised operational burden with structural vulnerability therefore provides a transparent and reproducible framework for hotspot identification, strategic resource allocation and evidence-informed spatial crime-risk assessment.

Keywords: *Kidnapping; Spatial crime analysis; Composite indicator; Crime risk mapping; Hotspot analysis; Nigeria*

A Feasible Generalized Ridge Estimator for Linear Regression Under Multicollinearity and Non-Spherical Errors

Y. H. Aliyu; M. K. Garba

University of Ilorin, Ilorin, Nigeria

**Corresponding Author: YUSUF Hassan Aliyu ; yusufaliko@gmail.com*

Reliable statistical inference from linear regression models depends on the assumptions of independent, homoskedastic errors and the absence of severe multicollinearity. In practice, however, explanatory variables are frequently highly correlated while the error process simultaneously exhibits heteroskedasticity and autocorrelation. Existing estimation techniques address these challenges separately: Ridge and Generalized Ridge Regression mitigate multicollinearity, whereas Generalized Least Squares (GLS) and Feasible Generalized Least Squares (FGLS) account for non-spherical error structures. Consequently, there remains no unified estimator that simultaneously regularizes ill-conditioned design matrices and accommodates unknown covariance structures within a single estimation framework. This paper proposes a Feasible Generalized Ridge (FGR) Estimator, a covariance-regularized estimation procedure that integrates generalized ridge regularization with feasible covariance adjustment for linear regression models. The proposed estimator is developed from a unified optimization framework and is shown theoretically to encompass the Ordinary Least Squares (OLS), Ridge Regression (RR), Generalized Ridge Regression (GRR), Generalized Least Squares (GLS), and Feasible Generalized Least Squares (FGLS) estimators as special cases under appropriate restrictions. Fundamental theoretical properties, including existence, uniqueness, bias, covariance matrix, and Mean Squared Error Matrix (MSEM), are derived to establish the mathematical validity of the estimator. The proposed methodology is further illustrated using a real-world dataset to demonstrate its practical applicability. The proposed estimator provides a unified and flexible framework for regression estimation under simultaneous violations of classical assumptions. Its ability to integrate coefficient regularization with covariance correction makes it a promising methodological contribution to modern statistical modelling and evidence-based decision-making in data-driven environments.

Keywords: *Covariance Regularization; Generalized Ridge Regression; Feasible Generalized Least Squares; Multicollinearity; Non-Spherical Errors.*

Agricultural Credit and Sustainable Food Systems in Nigeria: Implications for Agricultural Productivity, Food Security and Public Health

M. A. Umar; O. A. Oyegoke; O. O. Oladapo

Federal University of Health Sciences Azare, Bauchi - Nigeria; Federal Polytechnic Iree, Osun state

**Corresponding Author: UMAR ADAMU MUHAMMAD ; mumarad90@gmail.com*

Agricultural financing is widely recognized as a critical driver of agricultural productivity, rural development, and national food security. Despite substantial investments in agricultural credit schemes in Nigeria, concerns remain regarding the effectiveness of these interventions in stimulating agricultural growth and improving broader development outcomes. This study evaluates the relationship between agricultural credit, agricultural growth, and agriculture's contribution to the Gross Domestic Product (GDP) using annual secondary data for Nigeria covering the period 2010-2025. The study employs descriptive statistics, trend analysis, Pearson correlation analysis, multiple linear regression, binary logistic regression, and Firth penalized logistic regression to investigate the influence of agricultural credit and lending interest rates on agricultural performance. The binary response models classify agricultural growth into high- and low-growth categories, allowing comparison between conventional logistic regression and Firth's bias-reduced estimation under potential data separation. Beyond statistical evaluation, the study proposes a conceptual pathway linking agricultural credit to sustainable food systems, food security, improved nutrition, and public health outcomes. Results are expected to demonstrate that increased agricultural credit positively influences agricultural productivity, whereas high lending interest rates constrain farmers' access to financial resources. The study argues that agricultural financing policies should be viewed not only as economic instruments but also as strategic investments in food security, nutrition, and population health. The findings are expected to provide valuable evidence for evidence-based policy formulation and contribute to sustainable agricultural development in Nigeria.

Keywords: *Agricultural Credit; Sustainable Food Systems; Firth Logistic Regression; Food Security; Public Health*

A Nonlinear Scale Mixture Framework for Statistical Intelligence in Infectious Disease Surveillance

M. A. Umar; N. R. Zakariyau; W. B. Yahya

Federal University of Health Sciences Azare, Bauchi - Nigeria; Federal Polytechnic Oko, Anambra state; University of Ilorin, Ilorin, Nigeria

**Corresponding Author: UMAR ADAMU MUHAMMAD ; mumarad90@gmail.com*

Accurate disease surveillance is fundamental to health security, epidemic preparedness, and evidence-based public health decision-making. However, epidemiological surveillance data frequently exhibit heterogeneous variability arising from differences in disease transmission, environmental conditions, healthcare accessibility, reporting practices, and localized outbreaks. Classical Gaussian-based statistical models and conventional Scale Mixture of Normal (SMN) distributions often assume linear variance scaling, which may be inadequate for representing these complex sources of uncertainty. This paper introduces a Modified Scale Mixture of Normal (mSMN) framework in which the conventional linear scale component is replaced by a flexible nonlinear variance scaling function. Specifically, the proposed model contains a positive nonlinear scaling function governing latent variability. The proposed framework substantially extends the classical SMN family by accommodating nonlinear heterogeneity, improving robustness to heavy-tailed observations, and providing greater flexibility in modeling complex epidemiological processes. Fundamental theoretical properties, including conditions for properness, stochastic representation, moments, and special cases, are established. Maximum likelihood and Bayesian estimation procedures are discussed, together with simulation studies evaluating parameter recovery and robustness under varying levels of heterogeneity. An application to infectious disease surveillance data illustrates the practical advantages of the proposed framework over existing SMN-based models in capturing outbreak variability and supporting statistical intelligence for health security. The proposed methodology offers a flexible statistical framework for analyzing surveillance data characterized by complex uncertainty and contributes to the development of more reliable evidence-based public health monitoring systems.

Keywords: *Modified Scale Mixture of Normal; Disease Surveillance; Statistical Intelligence; Health Security; Nonlinear Variance Scaling.*

Predicting Naira Stress: Inflation and External Reserve Adequacy as Early-Warning Indicators in Nigeria

G. T. ZEBULUN; L. A. ADEYEYE

Ignatius Ajuru University of Education; Federal College of Animal Health and Production Technology Ibadan

**Corresponding Author: Gladys ThankGod Zebulun ; gladys.zebulun@iaue.edu.ng*

AbstractNigerias foreign-exchange market has experienced repeated episodes of abrupt naira depreciation, persistent inflation, and fluctuating reserve cover. This study evaluates whether inflation and external reserve adequacy provide useful early-warning information for monthly naira stress. The analysis uses a time-series dataset consisting of 204 monthly observations from January 2008 to December 2024 obtained from the Central Bank of Nigeria Statistical Bulletin. Naira stress is measured as 100 times the monthly logarithmic change in the end-period official naira to United States dollar exchange rate. Headline inflation and months of import cover serve as candidate predictors. The empirical procedure combines descriptive analysis, seasonal-trend decomposition, rolling volatility, unit-root tests, multiple-break detection, cross-correlation analysis, SARIMA and ARIMAX estimation, VAR-based causality analysis, impulse responses, forecast-error variance decomposition, objective stress thresholds, and rolling one-step out-of-sample forecasts. The results show weak seasonality but pronounced trend and regime effects. The biggest periods of stress were in June 2023, January 2024, and June 2016. Inflation has a positive concurrent association with stress, while reserve adequacy shows a weak unconditional association. The selected ARIMAX specification assigns the expected positive sign to lagged inflation and the negative sign to lagged reserve cover, but neither coefficient is statistically significant. VAR results provide no evidence of Granger causality from either predictor to monthly stress. Over the final 39-month test period, ARIMAX reduces stress RMSE from 11.53 to 11.10 percentage points and exchange-level RMSE from 108.53 to 107.95, yet it records higher MAE and MAPE, and the forecast-loss difference is not statistically significant. Inflation and reserve adequacy therefore function as contextual risk indicators rather than sufficient standalone alarms. Policy surveillance should combine them with market-premium, oil-receipt, liquidity, flow, and policy-regime indicators.

Keywords: *naira stress; inflation; external reserves; import cover; exchange-rate depreciation.*

A New Weibull-Gamma Distribution: Theory, Estimation and Application's

G. O. Suleiman ;W. B. Yahya

Al-Hikmah University ; University of Ilorin

**Corresponding Author: Ganiyu Olayiwola SULEIMAN ; agsuleiman19@gmail.com*

In the fields of reliability engineering, survival analysis, and lifetime data modeling, accurately representing the failure times and life durations of systems, components, and organisms is a central concern. Traditional lifetime distributions such as the exponential, Weibull, and gamma distributions have been widely used due to their mathematical tractability and interpretability. However, these classical models often struggle to capture the complexity of real-world data, particularly when hazard rate behaviors vary, including increasing, decreasing, bathtub-shaped, or unimodal patterns. Although the Weibull distribution is popular for its flexibility, it may not sufficiently model datasets where the hazard function deviates from its typical monotonic form. Similarly, while the gamma distribution is effective in many stochastic and queuing contexts, it lacks the versatility to represent certain tail behaviors and multimodal characteristics observed in practice. To address these limitations, statisticians have developed hybrid and compound distributions that merge features from multiple distributions, enhancing both flexibility and applicability. One such development is the Weibull-Gamma distribution, derived by mixing the Weibull and Gamma distributions. The first four moments about the origin, as well as the mean, were calculated for this new distribution. Derived expressions also include the coefficient of variation, skewness, kurtosis, and index of dispersion. In addition, the moment generating function, characteristic function, and Laplace transform were established. Key reliability functions such as the survival function, hazard rate function, and mean residual life were also derived. Parameter estimation was carried out using the Maximum Likelihood Estimation (MLE) method. The goodness-of-fit of the proposed distribution was evaluated against several existing related models using criteria such as the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC), and Bayesian Information Criterion (BIC). These comparisons were based on real-world datasets. The results demonstrated that the Weibull-Gamma distribution outperformed the competing models, making it a promising alternative for modeling real-life lifetime data.

Keywords: *Probability, Gamma, Weibull, Goodness of fit, maximum likelihood method.*

Adaptive Bayesian Polynomial Regression via Structured Gaussian Priors

Akinlolu Adeseye Olosunde; Oluwafemi Lawal Bisiriyu

Obafemi Awolowo University, Ile-Ife, Nigeria

*Corresponding Author: Bisiriyu Oluwafemi Lawal ; olbisiriyu@oauife.edu.ng

Bayesian polynomial regression has become an important framework for modeling nonlinear relationships by incorporating prior information into regression coefficients. The conventional Bayesian polynomial regression model proposed by Bishop (2006) assumes independent Gaussian priors with a common precision parameter for all regression coefficients. Although effective for low-dimensional problems, this assumption applies identical shrinkage across coefficients and may lead to instability and overfitting when high-degree polynomial terms are included. Motivated by this limitation, this paper proposes a Structured Prior Bayesian Polynomial Regression (SPBPR) model in which the scalar prior precision is replaced by a structured covariance matrix that imposes adaptive shrinkage on polynomial coefficients according to their order. The proposed prior allows lower-order coefficients to retain greater flexibility while progressively shrinking higher-order terms, thereby improving numerical stability and reducing unnecessary model complexity. The posterior distribution of the regression coefficients is derived analytically by combining the Gaussian likelihood with the proposed structured Gaussian prior. A closed-form maximum a posteriori (MAP) estimator is subsequently obtained, extending the classical Bayesian polynomial regression estimator as a special case. The performance of the proposed estimator is investigated through an extensive Monte Carlo simulation study under varying sample sizes, noise levels, and polynomial complexities. The proposed method is compared with ordinary least squares regression, ridge regression, and the classical Bayesian polynomial regression model using prediction accuracy, coefficient bias, root mean squared error (RMSE), mean absolute error (MAE), and predictive log-likelihood as evaluation criteria. Simulation results indicate that the proposed model provides more stable coefficient estimates and stronger regularization of higherorder polynomial coefficients while maintaining competitive predictive performance, particularly in highly parameterized polynomial regression settings. The proposed methodology offers a simple yet effective extension of Bayesian polynomial regression for complex nonlinear statistical modeling.

Keywords: *Bayesian polynomial regression; structured Gaussian prior; adaptive shrinkage; maximum a posteriori estimation; polynomial regression;*

Hybrid Robust Estimation Methods for Non-Linear Regression Models

Bassa S. Yakura; Abraham Okolo; Emmanuel Torsen; Akinrefon A. Adesupo

Department of General Studies Education, Federal College of Education, Yola, Nigeria; Department of Statistics, Faculty of Physical Sciences, Modibbo Adama University, Yola, Nigeria; Department of Statistics, Faculty of Physical Sciences, Modibbo Adama University, Yola, Nigeria; Department of Statistics, Faculty of Physical Sciences, Modibbo Adama University, Yola, Nigeria

**Corresponding Author: Shiwaye Yakura Bassa ; fcbassa@gmail.com*

Nonlinear regression models are widely employed across scientific disciplines; however, standard estimation methods such as Ordinary Least Squares (OLS) are highly sensitive to outliers, heavy-tailed error distributions, and predictor-space contamination, leading to biased and inefficient parameter estimates. This study develops and evaluates two hybrid robust estimation methods for nonlinear regression: MM-Estimation with L Regularisation (MM-L) and M-Estimation with Adaptive Huber Loss (MAH). The hybrid estimators are designed to combine the complementary strengths of existing robust techniques, achieving simultaneously high breakdown point, estimation efficiency, and sparsity, to overcome the limitations of individual robust methods. A comprehensive Monte Carlo simulation study was conducted across 240 scenarios covering two nonlinear regression model types (polynomial and exponential), four error distributions (Normal, Uniform, Cauchy, and Exponential), five contamination levels (0%, 5%, 10%, 20%, and 50%), and six sample sizes ($n = 20$ to 1,000). The simulation results demonstrate that MAH achieves the lowest MSE under mild-to-moderate contamination (below 20%), while MM-L becomes the superior estimator under severe contamination (above 20%); both hybrid estimators substantially outperform OLS and all four single robust alternatives (AH, L, M, and MM) across contamination conditions. The empirical validation uses two real-life datasets, the Concrete Compressive Strength dataset (exponential regression framework, $n = 86$, y-space contamination 4.7%, x-space leverage 8.1%) and the Boston Housing dataset (polynomial regression framework, $n = 506$, y-space contamination 11.9%, x-space leverage 3.6%). MAH achieved Rank 1 in both datasets (MSE = 18.270 and 28.643, respectively), with MM-L ranking second in both (MSE = 19.840 and 30.951), confirming that both contamination rates fall below the 20% threshold. The study concludes that both hybrid estimators are theoretically grounded, empirically validated, and practically applicable, with MAH recommended when contamination in either dimension is below 20% and MM-L recommended above that threshold.

Keywords: Robust Nonlinear Regression; MM-estimation; Adaptive Huber Loss; L1 regularization; Non-Normal Error Distribution

Statistical Evaluation of Equity in Access to the Nigerian Education Loan Fund (nelfund) Among Students in Nigerian Tertiary Institutions

Saminu Umar, Paul Sunday, Okwuobodo Monday

Department of Mathematics and Statistics, Umaru Ali Shinkafi Polytechnic, Department of Mathematics and Statistics, Umaru Ali Shinkafi Polytechnic, Department of Mathematics and Statistics, Delta State Polytechnic Ogwashi-Uku, Delta State

**Corresponding Author: Saminu Umar ; saminu.umar@gmail.com*

The Nigerian Education Loan Fund (NELFUND) represents a major policy initiative aimed at improving access to higher education in Nigeria. However, there is limited statistical evidence on whether eligible students have equitable access to the scheme across different categories of tertiary institutions. This study statistically evaluated equity in access to the NELFUND across Nigerian tertiary institutions. A national cross sectional survey was conducted among eligible students selected through a multistage stratified sampling technique from tertiary institutions, comprising federal, state and private universities, polytechnics and colleges of education across Nigeria's six geopolitical zones. Data were collected using a structured questionnaire comprising closed ended items measured on a five point Likert scale. The instrument was validated by experts and demonstrated satisfactory reliability (Cronbach's alpha = 0.70). Data were analysed using descriptive statistics, chi square tests and multivariate logistic regression at a 5% level of significance. The findings revealed significant disparities in access to NELFUND across institution ownership, institution type, geopolitical zone and socioeconomic status ($p < 0.05$), indicating that equitable access to the scheme had not yet been fully achieved. Students enrolled in federal institutions and those from relatively higher socioeconomic backgrounds demonstrated significantly greater access to the programme than their counterparts in other institution categories. The study concludes that although NELFUND represents a significant step toward expanding educational financing in Nigeria, targeted interventions are required to address existing disparities and ensure that eligible students benefit equitably irrespective of institution type or socioeconomic background. The findings provide an evidence based framework for strengthening programme implementation, improving equitable access to higher education and supporting data driven policy decisions aimed at promoting good governance in Nigeria.

Keywords: NELFUND, equity in access, educational financing, logistic regression, good governance, Nigeria.