

From the Editor

The March issue of *Statistics in Transition new series*, presented to our readers hereby, opens this year's series of our quarterly. It contains a set of twelve articles by twenty-eight authors from nine countries (in order of appearance): Poland, USA, Sri Lanka, India, Botswana, Pakistan, Malaysia, Nigeria, and France. The wide spectrum of issues discussed confirms that our journal consistently strives to cover a wide geographical scope while remaining open to a variety of statistical problems that are of interest to experts from different fields of research, world-wide.

It is with great satisfaction that I note that this issue opens with an article by a distinguished member of our Editorial Board, Professor Janusz Wywiał.

Invited papers

Janusz L. Wywiał's paper, *Generalised spatial autocorrelation coefficients*, focuses on properties of coefficients of spatial correlation generalised to the multidimensional case. The main result of the work is the decomposition of the introduced generalised autocorrelation coefficients into the sum of ordinary autocorrelation coefficients, but calculated on the basis of the principal components of the originally observed multidimensional variable. The development is illustrated with an empirical example. The coefficients provide a more detailed and useful description of the spatial relationships of a set of variables characterizing a population.

Research papers

In the article entitled *An expectation-maximization algorithm for logistic regression based on individual-level predictors and aggregate-level response*, **Zheng Xu** proposes an Expectation-Maximization (EM) algorithm to avoid the direct maximization of the complicated likelihood function. Simulation studies have been conducted to evaluate the performance of the EM estimator compared to different estimators proposed in the literature. Two real data-based studies have been conducted to illustrate the use of the different estimators. The EM estimator proves efficient for the logistic regression problem with an aggregate-level response and individual-level predictors.

The next paper *by D. Dilshanie Deepawansa and Priyanga Dunusinghe Selection criteria and targeting the poor for poverty reduction: the case of social safety nets in Sri Lanka* discusses a multidimensional selection criterion for the leading social safety net for Sri Lanka, Multidimensional Deprivation Score Test (MDST). The method

used has been applied to the HIES-2019 data. It showed that exclusion error is less than existing selection criteria when compared with different targeted groups. According to the selection cut-off, Samurdhi/welfare beneficiaries can be identified. In addition, in order to impact poverty, the transfer schemes should be varied concerning the severity of poverty. Otherwise, if all the beneficiaries get same amount of money, the impact on poverty is unlikely to change significantly. In addition to identifying the suitable beneficiaries, MDST helps to compute the contribution of deprivation in every dimension, which is taken into consideration by household or family, community, or geographical level.

Hemani Sharma's and **Parmil Kumar's** article *On survival estimation of Lomax distribution under adaptive progressive type-II censoring* compares the maximum likelihood (ML) estimation and the Bayesian approach for parameter estimation of the Lomax distribution. Additionally, the study aims to determine the approximate intervals for the parameters and the survival function based on adaptive progressive type-II censored data. The ML estimators of the probability distribution parameters were calculated using the Newton-Raphson method, while the delta method was used to compute the approximate confidence intervals for the survival function. The Bayesian approach was also used to estimate the unknown parameters and survival function. This was achieved through the construction of Bayesian estimators under an informative and non-informative prior based on the squared error loss function (SELF) and approximate credible intervals. The Markov Chain Monte Carlo (MCMC) method was employed to test the efficiency of the proposed method in various situations based on different criteria such as mean-squared error, bias, coverage probability, and expected length-estimated criteria.

In the paper entitled *A fuzzy hybrid MCDM approach to the evaluation of subjective household poverty*, **Aleksandra Łuczak** and **Sławomir Kalinowski** propose a comprehensive procedure for constructing a synthetic measure of subjective poverty. This involves aggregating factors describing the present, future, and past, which makes it easier to grasp the feeling of deprivation over time. Methods such as fuzzy TOPSIS and fuzzy hierarchical analysis (FHA) based on the fuzzy sets theory were used for this purpose. This innovative procedure was applied to assess the level of subjective household poverty in Poland based on data from survey research carried out in three stages in 2020 using the CAWI method. The results show that the assessment of household's current level of living conditions is also influenced by past events as well as projections of future developments. Changes in the values of the synthetic index illustrate the trajectory of switching from panic to negation, and attempting to cope with the situation or, alternatively, switching to a state of irritation.

The next paper, *Type I heavy-tailed family of generalized Burr III distributions: properties, actuarial measures, regression and applications*, by **Wilbert Nkomo**, **Broderick Oluyede**, and **Fastel Chipepa**, presents a new family of distributions (FoD)

called type I heavy-tailed odd Burr III-G (TI-HT-OBIII-G) distribution. Several statistical properties of the family are derived along with actuarial risk measures. The maximum likelihood estimation (MLE) approach is adopted in the parameter estimation process. The estimates are evaluated centered on mean square errors and average bias via the Monte Carlo simulation framework. A regression model is formulated and the residual analysis is investigated. Members of the new FoD are applied to heavy-tailed data sets and compared to some well-known competing heavy-tailed distributions. The practicality, flexibility and importance of the new distribution in modeling are empirically proven using three data sets.

Laba Handique, Farrukh Jamal, and Subrata Chakraborty in their article *On a family that unifies the generalized Marshall-Olkin and Poisson-G family of distributions* propose a unification of the generalized Marshall-Olkin (GMO) and Poisson-G (P-G) distributions into a new family of distributions. The density and survival function are expressed as infinite mixtures of an exponentiated-P-G family. The quantile function, asymptotes, shapes, stochastic ordering and Rényi entropy are derived. The paper presents a maximum likelihood estimation with large sample properties. A Monte Carlo simulation is used to examine the pattern of the bias and the mean square error of the maximum likelihood estimators. The utility of the proposed family is illustrated through its comparison with some important models and sub models of the family in terms of modeling real data.

Iwona Skrodzka's paper, *Impact of human capital on the innovation performance of EU economies* attempts to empirically determine the impact of human capital on the innovation performance of EU economies, given a gap in the literature regarding this issue. There are difficulties associated with the measurement as well as the limited number of methods to study the relationships between unobservable variables. In order to fill this gap, the partial least squares structural equation modelling (PLS-SEM) was used, covering the years 2014-2020.

The next article, *Improving detectability of the indicator saturation approach through winsorization: an empirical study in the cryptocurrency market*, by **Suleiman Dahir Mohamed, Mohd Tahir Ismail, and Majid Khan Bin Majahar Ali**, presents a hybrid approach called the Win-IS strategy, focusing on the influence of extreme outliers in the tail and subsequently identify breaks, trend breaks and outliers in cryptocurrencies. The study uses cryptocurrencies like Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Tether (USDT), and Ripple (XRP). This article improves the detectability of the IS approach by combining it with the winsorization strategy and hence proposes a technique known as Win-IS. The performance of Win-IS is then empirically compared to IS in five cryptocurrency markets. The Win-IS strategy outperformed the IS technique, as demonstrated by BIC scores. Furthermore, the Win-IS technique reduced severe outliers in four coins while revealing new outliers, breaks, and trend breaks, some of which were duplicated from the IS results. The repeated

outliers, breaks, and trend breaks show their importance in this market because they remained constant in both winsored and original returns.

M. Dileepkumar's, R. Anand's, and P. G Sankaran's paper, *Reliability properties and applications of proportional reversed hazards in reversed relevation transform*, describes important reliability properties of the reversed relevation transform under the proportional reversed hazards assumption. The results of research on information measures are presented. Various ageing concepts and stochastic orders are discussed. A new flexible generalization of the Fréchet distribution is introduced using the proposed transformation, and reliability properties and applications are discussed. The ageing and stochastic ordering properties of the model were derived.

In the paper *Analytical modelling for COVID-19 data (fatality): a case study of Nigeria for the period of February 2020 – April 2022*, **E. Torsen, U. M. Modibbo, M. Mijinyawa, L. L. Seknewna, and I. Ali** used univariate time series models to analyze the confirmed cases of COVID-19 fatalities (count data and having zero inflation) due to COVID-19 in Nigeria. Specifically, the Autoregressive Integrated Moving Average (ARIMA), Zero-Inflated Poisson Autoregressive (ZIPAR), and Zero-Inflated Negative Binomial Autoregressive (ZINBAR) models were employed. The findings indicate that ZINBAR having the lowest Root Mean Square Error (RMSE), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) outperforms the other two models: hence, the ZINBAR performs better than the ZIPAR and the ARIMA. This demonstrates and emphasized the fact that for count time series data, count time series models should be used, with indication to the ZINBAR to be used to predict and forecast COVID-19 in Nigeria.

Research Communicates and Letters

Stanisław Jaworski's article *Optimal sample size in a triangular model for sensitive questions* considers the nonrandomized response model (proposed by Tian et al., 2007) and introduces a novel CI for the fraction of sensitive questions in the triangular model. Unlike the widely used asymptotic CI, the new approach maintains the prescribed confidence level. The minimum sample size satisfying two criteria was considered: average length and almost sure length. To obtain such sample sizes, the restrictions on privacy protection were imposed, specifically the probability of discovering a YES answer to the sensitive question. This probability should be sufficiently small to ensure the interviewee's comfort in answering the questionnaire.

Włodzimierz Okrasa

Editor

