

From the Editor

It is with great pleasure that we present our readers with the September issue consisting of 12 articles arranged in three sections: *Original Research Papers*, *Conference Papers*, and *Research Communicates and Letters*. Altogether 24 authors from a large set of countries – Australia, Poland, Jordan, India, Nigeria, Indonesia, and Italy – present results of their investigations in a wide spectrum of research areas.

Original Research Papers

The section of original research papers starts with the article by **Ravi Dutta-Powell** entitled *The perils of premature evaluation: reassessing the application of Benford's Law to the USA's COVID-19 data*. In follow-up of a review of earlier applications of Benford's Law to the COVID-19 data in the United States – that claimed these data's non-conformity with Benford's Law – more recent granular data are used to demonstrate that this was likely due to the earlier data being unsuitable for such applications. It also demonstrates that the same dataset, analyzed in different ways, can show vastly different levels of conformity with Benford's Law. Specifically, most US states show high degrees of conformity for the COVID-19 cases and cumulative deaths when the Robust Order of Magnitude (ROM) is over 3 and data at the county level is used to analyze state outcomes. Conversely, when the county data is aggregated to the state level and analyzed (i.e. case totals for all counties are summed to create a single state figure for each day of the pandemic), every state shows non-conformity.

Leszek Morawski's paper *Education expansion and income inequality: evidence from Poland (2005–2019)* shows that educational change reduced poverty and income inequality incidence and depth, and describes the consequences of this change using a microsimulation decomposition based on a tax-benefit microsimulation model. Using a microsimulation approach, the author estimates that the impact of the abovementioned educational change on changing material poverty risk corresponds to 40% of the policy effect associated with changes in tax and benefit rules. For the Gini index, the educational effect amounted to 91% of the policy effect. The results show that educational changes in Poland between 2005 and 2019 significantly impacted income inequality and the risk of material poverty.

In the article *Area-biased one-parameter exponential distribution with financial applications* **Abdullah Hardan** and **Loai Alzoubi** propose the area-biased one-parameter linear exponential distribution the main properties of which – such as the

moments and the related measures, the harmonic mean and the mode – are derived and analyzed using the reliability analysis functions along with the pdfs of the minimum, maximum and the k th order statistics. Additionally, employed are the quantile function, the mean absolute deviations from the mean and the median jointly with the mean waiting and residual lifetime. A simulation study using the MLE, OLS, WLS, MPS, CVM and AD methods of estimating parameters is conducted showing that the estimators are unbiased and consistent. Three real financial data applications prove the goodness of fit for this distribution. They show that the suggested distribution fits the real data better than the competence distributions.

The next paper, *Bayesian estimation of two-parameter power Rayleigh distribution and its application* by Mohd Irfan and Anup Kumar Sharma explores classical and Bayesian approaches to the estimation of unknown parameters and reliability functions for the power Rayleigh distribution. The maximum likelihood estimator (MLE) method is considered in classical estimation. The Bayesian estimation, on the other hand, uses several loss functions under informative and non-informative prior distributions, utilizing the Lindley technique and Markov Chain Monte Carlo (MCMC) method for Bayesian computations. Approximate confidence intervals are established based on the MLEs using the delta technique, while Bayes credible intervals are determined using the MCMC method. A simulation study is conducted to compare the performance of these methods in terms of biases and mean square errors, showing that Bayesian estimators outperform their classical counterparts. Additionally, two real datasets are presented for illustrative purposes.

Krzysztof Brania's and Henryk Gurgul's article *The impact of the COVID-19 pandemic on forecast uncertainty of macroeconomic data releases* focuses on the uncertainty associated with macroeconomic data forecasts measured by the surprise indicator (SI). Moreover, the authors examine whether the distribution of SI depends on the economy, category of indicator or time, considering pre-pandemic, pandemic and post-pandemic periods in the context of the COVID-19 crisis. The construction of a sentiment indicator that is intended to aggregate all information that is jointly released through macroeconomic indicators was also proposed, Macroeconomic data releases are very important benchmarks of the economy. Therefore, the vast majority of financial market analysts and traders closely monitor both the projected estimates and the intuitively more impactful actual values.

The paper by Nureni Olawale Adeboye and Olumide Sunday Adesina discusses *Bayesian and frequentist modelling of West African economic growth: a dynamic panel approach*. As the empirical outcomes of previous studies examining the relationship between economic growth and socio-economic indicators have been inconclusive and contradictory the current research employed an alternative strategy. A dynamic panel model is estimated *via* three robust dynamic panel data estimators of

the generalized method of moment (GMM), frequentist instrumental variable (IV) and the Bayesian IV on real and simulated data. Various model performance criteria such as Wald statistics, leave-out-one cross-validation and the Pareto checks were used for validity verification. The results of the robust diagnostics checks and a model strength metric showed that the family of IV models outperformed the GMM. Thus, the estimation provided by the Bayesian IV is upheld and recommended in modelling dynamic panel data as it provides robust estimates of the parameters of interest.

In the work entitled *Exploring the stochastic production frontier in the presence of outliers: a simulation study*, **Anik Djuraidah** and **Ismail Pranata** present the results of a simulation conducted to compare five SPF models: Normal-half Normal, Normal-Gamma, Normal-Weibull, Normal-Rayleigh, and Student's-t-half Normal. Applying simulated data across nine scenarios with varying data amounts and outlier percentages, the findings prove that the SPF Student's t-half Normal model provides the most accurate prediction of technical efficiency. Using a heavy-tailed distribution, such as the Student's t distribution, for the disturbance component is more effective in handling outliers in the response variable than modifying the inefficiency of the component distribution.

The paper by **Alina Jędrzejczak**, **Małgorzata Misztal**, and **Dorota Pekasiewicz** entitled *Regional differentiation of income distributions in Poland* examines the regional differences in the total income distribution in Poland. Both average income levels and income inequality and poverty parameters are included in the analyzes. The study, based on individual data from the Household Budget Survey, used parametric and non-parametric methods for estimating inequality and poverty measures, as well as cluster analysis methods. In the parametric approach, the empirical income distributions in Poland were approximated using the theoretical Dagum distribution. This enabled the segmentation of voivodships in terms of the estimated characteristics of the equivalent household income distribution. The results confirmed anticipations that income distributions in Poland differ significantly across regions. The obtained clusters allowed identifying groups of regions that may require separate social policies aimed at upholding household income or at reducing income inequality.

Błażej Suproń's paper *Modelling the asymmetric relationship between energy and CO₂ emissions in the Visegrad group: empirical evidence from a panel NARDL approach* presents an attempt to assess the impact of renewable energy consumption, non-renewable energy consumption and economic growth on the volume of carbon dioxide (CO₂) emissions in the Visegrad countries between 1991 and 2021. Using a Nonlinear Autoregressive Distributed Lag (NARDL) model for panel data, the research captures both long-term dependencies and short-term dynamics. The results show that a reduction in CO₂ emissions yielded by a significant long-term decrease

in non-renewable energy consumption is proportionally larger than the increase in the emissions caused by the growth in the consumption of such energy. GDP growth in the V4 countries increases CO₂ emissions, but GDP decline contributes significantly more to the reduction in emissions. On the contrary, renewable energy consumption consistently reduces CO₂ emissions over the long term, with no significant asymmetry detected. In the short term, both economic growth and non-renewable energy consumption increase CO₂ emissions.

Conference Papers

XXXXII Multivariate Statistical Analysis 2024, Lodz, Poland

Stefano Bonnini's and **Michela Borghesi's** paper entitled *Multivariate two-sample permutation test with directional alternative for categorical data* presents a distribution-free test, based on the permutation approach, on treatment effects with a multivariate categorical response variable. It refers to a typical case-control biomedical study, performed to investigate the effect of the treatment called "assisted motor activity" (AMA) on the health of comorbid patients affected by "low back pain" (LBP), "hypertension" and "diabetes". Specifically, the goal was to test whether the AMA determines an improvement in the functionality and the perceived health status of patients. Two independent samples (treated and control group) were compared according to 13 different binary or ordinal outcomes. The null hypothesis of the test assumes the equality in the distribution of the multivariate responses of the two groups, while under the alternative hypothesis, the health status of the treated patients is better. The approach proposed in this work is based on the Combined Permutation Test (CPT) method, which is suitable for analyzing multivariate categorical data in the presence of confounding factors.

XV Scientific Conference MASEP 2024 – Measurement and Assessment of Social and Economic Phenomena, Warsaw, Poland

Abu Feye Bantu, Andrzej Kozyra, and Józef Wiora discuss *Normality tests for transformed large measured data: a comprehensive analysis*. The study emphasizes the importance of assessing and transforming large datasets, such as GNSS measurements, to ensure normality for the validation of parametric statistical tests. The untransformed GNSS latitude data were identified as non-normal using various visual and statistical tests, including histograms, Q-Q plots, skewness, kurtosis, and Statistical tests: KS, AD, DA, SW, JB, CVM, Chi2, and LF. From among the transformation techniques, the rank-based Inverse Normal Transformation (INT) shown relatively higher effectiveness in enhancing data normality, as validated by various testing methods. The efficiency of Statistical tests was assessed using ROC and AUC analysis, which successfully categorized untransformed data as non-normal and transformed data as normal.

These findings underscore the necessity of using tailored transformation methods in large-scale data applications, particularly in geospatial and industrial fields, to enhance the reliability and applicability of parametric statistical methods.

Research Communicates and Letters

Archana Panigrahi, Priyaranjan Dash, and Gopabandhu Mishra present *A minimum variance unbiased estimator of finite population variance using auxiliary information*. A class of estimators of finite population variance (S_y^2) using auxiliary information has been developed under simple random sampling without replacement (SRSWOR) scheme. An attempt has been made to derive the minimum variance unbiased estimator of finite population variance from the proposed class of unbiased estimators. The efficiency of the class of estimators under optimality is compared with the usual unbiased estimator, ratio type estimator, product type estimator, regression type estimator, exponential ratio type estimator, exponential product type estimator, and ratio-in-regression estimator, both theoretically and empirically under general conditions and under bivariate normality. The proposed class of estimator performs better than these estimators under certain realistic conditions. The proposed class of estimators is generalized for the case of multi-auxiliary variables.

Włodzimierz Okrasa

Editor

