From the Editor

The June issue presents a set of 12 articles. Their authors, thirty in total, come from nine countries (in the order of their appearance): Poland, Indonesia, India, Morocco, Azerbaijan, Malaysia, Nigeria, Ukraine, and Canada. Such a global scope of cooperation with our partners – authors, readers and reviewers – confirms the validity of the phrase in the title of the journal, "statistics in transition", with respect to both possible interpretations of its meaning. Highlighting the durability of the dynamic development of the discipline itself, it reflects differences in the institutional development profiles on the one hand and research interests of experts in such a broad international environment on the other. It therefore adds to the thematic richness and diversity of the entire statistical community. On behalf of everyone involved in the preparation of this volume, I wish you pleasant reading.

Research articles

In the first paper, *Volatility and models based on the extreme value theory for gold returns*, **Dominik Krężołek** and **Krzysztof Piontek** focus on the daily gold log-returns process to analyze the quality of forecasting expected shortfalls (ES) using volatility and models based on the extreme value theory (EVT). ES forecasts were calculated for conditional APARCH models formed on the entire distribution of returns, as well as for EVT models. The results of ES forecasts for each model were verified using the backtesting procedure proposed by Acerbi and Szekely. The results show that EVT models provide more accurate one-day ahead ES forecasts compared to the other models. Moreover, the asymmetric theoretical distributions for innovations of EVT models allow the improvement of the accuracy of ES forecasting. The study shows that the process of gold returns is characterized by significant, unpredictable, and heterogeneous volatility (e.g. the apparent effect of the COVID-19 pandemic). Moreover, empirical distributions of gold returns were found to be characterized by clustering of variance, leptokurtosis, asymmetry, and fat tails (compared to a normal distribution).

The article entitled *Robust Spatial Durbin modelling on tuberculosis data using the MM-estimator method* by Ummul Auliyah Syam, Siswanto Siswanto, and Nurtiti Sunusi presents the modelling of a Robust Spatial Durbin Model (RSDM) on tuberculosis data in South Sulawesi Province. The results indicate that RSDM performs better than the Spatial Durbin Model (SDM) in explaining the number of tuberculosis cases in South Sulawesi Province in 2020, as it has a larger Adjusted R value and a smaller MSE value than SDM. Factors that significantly affect the number of tuberculosis cases in South Sulawesi Province in 2020 are population density, the percentage of households with PHBS, the percentage of BCG (Bacillus Calmette-Guérin) immunization, and the percentage of malnutrition. In terms of comparison with other spatial models for tuberculosis or infectious diseases in general, it is important to note that different spatial models may be more appropriate depending on the specific context and available data.

Neha Garg's and Menakshi Pachori's manuscript *Ratio-type estimator of the population mean in stratified sampling based on the calibration approach* presents an improved ratio-type calibrated estimator that was developed using the logarithmic mean in the calibration constraint for the stratified random sampling scheme. The proposed estimator was extended in the case of stratified double sampling and compared with the estimators given by Tracy et al. (2003) together with its ratio-type estimators, as well as Nidhi et al. (2017) and Khare et al. (2022). A simulation study was also carried out on both a real and artificial dataset in order to evaluate the performance of the proposed estimator compared to the existing estimators. The simulation study conducted on real as well as on artificial populations showed that the proposed estimators have less %RRMSE and their value decrease as the sample sizes increase, which in turn explains their performance better than the existing estimators.

In the paper *Modeling the impact of an ISO 9001 certified quality management system on the organizational performance of Moroccan services firms* Ibtissam El Moury, Houda Kacimi, Sara Fennane, and Adil Echchelh discuss the consequences of a noticeable absence of such a system (ISO 9001 certified quality management systems) for organizational performance of companies in Moroccan context, where the services sector contributes more than 50% to the national wealth. The authors constructed a causal model to quantify the strength of cause-and-effect relationships among the following key elements: processes within a quality management system and their impact on organizational performance, the influence of management responsibility process on all aspects of the quality management system, and the relationship between the organizational and financial performance. This model is developed using Structural Equation Modeling (SEM) and estimated through the Partial Least Squares (PLS) approach, utilizing the XL-Stat software.

Ibrahim Niftiyev's paper Dimensionality reduction analysis of the renewable energy sector in Azerbaijan: nonparametric analyses of large datasets analyzes the renewable energy sector in Azerbaijan using Principal Component Analysis (PCA) and Multiple Correspondence Analysis (MCA). The PCA procedure yielded four distinct principle components reflecting the main macroeconomic variables, renewable energy production, industry-energy relations and natural resource revenues. Meanwhile, the MCA-based cross-country assessment of Azerbaijan's wind, solar and hydropower has struck somewhat pessimistic notes, as the country lags behind neighboring and other post-Soviet countries (e.g. Estonia, Iran, Latvia, Russia) in developing its green energy sector. These findings are of great interest to policymakers, businesses and academics who wish to gain deep insight into the Azerbaijani economy in terms of renewable energy production.

In the next article **On the Poisson-transmuted exponential distribution and its application to frequency of claim in actuarial science**, **S. R. M. Sabri** and **A. A. Adetunji** propose a new discrete distribution in the mixed Poisson paradigm to obtain a distribution that provides a better fit to skewed and dispersed count observation with excess zero. The cubic transmutation map is used to extend the exponential distribution, and the obtained continuous distribution is assumed for the parameter of the Poisson distribution. Various moment-based properties of the new distribution are obtained. The Nelder-Mead algorithm provides the fastest convergence iteration under the maximum likelihood estimation technique. The shapes of the proposed new discrete distribution are similar to those of the mixing distribution. Frequencies of insurance claims from different countries are used to assess the performance of the new proposition (and its zero-inflated form). Results show that the new distribution outperforms other competing ones in most cases.

Muzamil Jallal, Aijaz Ahmed, and Rajnee Tripathi in their paper *Extended odd Frechet-exponential distribution with applications related to the environment* attempt to expand the Frechet distribution by employing the T-X family of distributions and named the newly formulated model Extended odd Frechetexponential distribution (EOFED). Several structural properties, reliability measurements and characteristics were estimated and discussed. The authors present graphs which depict the behavior of the probability density function, cumulative distribution function and the hazard rate function. The adaptability and flexibility of this novel distribution were achieved through the application of real-world data sets. A simulation study was also performed to evaluate and compare the output efficacy of the estimators.

In the next paper, *Spatial and component structure analysis of the inclusive circular economy: SGICE* Oleksandr Osaulenko, Alla Shlapak, Iryna Zvarych, Oksana Brodovska, and Kateryna Krysovata propose a methodology (a conceptualisation) of the global inclusive circular economy, which can be considered as a complex

multidimensional system. The main components of it are the economic, sociological, ecological and circular aspects of the country's life. To achieve this goal, the GNU regression, econometrics and time-series library was used – an applied software package for econometric modeling, a part of the GNU project. Accordingly, the authors define the global inclusive circular economy as the SGICE (Global Inclusive Circular Economy) system, characterized by the vector of functions. For the most exhaustive consideration of the entire range of opportunities of the global inclusive circular economy, the study developed and accordingly analyzed the integrated index of the development of the global inclusive circular economy (IGICE) by ecological, economic, social and circular components with isolated weakly correlated indicators.

Devendra Kumar, Maneesh Kumar, Sapna Yadav, and Anju Goyal present *A new parameter estimation method for the extended power Lindley distribution based on order statistics with application*. The authors analyze inference procedures for the estimation of parameters by using order statistics and derive some new expressions for single and product moments of the order statistics from the extended power Lindley distribution. A simulation study is conducted for calculating the best linear unbiased estimates (BLUEs) for the location and scale parameters based on Type-II right-censored samples. Finally, one real data set has been used to obtain the MLEs of the model parameters, BLUEs of the parameters and the EPL distribution is also compared with some existing distributions. It was concluded that the EPL distribution provides the best fit among the compared distributions. A real data set is analyzed to illustrate the flexibility and importance of the model.

Katarzyna Kuryło's and Łukasz Smaga's article Functional repeated measures analysis of variance and its application is devoted to a medical problem when the same group of patients with multiple sclerosis are examined at several successive visits (doctor's appointments) and described by fractional anisotropy tract profiles, which can be represented as functions. Since the observations for each patient are dependent random processes, they follow a repeated measures design for functional data. To compare the results for different visits, functional repeated measures analysis of variance is thus considered. A pointwise test statistic is constructed by adapting the classical test statistic for one-way repeated measures analysis of variance to the functional data framework. By integrating and taking the supremum of the pointwise test statistic, two global test statistics are created. In addition to verifying the general null hypothesis of the equality of mean functions corresponding to different objects, authors propose a simple method for ex-post hoc analysis. The authors illustrate the finite sample properties of permutation and bootstrap testing procedures in an extensive simulation study, and analyze a real data example in detail [All methods are implemented in the R package rmfanova, available on CRAN.].

Other articles

XXXXI Multivariate Statistical Analysis 2023, Lodz, Poland. Conference Papers

Błażej Kochański's paper *The shape of an ROC curve in the evaluation of credit scoring models* focuses on the AUC, i.e. the area under the receiver operating characteristic (ROC) curve, or its scaled version, the Gini coefficient, which are the standard measures of the discriminatory power of credit scoring. Using binormal ROC curve models, it is shown how the shape of the curves affects the economic benefits of using scoring models with the same AUC. It was proposed next that the shape parameter of the fitted ROC curve be reported alongside its AUC/Gini whenever the quality of a scorecard is discussed. This approach can be useful when discussing the quality of existing or newly created scoring models. Knowing the shape of the curve allows one to assess which tasks the model is best suited for. Measuring the shape of the ROC curve can also be useful when evaluating individual component variables or component scorecards included in the combined master model.

Research Communicates and Letters

In the Research Communicates & Letters section a paper Monkeypox obeys the (Benford) law: a dynamic analysis of daily case counts in the United States of America by Leonardo Campanelli is presented. The article analyses the first-digit distribution of the monkeypox daily cases in the United States of America, from May 17 to September 21, 2022. The overall data follow Benford's law, a conclusion substantiated by eight different statistical tests, including the "Euclidean distance test", which has been designed to specifically check Benford's distribution in data. This result aligns with those of other infectious diseases, such as COVID 19, whose Benfordness has already been confirmed in the literature. Daily counts of monkeypox cases, and in general death and confirmed cases counts for any infectious disease, evolve in time. In order to follow the spread of monkeypox dynamically, the author analyzed the temporal deviation of monkeypox counts from Benford's law. In the case of monkeypox in the USA, no anomalies were detected, with the temporal series of daily cases conforming to Benford's distribution to a remarkably high significance level of about 99.96%. [The Statistical test used for the dynamic analysis was the Euclidean distance test.].

Włodzimierz Okrasa Editor

© Włodzimierz Okrasa. Article available under the CC BY-SA 4.0 licence 😇 💓 💿