

From the Editor

A set of eleven articles written by twenty two authors from ten countries places *Statistics in Transition new series (SiTns)* among the top scientific journals whose aim is the internationalisation of scientific research. This leads to a more general reflection on the nature and place of our publication in the constellation of the existing scientific periodicals. The studies provided by authors from outside of Europe and North America, with experts from African countries, the Middle East and South East Asia being particularly active, complement geographically one of the transition-related components emphasised in *SiTns*' mission objectives; the other one has traditionally meant the permanent process of the inherent development of the discipline. Recognising both directions of the promotion of articles (as the title of the journal suggests) – focusing either on the development of the discipline or on the statistical analysis of real issues in various environments (including developed and developing countries) – *SiTns* has been long pursuing the free-access and free-of-charge (for authors/APC) policy. We strive to continue our work in this very way, believing that we provide an important platform (a niche), essential for the presentation of a wide spectrum of research within both aspects, i.e. the authors' background and research topics, ideally combined and introduced in an innovative manner. The papers contained in the issue are examples of such an approach and policy.

Research articles

The article entitled *Acceptance sampling plans from a truncated life test based on the power Lomax distribution with application to manufacturing* by **Amjad D. Al-Nasser** and **Mohammad A. ul Haq** starts with the assumption that the quality characteristic follows the power Lomax distribution. The operating characteristic function values are calculated for the proposed sampling plan, jointly with the optimal sample size and the producer's risk for a selection of distribution parameters. Furthermore, a comparative study with other sampling plans is introduced to demonstrate the advantages of the proposed plan. Finally, a real-life example illustrating the applicability of the proposed sampling plan in a manufacturing company is discussed. Comparisons with other lifetime distributions showed that sampling plans based on PLxD are more efficient than their other counterparts.

In the next paper, *A calibrated synthetic estimator for small area estimation*, **Matthew J. Iseh** and **Ekaette I. Enang** discuss synthetic estimators used to produce

estimates of population mean in areas where no sampled data are available. Given that such estimates are usually highly biased with invalid confidence statements, the paper presents a calibrated synthetic estimator of the population mean which addresses several problematic issues. Two known special cases of this estimator were obtained in the form of combined ratio and combined regression synthetic estimators, using selected tuning parameters under stratified sampling. In result, their biases and variance estimators were derived. The empirical demonstration of the usage involving the proposed calibrated estimators shows that they provide better estimates of the population mean than the existing estimators discussed in the study. In particular, the estimators were examined through simulation under three distributional assumptions, namely the normal, gamma and exponential distributions. The results show that they provide estimates of the mean displaying less relative bias and greater efficiency. Moreover, they prove more consistent than the existing classical synthetic estimator. The further evaluation carried out using the coefficient of variation provides additional confirmation of the calibrated estimator's advantage over the existing ones in relation to small area estimation. The proposed combined ratio synthetic estimator has shown dominance over the combined regression synthetic estimator suggesting that the latter is not suitable for any real-life data that follow exponential distribution for small domains under stratified sampling.

Andrzej Szymański's and **Agnieszka Rossa's** paper entitled *The Complex-Number Mortality Model (CNMM) based on orthonormal expansion of membership functions* deals with a new fuzzy version of the Lee-Carter (LC) mortality model. Mortality rates as well as parameters of the LC model are treated in this model as triangular fuzzy numbers. As a starting point, the fuzzy Koissi-Shapiro (KS) approach is recalled. Based on this approach, a new fuzzy mortality model – CNMM – is formulated using orthonormal expansions of the inverse exponential membership functions of the model components. The paper includes numerical findings based on a case study applying the new mortality model compared to the results obtained with the standard LC model. Moreover, the confidence intervals for log-central mortality rates can also be derived, but they reflect the error term in the random walk model, ignoring the estimation errors of other parameters; thus, the confidence intervals can only be derived for the prediction window.

In the paper *Bayesian estimation and prediction based on Rayleigh record data with applications*, **Abu Awwad R. R.**, **Bdair O. M.**, and **Abufoudeh G. K.**, consider the problem of estimating the scale and location parameters of the model and predicting the future unobserved record data using a record sample from the Rayleigh model. Maximum likelihood and Bayesian approaches under specific loss functions are used to estimate the model's parameters. The Gibbs sampler and Metropolis-Hastings

methods are applied within the Bayesian procedures to draw the Markov Chain Monte Carlo (MCMC) samples, used in turn to compute the Bayes estimator and the point predictors of the future record data. Two examples of real-life data have been analysed to illustrate the developed procedures.

Vipin Sharma and **Vinod Kumar** in their article entitled *Trade potential under SAFTA between India and other SAARC countries: the augmented gravity model approach* attempt to assess the trade potential for the years 1992 to 2019 at annual frequency in general, and for 2004 to 2019 in detail. The findings of this paper show that intra-regional trade volumes between SAARC nations can be increased and encouraged. It is important to undertake structural reforms in order to boost trade with non-member countries. The authors suggest that research on this issue should take into account the effect of locational and infrastructural advantages on transportation costs using a gravity model. Previous research has also shown that an augmented gravity model may help in explaining some key features of South Asian trade which may not be explained by traditional gravity models.

Haitham A. Yousof, **M. Masoom Ali**, **Hafida Goual**, and **Mohamed Ibrahim** in their article *A new Reciprocal Rayleigh Extension: properties, copulas, different methods of estimation and a modified right censored Test for validation* derive the new reciprocal Rayleigh extension's relevant statistical properties. The authors emphasise the results of their research related to convexity and concavity and discuss their estimation of the model's parameters using different estimation methods such as the maximum likelihood estimation method, the ordinary least squares estimation method, the weighted least squares estimation method, the Cramer-Von-Mises estimation method, and the bootstrapping method. The performances of the proposed estimation methods are investigated through a simulation study. Several bivariate- and multivariate-type models have also been derived based on the Farlie Gumbel Morgenstern copula, the Clayton copula, Renyi's entropy copula and the Ali-Mikhail-Haq copula. The modified Nikulin-Rao-Robson test for right censored validation is applied to a censored real data set.

Agata Boratyńska's paper *Robust Bayesian insurance premium in a collective risk model with distorted priors under the generalised Bregman loss* presents a collective risk model relating to insurance claims. The objective is to calculate a premium, which is defined as a functional specified up to unknown parameters. The Bayesian methodology, which combines the prior knowledge about certain unknown parameters with the knowledge in the form of a random sample, has been adopted, along with the generalised Bregman loss function. The results, however, can be applied to numerous loss functions, including the square-error, LINEX, weighted square-error, Brown, entropy loss. Some uncertainty about a prior is assumed by a distorted band class of

priors. A range of collective and Bayes premiums is calculated and posterior regret Γ -minimax premium as a robust procedure has been implemented. Two examples are provided to illustrate the issues considered – the first one with an unknown parameter of the Poisson distribution and the second one with unknown parameters of distributions of the number and severity of claims.

Hajar Bouazzaoui, Mohamed Abdou Elomary, and My Ismail Mamouni discuss *An application of persistent homology and the graph theory to linguistics: The case of Tifinagh and Phoenician scripts* using tools from within topological data analysis and the graph theory for identifying the similarity between the two scripts (Tifinagh and Phoenician). The clustering of their letter shapes is performed based on the pairwise distances between their topological signatures. The ideas presented in this work can be extended to study the similarity between any two writing systems and as such can serve as the first step for linguists to determine the possibly related scripts before conducting further analysis. For instance, a future work might explore the nature of this relatedness, i.e. whether one script is derived from the other or one was built under the other's influence.

In the paper *A new count data model applied in the analysis of vaccine adverse events and insurance claims*, **Showkat Ahmad Dar, Anwar Hassan, Peer Bilal Ahmad and Sameer Ahmad Wani** present a new probability distribution, created by compounding the Poisson distribution with the weighted exponential distribution. Important mathematical and statistical properties of the distribution are derived and discussed. The paper describes the proposed model's parameter estimation, performed by means of the maximum likelihood method. Finally, a real data set is analysed to verify the suitability of the proposed distribution in modelling a count dataset representing vaccine adverse events counts and insurance claims.

Other articles:

38th Multivariate Statistical Analysis 2019, Łódź. Conference Papers

The last paper, by **Wioletta Grzenda**, entitled *Modelling the occupational and educational choices of young people in Poland using Bayesian multinomial logit models* is based on a presentation delivered at the 2019 Multivariate Statistical Analysis conference. A binomial logit model is applied to two states: of being unemployed and employed, or economically inactive and active. The paper focuses on the situation of young people aged 18 to 29 on the labour market in Poland. They were divided into the following groups: the employed and not learning, those combining education with work, the unemployed, learners/students only, and those economically inactive and not at school. All the different models were estimated within the Bayesian approach. The findings show that continuing education by young people may result from their

problems with finding a job; moreover, combining work with education is not the preferred form of professional activity. In addition, the study examines the inequalities observed on the Polish labour market. The paper provided a new insight into how young people enter the labour market in Poland.

Research Communicates and Letters

The Research Communicates and Letters section presents an article entitled ***On improvement of paired ranked set sampling to estimate population mean*** by **Syed Abdul Rehman** and **Javid Shabbir**, who discuss the difficulties involved in the quantification of units in ecological and environmental sampling, relating to time, money, workload, etc. Therefore, the need for efficient and cost-effective sampling methods was identified and addressed by the authors who propose a sampling scheme called Improved Paired Ranked Set Sampling (IPRSS) to estimate the population mean. The performance of the proposed IPRSS is evaluated under perfect and imperfect rankings. A simulation study based on selected hypothetical distributions and a real-life data set show that IPRSS is more precise than RSS, Paired RSS (PRSS) or Extreme

Włodzimierz Okrasa

Editor