STATISTICS IN TRANSITION new series, June 2025
Vol. 26, No. 2, pp. VII–X

From the Editor

The June issue of *Statistics in Transition new series* contains 11 articles by 32 authors from seven countries, namely (in order of appearance): the USA, Poland, India, Egypt, Libya, Indonesia and Nigeria. I am convinced that they will be excellent reading for both theoretically- and practically-oriented experts and enthusiasts of several fields of statistics – just right for the summer holidays.

Invited papers

In the first paper, *Thresholding nonprobability units in combined data for efficient domain estimation,* **Terrance D. Savitsky, Matthew R. Williams, Vladislav Beresovsky** and **Julie Gershunskaya** propose a new method for thresholding or excluding convenience units to minimise the variance of the resulting survey-weighted domain estimator. The authors compare their thresholding method with other such constructions in a simulation study for two classes of datasets based on the degree of the overlap between survey and convenience samples on covariate support. The research results show that excluding convenience units that each express a low probability of appearing in both reference and convenience samples reduces the risk of the occurrence of an estimation error.

Original Research papers

Andrzej Młodak and **Tomasz Józefowski's** article ***Application of Statistical Disclosure Control methods to protect the confidentiality of the 2020 agricultural census microdata*** presents the attempt to develop an efficient disclosure control algorithm for microdata in a statistical portal used for releasing detailed statistical information at various levels of spatial aggregation. The proposed algorithm is based on perturbative methods, such as microaggregation with Gower's distance for categorical variables and the addition of correlated noise for continuous variables. The proposed algorithm can be used to assess the loss of information by measuring distribution disturbances (based on a complex distance that accounts for all measurement scales) and the impact of the Statistical Disclosure Control (SDC) on the strength of correlations between variables (for continuous variables).

Formulation of estimator for population mean in stratified successive sampling using memory-based information by **Sanjoy Majumder**, **Arnab Bandyopadhyay** and **Arindam Gupta** discusses the development of a memory-type estimator for a population mean in stratified successive sampling. The authors use the past sample information together with the current sample information through hybrid, exponentially-weighted moving-averages statistics, and employ the information available on auxiliary variable to construct the proposed estimator. The performance of this estimator is compared with a conventional estimator of a population mean. The results are obtained using both the simulated as well as population survey data. In the light of the findings of the study, the proposed estimator is recommended for survey statisticians for solving real-life research problems.

In ***Inference of dynamic weighted cumulative residual (DWCR) entropy for Burr XII distribution based on progressive censoring*** by **Amal S. Hassan**, **E. A. Elsherpieny** and **Wesal E. Aghel,** the authors introduce the DWCR entropy performing as an additional measure of uncertainty related to the residual lifetime function in several disciplines, including survival analysis and reliability. In this work, the DWCR formula is based on Havarda and Charvat and uses progressive Type II censoring to investigate the implications of the DWCR Tsallis entropy (DWCRTE), DWCR Rényi entropy (DWCRRE), and DWCRHCE for the Burr XII distribution. Both classical and Bayesian methods are used to derive the estimators of these entropy metrics. Assuming independent gamma priors, the authors obtain the Bayes estimator of the suggested measures. Due to the lack of explicit forms, the Metropolis-Hastings approach is used to determine the Bayes estimates for symmetric and asymmetric loss functions. To assess the efficacy of the suggested estimating techniques, several simulations are run for different censoring schemes.

The paper ***Synchronization and similarity between regional and sectoral output gaps in the Polish manufacturing industry*** by **Mirosław Błażej, Mariusz Górajski** and **Magdalena Ulrichs** presents a new micro-econometric model to calculate regional and sectoral output gaps in Poland's manufacturing industry in the years 2009–2020. The model was estimated using official data sources, which included business tendency surveys and annual enterprise activity reports. The authors assessed the synchronization and similarities between the sectoral and regional output gaps in Poland’s manufacturing industry. They also analyzed the impact of the COVID-19 pandemic on the coherence of the regional and sectoral output gaps. The study contributes to the unification of the method for estimating and assessing the levels of unobservable potential production.

Sapriadi Rasyid, **Siswanto Siswanto** and **Sitti Sahriman’s** article, ***Clustering based on poverty indicator data using K-Means cluster with Density-Based Spatial Clustering of Application with Noise*** describes a statistical method to cluster people affected by poverty on the basis of error indicators for each region, serving as a reference for providing assistance. The cluster analysis shown to be appropriate due to minimizing object differences within one cluster and maximizing them between clusters. The study employs two methods, namely K-Means and Density-Based Spatial Clustering of Application with Noise (DBSCAN), to compare their effectiveness based on the Silhouette Coefficient. The data used for the analysis comes from eight poverty indicators for the South Sulawesi province in 2022. The study demonstrates that the K-Means method is more effective than the DBSCAN in helping the government to group the poverty characteristics of each region.

In **Mbanefo S. Madukaife**, **Uchenna C. Nduka** and **Everestus O. Ossai's** paper ***Testing for multinormality with goodness-of-fit tests based on phi divergence measures***, the authors use the phi divergence measure, DΦ(F,G), between F and G distributions to obtain a goodness-of-fit test to multivariate normality (MVN) based on the theoretical density function of the beta transformed random variable and a window-size spacing-sample density function. Three versions of the statistic are derived from the three known phi divergence measures that are based on the sum of squares. The empirical critical values of the statistics are obtained and the empirical type-one-error rates as well as powers of the statistics in comparison with those of other well-known competing statistics are computed through extensive simulation study. The study shows that the new statistics have good control over type-one-error and are highly competitive with the existing well-known ones in terms of power performance.

Marta Marszałek's article ***Households' invisible input to the economy: a review of its measurement methods and results*** discusses the unpaid domestic work as the main part of the non-market household production, which is not covered by official statistics (GDP). The monetary value of unpaid work is identified within the gross value added (GVA), which is 60-80% of (the invisible) non-market household production. This paper shows that different wages used in input methods do not change the final proportion of the gross value added GVA of unpaid work to total household production. The analysis confirms that a regular implementation of the Household Satellite Accounts (HHSA) jointly with the core system – the European System of Accounts (ESA) – is a valuable tool for assessing the total output of household production.

In the paper ***Ratio regression type estimators of the population mean for missing data in sample surveys*** by **Prachi Garg, Namita Srivastava** and **Manoj Kumar Srivastava,** new ratio regression-type estimators with imputation are proposed as a means to overcome the problem of missing data for a studied variable in a survey. It is shown that the suggusted estimators are more efficient than the mean method of imputation, the ratio method of imputation, the regression method of imputation, or several other estimators. The authors derive the bias and the mean square errors of the suggested estimators and conduct a comprarative study using real and simulated data. The results seem tobe improvement as compared to all the methods discussed.

Conference papers

XV SCIENTIFIC CONFERENCE

MASEP 2024 -Measurement and Assessment of Social and Economic Phenomena, Warsaw, Poland

In the paper ***Harnessing AI for business transformation: strategies for effective implementation and market advantage***, **Jakub Kubiczek, Elżbieta Roszko-Wójtowicz, Julianna Koczy, Izabela Waszkiewicz** and **Klaudia Woś** check in what ways AI-driven tools contribute to business transformation, focusing on their impact on the operational efficiency, customer engagement and market competitiveness. The research employs a multi-method approach, including literature reviews, secondary data analyses and case studies of the AI use in selected enterprises in Poland. The authors' findings highlight both the positive and the negative results of using AI. While it enhances productivity, accuracy and sustainability, businesses must also navigate risks related to data security, compliance, and financial feasibility. The study underscores the importance of dynamic capabilities in leveraging AI for strategic growth while mitigating associated challenges. The results contribute to the discourse on AI's role in shaping modern commerce, offering practical insights for companies seeking to integrate AI-driven solutions effectively.

Research Communicates and Letters

Mateusz Borkowski and **Ewa Gruszewska's** study presented in their paper ***The integrity of the innovation process on the example of EU countries: a PLS-SEM approach*** assesses the integrity of the elements of innovation process and their efficiency with regard to EU countries on the basis of the Schumpeter concept. The study applies the partial least squares structural equation modeling (PLS-SEM), which allows the analysis of the latent variables (LVs). On the basis of the PLS-SEM models for 2010 and 2020, it is concluded that innovation proceed in an integrated manner in EU countries. Not only did the modeling results indicate a positive and moderate effect of the invention inputs on the innovation efficiency LV, but they also showed a positive and strong influence of innovation efficiency LV and the innovation diffusion LV in the analyzed countries. The technological process integrity of EU economies was lower in 2020 than in 2010. In order to improve the functioning of innovation activities, it is necessary to increase technology inputs and the efficiency of their use in R&D.

Włodzimierz Okrasa
Editor