STATISTICS IN TRANSITION new series, December 2024
Vol. 25, No. 4, pp. VII–XI

From the Editor

With this issue of Statistics in Transition new series, which we present to our Readers as the last one in 2024, we begin the first year of a new decade of its publication – the fourth since its launch in 1993. With 10 articles by authors from 8 countries – Slovakia, Poland, Ghana, USA, India, Nigeria, Bangladesh, and Iraq – this issue demonstrates the continuity in serving a broad, international community of experts in various fields of statistics. Both as an academic discipline and as application-oriented fact-based knowledge useful for solving problems in various areas of decision-making and other practical activities.

Wishing you a Happy New Year, we hope that a sense of true professional satisfaction will accompany our Readers while reading this issue, as well as all subsequent ones in the upcoming 2025.

Research articles

The first paper by **Erik Šoltés**, **Silvia Komara**, **Martina Košíková**, and **Tatiana Šoltésová** entitled ***Comparison of household work intensity in Slovakia and Czechia through least squares means analysis based on GLM*** addresses the problem of assessing the impact of relevant factors and their interactions on the *work intensity* (WI) of households in Slovakia and Czechia. For this purpose, general linear models, contrast analysis and estimates of marginal means are employed. The presented analyses are based on the EU-SILC 2021 data and carried out for Slovakia and Czechia. The paper reveals the common and different features of these countries in terms of WI of households. Particular attention is devoted to the identification of the profiles of persons at high risk of living in QJ households. The paper provides estimates of the marginal means of WI households for employed, unemployed and disabled persons with different education, age and from different types of households in Slovakia and in Czechia.

Tomasz Panek, **Jan Zwierzchowsk**i, and **Jan Kroszka** in the article ***The impact of the COVID-19 pandemic on the financial situation of people aged 50+ based on SHARE data*** describe the changes in the financial situation of households of people aged 50+ during the COVID-19 pandemic. The authors evaluate the outcome of the introduced national policies, the EU countries' economic performance, labor market conditions and the individual characteristics of the financial situation of the members

of the examined households. To achieve this goal, an original synthetic index was constructed to measure the changes in the overall financial situation of the surveyed group of households. This index combines various indicators, including income, subjective income assessment, the use of savings to finance current consumption and the postponement of bill payments, allowing a comprehensive evaluation of the shifts in the financial status of the 50+ population during the pandemic. In addition, the study aimed at examining how the age of respondents is interlinked with the changes in their financial situation using data from the Survey of Health, Ageing, and Retirement in Europe (SHARE). The study's findings show that during the pandemic, the changes in the financial situation of households with people aged 50+ varied across the selected countries.

In the next paper, ***Modelling Tinnitus Functional Index reduction using supervised machine learning algorithms***, prepared by **Edmund Fosu Agyemang**, the reduction in the Tinnitus Functional Index (TFI) is attempted utilizing supervised machine learning algorithms, focusing primarily on Ordinary Least Squares (OLS), K-Nearest Neighbor (KNN), Ridge, and Lasso regressions. The analysis highlighted Group, ISI, and SWLS as significant predictors of TFI reduction, identified through the best subset selection and confirmed by both forward and backward selection criteria in the OLS regression. Notably, the shrinkage methods, Ridge and Lasso regressions, demonstrated superior performance compared to OLS and KNN, with the Ridge regression presenting the smallest test mean square error (MSE). This finding establishes the Ridge regression as the best model for analyzing our Tinnitus dataset relative to the other methods. This research highlights the potential of supervised machine learning algorithms in advancing personalized Tinnitus treatment, reflecting broader trends in the field as evidenced by studies in the literature.

Housila P. Singh's, **Rajesh Tailor's**, and **Priyanka Malviya's** paper entitled ***Efficient use of auxiliary information in estimating finite population variance in sample surveys*** discusses the problem of estimating the finite population variance of the study variable y using information on the known population variance of the auxiliary variable x in sample surveys. The bias and mean squared error of the suggested class of estimators up to the first order of approximation was obtained. Preference regions were derived under which the suggested class of estimators is more efficient than the usual unbiased estimator, Das' and Tripathi's estimators (1980), Isaki's ratio estimator (1983), Singh's et al. estimator (1973, and Gupta's and Shabbir's estimator (2007). An empirical study as well as simulation study were carried out in support of the presented study.

Sebastian Wójcik's paper ***AMUSE: Analysis of mobility using simultaneous equations. Present population of refugees in Poland*** describes the significant influx of Ukrainian refugees into Poland following the escalation of the conflict in Ukraine after

February 2022. It highlights the challenges in tracking refugee movements using traditional statistical and administrative data sources due to problems associated with timeliness and spatial granularity. As a result, official statistics are turning to big data sources, such as mobile network operator (MNO) data, to supplement existing data. The paper focuses on utilizing synthetic MNO daily data from SIM cards issued to Ukrainian refugees by a Polish MNO. It proposes AMUSE model: mobility model for data de-duplication and a simple estimator for estimating the present refugee population based on aggregated signaling data over time and areas. Further research shall be focused on including data variability over time into modelling.

The article ***Forecasting under-five child mortality in Bangladesh: progress towards the SDGs target by 2030*** by **Sacchidanand Majumder** and **Soma Chowdhury Biswas** explores the under-five child mortality trends and estimates projection by 2030 to make progress towards achieving the SDGs target regarding the under-five child mortality in Bangladesh. Child mortality is a crucial indicator of a nation's socio-economic advancement and the well-being of mothers, reflecting the overall quality of life within a society. The yearly dataset regarding mortality among children aged five and under (per 1,000 live births) in Bangladesh employed in this study was collected from the World Bank Databank (https://data.worldbank.org/indicator) for the 1972–2022 period. The selection of the best-fitted model for the purpose of forecasting was between the ARIMA model and the Double Exponential Smoothing Holt's Method. Compared with the ARIMA (1,2,1) model, the Double Exponential Smoothing Holt's method proved the best-fitted model for forecasting the under-five child mortality in the future. The results show that under-five child mortality in Bangladesh is an annually declining trend. The average under-five child mortality is forecasted to drop by one during the 2023–2035 period. Thus, the predicted value of under-five child mortality would be 26 in 2025 and 22 in 2030; the national target is 27 (per 1,000 live births) in 2025 and the SDGs target (25 deaths per 1,000 live births).

Kuntal Bera and **M. Z. Anis** in the paper ***On some statistical properties of a stationary Gaussian process in the presence of measurement errors*** discuss some statistical properties, including the mean and variance of a stationary Gaussian process when observed data are affected by measurement errors. As a special case, the authors debate a stationary autoregressive process of order one with Gaussian white noise where measurement error follows an independent Gaussian distribution. To estimate some inferential results based on the collected sample, sometimes the mean and variance of the sample are required to be estimated. The results of this study show that the theoretical values based on the obtained results and the estimated sample values are reasonably close. Here, the authors have considered a particular case of a stationary Gaussian process, namely an AR(1) process. But there are more stationary Gaussian processes other than the AR(1) process. Also, in this paper, it is assumed that the

measurement error follows an independent Gaussian distribution but there are some situations where the measurement error does not follow Gaussian distribution.

The next paper, by **Arkadiusz Derkacz,** entitled ***A method of estimating the Return on Housing Investment (ROHI)*** presents a method for estimating the profitability level of housing investments. Market practice shows that the profitability of this type of investment is influenced by specific determinants that are absent in the classical approach to profitability analysis. The most commonly used method is the Return on Equity (ROE) ratio, which is dedicated to enterprises. However, housing investments are becoming an increasingly popular form of investment among private individuals. This makes the classical ROE method proved suboptimal for such investments ventures (i.e. those that involve the purchase of a residential property and its subsequent rental to third parties). In this context,-an attempt was made to develop a method that would allow to estimate the profitability level of this type of investment. It was found that the ROHI method enables the estimation of the profitability level, taking into consideration the most important determinants characteristic of this type of investment.

Other articles

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

Piotr B. Nowak's paper ***Estimation of the Cox model with grouped lifetimes*** presents how random numbers can be used to transform grouped lifetimes into a pseudo-complete sample. The aim of the study is to investigate the Fisher consistency of the partial likelihood estimator of the regression parameters in the Cox model based on the restored sample. It has been proven that for elliptical-type distributional assumptions about explanatory variables the estimators of the regression parameters in the Cox model based on the pseudo-complete sample are consistent up to a scaling factor. A simulation study illustrates the asymptotic properties of the estimates. In addition, real data case analysis is presented. The importance of the discussed problem is due to the fact that initial data are often aggregated and then classical methods based on the assumption of continuity of the dependent variable are limited. Therefore, the presented randomization method can also be used in other regression models, where a dependent variable is grouped.

Research Communicates and Letters

Ahmed Mahdi Salih and **Murtadha Mansour Abdullah** in their article entitled ***Comparison between classical and Bayesian estimation with joint Jeffrey's prior to Weibull distribution parameters in the presence of large sample conditions*** proposed a comparison of Weibull distribution parameters under large sample conditions. Three methods for estimating parameters: Maximum Likelihood Estimator, Moment Estimator, and Bayesian Estimator with a non-informative prior (weak prior with

minimal influence) are being evaluated. The classical estimation methods of Weibull distribution parameters have been chosen by the authors to the study, including the maximum likelihood estimator and moments estimation (ME). These methods were compared with the Bayesian estimation method (BE) with Jeffrey's prior function. The authors validated the proposed study via simulation using both small and large samples. To determine the best estimation method, mean square errors (MSEs) were used. The simulation findings suggest that maximum likelihood estimators are reasonably effective when using small sample sizes; in cases where the sample size is larger, the BE performs more effectively for both scale and shape parameters of the Weibull distribution function.

Włodzimierz Okrasa
Editor