symmetry

MDPI

Editorial

Symmetrical and Asymmetrical Distributions in Statistics and Data Science

Arne Johannssen [1,*], Nataliya Chukhrova [1] and Quanxin Zhu [2]

[1] Faculty of Business Administration, University of Hamburg, 20146 Hamburg, Germany; nataliya.chukhrova@uni-hamburg.de
[2] School of Mathematical Sciences and Statistics, Hunan Normal University, Changsha 410081, China; zqx22@126.com
* Correspondence: arne.johannssen@uni-hamburg.de

Probability distributions are a fundamental topic of Statistics and Data Science that is highly relevant in both theory and practical applications. There are numerous probability distributions that come in many shapes and with different properties. In order to identify an appropriate distribution for modeling the statistical properties of a population of interest, one should consider the shape of the distribution as a crucial factor. In particular, the symmetry or asymmetry of the distribution plays a decisive role.

The objective of this Special Issue, entitled "Symmetrical and Asymmetrical Distributions in Statistics and Data Science", is to highlight the importance of symmetrical and asymmetrical distributions in its thematic breadth and with applications in many fields. We welcomed submissions related to the latest developments in the area of symmetrical and asymmetrical distributions in Statistics and Data Science. The response from the scientific community was remarkable: 39 papers were submitted for consideration, and 13 papers were finally accepted after a rigorous peer-review process. The remainder of this editorial contains a summary of the contributions to this Special Issue, ordered by date of publication.

Hu et al. [1] study the performance of the modified one-sided Exponentially Weighted Moving Average (EWMA) $\bar{X}$ control chart for monitoring normally distributed processes that are characterized by a perfectly symmetrical shape. Using Monte Carlo simulations and a real data application in semiconductor manufacturing, the authors demonstrate the properties and features of the proposed chart. They also show that the performance of the chart can be further increased by adding the Variable Sampling Interval (VSI) feature to the monitoring procedure.

Shih et al. [2] are concerned with copulas that can be either symmetric (e.g., Gaussian copula) or asymmetric (e.g., Clayton copula). As there are only a few studies on copula-based bivariate meta-analysis, the authors develop the corresponding methodology and theory, specifically to estimate the common mean vector, in this paper. For direct implementation, the authors also provide the `R` package `CommonMean.Copula`. To illustrate the practical applicability of the proposed methods, the authors conduct two applications with real data in the educational and medical sectors.

Alyami et al. [3] present a new class of statistical distributions in their paper, the so called type II half-Logistic odd Fréchet-G class. This class contains various distributions of symmetrical or asymmetrical shapes. The authors discuss four of these specific distributions in the paper, derive their statistical properties, and conduct real-life applications using biomedical, engineering, environmental, and manufacturing data sets.

Hu et al. [4] propose a modified EWMA control chart to monitor the ratio of two normally distributed random variables. As in [1], the authors consider their chart with and without the VSI feature. Based on a simulation study and a real data application in the food industry, the authors illustrate the properties and features of the control chart. They conclude that the proposed chart (1) with VSI is superior and (2) the VSI-based chart also leads to better results than previous approaches in the literature.

Citation: Johannssen, A.; Chukhrova, N.; Zhu, Q. Symmetrical and Asymmetrical Distributions in Statistics and Data Science. *Symmetry* **2023**, *15*, 2140. https://doi.org/10.3390/sym15122140

Received: 5 September 2023
Accepted: 23 November 2023
Published: 1 December 2023

Haj Ahmad and Almetwally [5] discuss three discretization methods to formulate discrete analogues of the continuous generalized Pareto distribution. The authors use Bayesian inference techniques to estimate the discrete models with different symmetric (squared error) and asymmetric (linear exponential and general entropy) loss functions. They perform a comparative analysis and find that the discretized generalized Pareto distribution is a promising alternative. The practical applications are conducted using medical data (COVID-19 daily deaths in the United States and in Italy).

Harke et al. [6] present a data-driven method to estimate the spatial autoregressive dependence on irregular lattices. As the structure of the spatial weights matrix **W** is difficult to obtain in practical applications, they introduce a method to obtain **W**, whether it is symmetric or asymmetric. The authors verify their method using simulation and comparative studies. They also discuss a practical application on the evolution of sales prices for building land in Brandenburg, Germany, and conclude that this evolution and its spatial dependence are mainly driven by the orientation towards Berlin.

Khan and Paramasivam [7] explore how feedback, balking, retaining reneged clients, and the quality control technique affect the encouraged arrival queuing model. The encouraged arrival is valuable for many different businesses in terms of managing operations, deliberating, outlining, implementation, and service development. The authors derive performance measures for the expected number of units in the system, the average number of occupied services and the expected waiting time in the system as well as in the queue.

Haj Ahmad et al. [8] provide a statistical analysis of the alpha power inverse Weibull distribution under a hybrid type II censoring scheme. Their study is motivated by applications in the medical field, namely ball bearings and the resistance of guinea pigs. In these applications, type II censored schemes are recommended to minimize the experimental time and cost where the components are following the alpha power inverse Weibull distribution. The authors conclude that this distribution under a hybrid type II censoring scheme is suitable to model real biomedical data.

Alotaibi et al. [9] investigate parameter estimation, reliability, and hazard rate functions of the Fréchet distribution based on generalized type II progressive hybrid censored data. The Bayesian estimators are computed with independent gamma conjugate priors using the symmetrical squared error loss function. The authors perform comprehensive Monte Carlo simulations and discuss practical applications in physics (precipitation in Minneapolis–Saint Paul) and engineering (vehicle fatalities in South Carolina).

Arslan et al. [10] introduce a control chart for monitoring the process mean based on two supplementary variables. These variables are correlated with the study variable in the form of a regression estimator that is an efficient and unbiased estimator for the process mean. The authors state that the proposed charting scheme performs effectively when both supplementary variables are uncorrelated. The applicability of the proposed chart is shown within a real-data example based on carbon fiber manufacturing data.

Rehman et al. [11] deal with parametric regression analysis of survival data using the additive hazards model with competing risks in the presence of independent right censoring. The baseline hazard function is parameterized using a modified Weibull distribution as a lifetime model. The parameters are estimated using maximum likelihood and Bayesian methods, and the asymptotic confidence interval and the Bayes credible interval of the parameters are derived. The finite sample behavior of the estimators is investigated through simulations, and the model is applied to liver transplant data.

Haj Ahmad et al. [12] analyze a multi-component stress–strength system that provides a useful framework to evaluate the reliability of dams and their ability to cope with external influences such as water pressure, earthquake activity, and erosion. The authors suggest the Gumbel type II distribution as a suitable model for fitting related data. Both classical and Bayesian approaches are used to estimate the reliability function, and Monte Carlo simulations are employed for parameter estimation. As a real case study, the paper considers data from the Shasta reservoir in the United States.

Alotaibi et al. [13] investigate the difficulties associated with estimating the model parameters and reliability time functions of the Kavya Manoharan Kumaraswamy distribution based on generalized type II progressive hybrid censoring. Using the symmetrical squared error loss function, independent gamma conjugate priors are employed to compute the Bayesian estimators. As the Bayesian estimators cannot be derived analytically, the authors implement Markov Chain Monte Carlo (MCMC) techniques. A practical application is provided using a data set on the tensile strength of polyester fibers.

Finally, we would like to congratulate all the above authors on the acceptance of their paper(s) to this Special Issue. We hope that this Special Issue will inspire further researchers to make important discoveries and contributions in the field of "Symmetrical and Asymmetrical Distributions in Statistics and Data Science". To provide the best possible platform for further significant contributions, Volume II of this Special Issue in *Symmetry* is already established (`https://www.mdpi.com/journal/symmetry/special_issues/A28X1PLF7Y`).

We would therefore like to invite researchers to submit their contributions to Volume II of this Special Issue. A potential topic of interest could be the remarkable revival of the hypergeometric and the negative hypergeometric distribution in various areas such as healthcare [14], manufacturing [15], risk management [16], statistical inference [17], and beyond. We look forward to many high-quality papers and a successful second edition of this Special Issue.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hu, X.; Zhang, S.; Sun, G.; Zhong, J.; Wu, S. Modified One-Sided EWMA Charts without- and with Variable Sampling Intervals for Monitoring a Normal Process. *Symmetry* **2022**, *14*, 159. [CrossRef]
2. Shih, J.-H.; Konno, Y.; Chang, Y.-T.; Emura, T. Copula-Based Estimation Methods for a Common Mean Vector for Bivariate Meta-Analyses. *Symmetry* **2022**, *14*, 186. [CrossRef]
3. Alyami, S.A.; Babu, M.G.; Elbatal, I.; Alotaibi, N.; Elgarhy, M. Type II Half-Logistic Odd Fréchet Class of Distributions: Statistical Theory and Applications. *Symmetry* **2022**, *14*, 1222. [CrossRef]
4. Hu, X.; Sun, G.; Xie, F.; Tang, A. Monitoring the Ratio of Two Normal Variables Based on Triple Exponentially Weighted Moving Average Control Charts with Fixed and Variable Sampling Intervals. *Symmetry* **2022**, *14*, 1236. [CrossRef]
5. Haj Ahmad, H.; Almetwally, E.M. Generating Optimal Discrete Analogue of the Generalized Pareto Distribution under Bayesian Inference with Applications. *Symmetry* **2022**, *14*, 1457. [CrossRef]
6. Harke, F.H.; Merk, M.S.; Otto, P. Estimation of Asymmetric Spatial Autoregressive Dependence on Irregular Lattices. *Symmetry* **2022**, *14*, 1474. [CrossRef]
7. Khan, I.E.; Paramasivam, R. Reduction in Waiting Time in an M/M/1/N Encouraged Arrival Queue with Feedback, Balking and Maintaining of Reneged Customers. *Symmetry* **2022**, *14*, 1743. [CrossRef]
8. Haj Ahmad, H.; Almetwally, E.M.; Rabaiah, A.; Ramadan, D.A. Statistical Analysis of Alpha Power Inverse Weibull Distribution under Hybrid Censored Scheme with Applications to Ball Bearings Technology and Biomedical Data. *Symmetry* **2023**, *15*, 161. [CrossRef]
9. Alotaibi, R.; Rezk, H.; Elshahhat, A. Computational Analysis for Fréchet Parameters of Life from Generalized Type-II Progressive Hybrid Censored Data with Applications in Physics and Engineering. *Symmetry* **2023**, *15*, 348. [CrossRef]
10. Arslan, M.; Anwar, S.; Gunaime, N.M.; Shahab, S.; Lone, S.A.; Rasheed, Z. An Improved Charting Scheme to Monitor the Process Mean Using Two Supplementary Variables. *Symmetry* **2023**, *15*, 482. [CrossRef]
11. Rehman, H.; Chandra, N.; Emura, T.; Pandey, M. Estimation of the Modified Weibull Additive Hazards Regression Model under Competing Risks. *Symmetry* **2023**, 15, 485. [CrossRef]
12. Haj Ahmad, H.; Ramadan, D.A.; Mansour, M.M.M.; Aboshady, M.S. The Reliability of Stored Water behind Dams Using the Multi-Component Stress-Strength System. *Symmetry* **2023**, *15*, 766. [CrossRef]
13. Alotaibi, R.; Almetwally, E.M.; Rezk, H. Reliability Analysis of Kavya Manoharan Kumaraswamy Distribution under Generalized Progressive Hybrid Data. *Symmetry* **2023**, *15*, 1671. [CrossRef]
14. Chukhrova, N.; Johannssen, A. Monitoring of high-yield and periodical processes in health care. *Health Care Manag. Sci.* **2020**, *23*, 619–639. [CrossRef] [PubMed]
15. Johannssen, A.; Chukhrova, N.; Celano, G.; Castagliola, P. A number-between-events control chart for monitoring finite horizon production processes. *Qual. Reliab. Eng. Int.* **2022**, *38*, 2110–2138. [CrossRef]

16. Johannssen, A.; Chukhrova, N.; Castagliola, P. Efficient algorithms for calculating the probability distribution of the sum of hypergeometric-distributed random variables. *MethodsX* **2021**, *8*, 101507. [CrossRef] [PubMed]
17. Chukhrova, N.; Johannssen, A. Randomized vs. non-randomized hypergeometric hypothesis testing with crisp and fuzzy hypotheses. *Stat. Pap.* **2020**, *61*, 2605–2641. [CrossRef]