scientific reports

OPEN

A study on the new logarithmic-G family with statistical properties, simulations, and different data analysis

Aijaz Ahmad[1], Afaq Ahmad[2], Yusra A. Tashkandy[3], M. E. Bakr[3], Getachew Tekle Mekiso[4]✉, Oluwafemi Samson Balogun[5] & Ahmed M. Gemeay[6]

In probability theory, it is acknowledged that traditional probability distributions often struggle to represent real-world data characterized by non-monotonic hazard rate behaviour accurately. Researchers are actively working to enhance these distributions to better capture real-life datasets' complexities. This study presents an innovative modelling technique that utilizes logarithmic functions within a family of distributions, providing greater adaptability and potential. Various mathematical properties were investigated and defined for the proposed methodology. Several estimation methods were employed to estimate the model parameters. To assess the effectiveness of these estimators, a comprehensive Monte Carlo simulation was conducted, focusing on bias, mean square error and mean relative error. The New Log-Weibull (NLW) distribution was then applied to analyze multiple COVID-19 and engineering datasets and compare its performance against alternative distributions.

Keywords Logarithmic function, moments, Reliability indicators, Rényi entropy, Estimations

Probability is vital for numerous scientific areas, especially in biology, medicine, and engineering, where it is employed to evaluate and predict lifelong events. Probability offers a framework for measuring uncertainty, allowing investigators to study complicated data and make informed decisions based on statistical findings. In biological sciences, for example, probabilistic models are essential for understanding and projecting lifetime outcomes such as survival rates, rehabilitation timeframes, and the progression of diseases. To analyze lifespan datasets, researchers have proposed several parametric continuous distributions, including the log-normal, Weibull, gamma, Rayleigh, and exponential distributions. Notably, the survival functions of the log-normal and gamma distributions do not possess closed mathematical formulas, which requires the use of computational integration to fully characterize their properties. In lifespan analysis, the exponential and Rayleigh distributions are commonly employed; however, these distributions may not provide the necessary flexibility to accommodate complex datasets effectively. While the exponential distribution represents data with a constant failure rate, the Rayleigh distribution is suitable only for datasets exhibiting an increasing failure rate. In contrast, the Weibull distribution-often referred to as the super-exponential distribution-offers greater adaptability by successfully modelling data with a monotonic hazard rate function, whether it is increasing, decreasing, or constant. This flexibility stems from its ability to combine features of both the exponential and Rayleigh distributions. Nonetheless, the Weibull distribution has its limitations, particularly in modelling datasets with monotonic failure rate patterns, such as unimodal, modified unimodal, or bathtub-shaped data. In our study, we focus on analyzing COVID-19 data, where understanding lifespan phenomena-such as recovery times and patient survival rates-necessitates robust statistical modelling. This context underscores the importance of employing versatile distributions that can effectively capture the unique characteristics of the data under examination. Investigators have recently made significant contributions to the development of new families of adaptable continuous probability distributions. These adaptable distributions are formulated by introducing additional parameters to the fundamental distributions. Among the recently proposed families of lifetime distributions, some noteworthy ones include Eghwerido et al.[1], Reyad et al.[2], Eugene et al.[3], Zagrofos and Balakrishnan[4],

[1]Department of Mathematics, Bhagwant University, Ajmer, India. [2]Department of Mathematical Science IUST, Awantipora Kashmir, India. [3]Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia. [4]Department of Statistics, Wachemo University, Hosaina, Ethiopia. [5]Department of Computing, University of Eastern Finland, 70211 Kuopio, Finland. [6]Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt. ✉email: getch.55tekle@gmail.com