विद्ययाऽमृतमश्नुते
नेपाल प्राध्यापक संघ
२०१८
NJ: NUTA

Multiple Parameter Generalized Rayleigh distribution with Application to Real Dataset

Lal Babu Sah Telee[1], Arun Kumar Chaudhary[2*], Umesh Kumar Yadav[3] and Vijay Kumar[4]
[1,2]Department of Management Science, Nepal Commerce Campus, Tribhuvan University, Nepal
[3]Department of statistics, Mahendra Bindeshwori Multiple Campus, Tribhuvan University, Nepal
[4]Department of Mathematics and statistics, DDU Gorakhpur University Gorakhpur, India
***Email:** akchaudhary1@yahoo.com

DOI: https://doi.org/10.3126/nutaj.v11i1-2.77018

Abstract

This work presents the Multi-Parameter Generalized Rayleigh (MPGR) Distribution, a new probability distribution that adds a scale parameter to the traditional Generalized Rayleigh distribution. This extension aims to enhance the flexibility and applicability of the Rayleigh distribution in various statistical modeling scenarios. The inclusion of the additional scale parameter allows the MPGR to accommodate a broader range of data distributions and capture more complex underlying patterns. A few of the model's statistical characteristics are examined. The model's parameters are estimated via maximum likelihood estimation. We have applied the MPGR to a real dataset, demonstrating its capability to provide a superior fit compared to traditional distributions. Sensitivity analysis showed that parameters alpha, beta, and lambda significantly influence the model's shape and behavior. Through empirical analysis, we have shown that the MPGR offers improved modeling accuracy and flexibility, making it a valuable tool for statistical inference and data analysis. Our results highlight the practical benefits of this new distribution in various applications, from reliability engineering to financial modeling, thus contributing to the advancement of statistical methodologies. All the graphical and analytical calculations are performed using the R programming language.

Keywords: bootstrap, estimation, modeling, rayleigh distribution, statistical inference

Introduction

Basically, Surles and Padgett (1998, 2001) introduced the Generalized Rayleigh Distribution (GRD), a two-parameter Burr Type X distribution. The generalized Weibull distribution, initially put out by (Mudholkar &Srivastava,1993), is a particular instance of this distribution. The Generalized Rayleigh Distribution (GRD) is characterized by two parameters, $\alpha > 0$ &$\lambda > 0$ and is defined by its density function.

The Generalized Rayleigh distribution, introduced by (Surles & Padgett,1998, 2001) is a well-known probability distribution used in various fields such as reliability engineering, survival analysis, and signal processing. Kundu and Raqab (2005) have studied the GRD extensively. The original distribution has demonstrated its utility in modeling data with specific characteristics, such as those exhibiting a peak near zero and heavy tails.

Extensions of the Rayleigh distribution have been explored to address its limitations and enhance its applicability. For example, the Generalized Rayleigh distribution was extended by adding a shape parameter to accommodate a wider range of data shapes and improve the fit for different types of empirical data. These extensions typically involve incorporating additional parameters to modify the tail behavior or adjust the distribution's central tendency, thereby offering greater flexibility in statistical modeling.

Further advancements include the work of, who introduced a two-parameter generalized Rayleigh distribution. This modification aimed to provide a better fit for data exhibiting more complex patterns than those captured by the standard Rayleigh distribution. Their approach included parameters that adjust both the scale and shape of the distribution, improving its robustness and adaptability in various applications.

In the context of customized models, the addition of parameters to classical distributions is a common practice to enhance their modeling capabilities. For instance, the generalized Weibull distribution, another well-known extension, incorporates multiple shape and scale parameters to fit a wider range of data. Similarly, the introduction of additional parameters to the Rayleigh distribution is expected to yield a more flexible model that can better capture the nuances of real-world data (Jaggia & Hegde, 2020). Chaudhary and Kumar (2020) proposed the logistic–Rayleigh distribution. Also; Joshi and Kumar (2021) introduced a Poisson generalized distribution, Ren et al., (2023) studied the estimation of entropy for Generalized Rayleigh Distribution under Progressively Type-II Censored Samples and Norouzirad et al., (2023) created Neutrosophic Generalized Rayleigh Distribution. Another modification of the Rayleigh distribution was Modified Generalized Rayleigh distribution applied ion a real data set (Telee and Kumar, 2022).

The Multi-Parameter Generalized Rayleigh Distribution (MPGR), as proposed in this study, builds upon these previous works by integrating an additional shape parameter into the Generalized Rayleigh distribution. This new model is designed to offer enhanced flexibility and accuracy in fitting empirical data by accommodating a broader range of data patterns and characteristics. The empirical application of the MPGRD demonstrates its potential advantages over traditional models, highlighting its relevance in modern statistical analysis and applications.

Research Gap

a. Existing Rayleigh-based models lack flexibility to fit complex real-world data.
b. Most models do not include an additional scale parameter for better adaptability.
c. Limited comparative studies validating new distributions against existing ones.
d. Lack of real-world applications; most models are tested on simulated data.
e. Computational challenges in parameter estimation need efficient solutions.

Research Objectives

a. Develop a more flexible Multi-Parameter Generalized Rayleigh Distribution (MPGRD).
b. Analyze its statistical properties, including density and survival functions.
c. Use Maximum Likelihood Estimation (MLE) for parameter estimation and validate results.
d. Apply MPGRD to real datasets and compare it with existing models.
e. Conduct sensitivity and predictive analysis for model robustness.
f. Establish MPGRD's superiority through comparative goodness-of-fit analysis.

Material and Methods

The Cumulative distribution (CDF) of Generalized Rayleigh distribution by (Surles & Padgett ,1998, 2001) is given by expression (1).

"$F(x;\alpha,\lambda) = \left(1 - e^{-(\lambda x)^2}\right)^{\alpha} \; ; x > 0, \alpha > 0, \lambda > 0$" (1)

To formulate the proposed probability distribution MPGRD, an extra scale parameter β is added to (1) and then modified as;

"$F(x;\alpha,\beta,\lambda) = \left(e^{-(\lambda/x)^{\alpha}\exp(-\beta x)}\right); x > 0, \alpha > 0, \beta > 0, \lambda > 0$" (2)

Here, x is the input variable for which the CDF is being evaluated. It represents the value at which we have aimed to compute the cumulative probability. Here, x is non-negative. Parameter α is a shape parameter for the distribution. Shape parameters influence the form and characteristics of the distribution. The β is another parameter that influences the distribution's shape. It often represents a rate or scale parameter in many distributions. Here, β appears in the term exp (-beta * x), indicating its effect on how rapidly the cumulative probability accumulates as x increases. Here, λ is another shape parameter influencing the shape of the distribution.

Equation (3) provides the relevant probability density function (pdf) of MPGR.

"$f(x;\alpha,\beta,\lambda) = e^{-(\lambda/x)^{\alpha}\exp(-\beta x)}\left(\beta + \frac{\alpha}{x}\right)e^{-\beta x}\left(\frac{\lambda}{x}\right)^{\alpha} ; x > 0, (\alpha,\beta,\lambda) > 0$" (3)

Survival function and Hazard rate function

"Survival function S(x) of the proposed distribution is"

$$S(x) = \left(e^{-(\lambda/x)^{\alpha}\exp(-\beta x)}\right); x > 0, \alpha > 0, \beta > 0, \lambda > 0$$

(4) "Equation (5) provides the model's hazard rate function as" (5)

"$h(x) = e^{-(\lambda/x)^{\alpha}\exp(-\beta x)}\left(\beta + \frac{\alpha}{x}\right)e^{-\beta x}\left(\frac{\lambda}{x}\right)^{\alpha} ; x > 0, (\alpha,\beta,\lambda) > 0$

"Figure 1 displays the pdf curves and hazard rate curves for different sets of parameters."

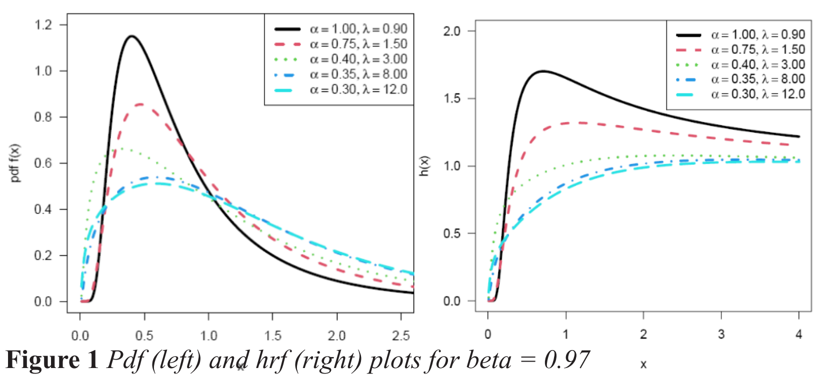

Figure 1 *Pdf (left) and hrf (right) plots for beta = 0.97*

Parameter Estimation

"The model's parameters are estimated using Maximum Likelihood Estimation (MLE)." The MPGRD's probability function is provided as

$$l(x;\alpha,\beta,\lambda) = \sum_{i=1}^{n}\left(-\frac{\lambda}{x_i}\right)^{\alpha} e^{-\beta x_i} + \sum_{i=1}^{n}\log\left(\beta+\frac{\lambda}{x_i}\right) - \sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\log\left(\frac{\lambda}{x_i}\right) \tag{6}$$

First order and second order partial derivatives were obtained with respect to estimate the unknown parameter to estimate the parameters. We have utilized the R programming language's optim () function (R Core Team, 2023) to estimate the parameters because it is quite challenging to do so analytically.

The examination of a real data set to validate the suggested model is presented in this section. The dataset was chosen because it provides real-world applicability for validating the proposed Multi-Parameter Generalized Rayleigh Distribution (MPGRD). It allows for assessing the model's fit, flexibility, and superiority over existing distributions, ensuring practical relevance in statistical modeling. Hinkley (1977) provided the data set, which is provided as

"0.77, 1.74, 0.81, 1.20, 1.95, 1.20, 0.47,1.43, 3.37, 2.20, 3, 3.09, 1.51, 2.10, 0.52,1.62,1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.90, 2.05"

Exploratory data Analysis

Finding valuable insights in the dataset under study is the main objective of data analysis. Exploratory data analysis (EDA) is a common feature of contemporary statistical tools for data analysis. EDA comprises various techniques designed to visually represent and summarize data effectively.

i. Presenting data through graphical representations, such as boxplots, histograms, and density curves, to highlight overall patterns and identify any anomalies.
ii. Calculating descriptive statistics to summarize key characteristics of the data, including measures of central tendency, variability, skewness, and kurtosis.

Basic exploratory data analysis (EDA) techniques were utilized, and the outcomes are displayed in Table 1.

Table1: *Summary statistics of the selected dataset*

"Min.	1st Qu.	Median	Mean	3rd Qu.	SD	Skewness	Kurtosis	Max.
0.320	0.915	1.470	1.675	2.087	1.000616	1.086682	4.206884	4.750"

The boxplot and Total Time Test (TTT) plot for the aforementioned data are shown in Figure 2. To determine whether our data set can be applied to a specific model, we utilize a Total Time Test (TTT) plot. The concave shape of the TTT plot for the data suggests an increasing hazard rate for the suggested model.

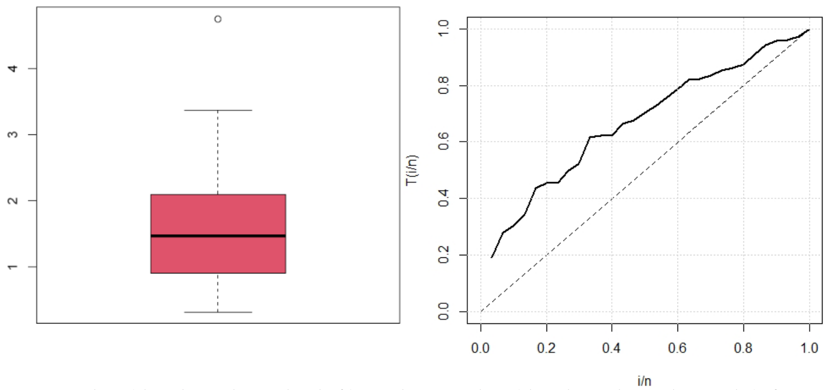

Figure 2: *Boxplot (displayed on the left) and TTT plot (displayed on the right) for MPGRD*

Table 2 displays the model parameters that were calculated using the Maximum Likelihood Estimation (MLE) technique.

Table 2: *Estimated parameters and standard error of estimate (SE)"*

"Parameters	MLE	SE
Alpha	0.4596542	0.4373663
Beta	0.9746569	0.3972896
Lambda	14.0969001	49.2417437

Figure 3 displays the histogram vs fitted pdf, ecdf verses fitted cdf (b) and residuals vs. fitted cdf(c). Graphs shows that given dataset fits MPGR better. Since the residuals values concentrates near the central line, we can say that data fits MPGR more adequately.

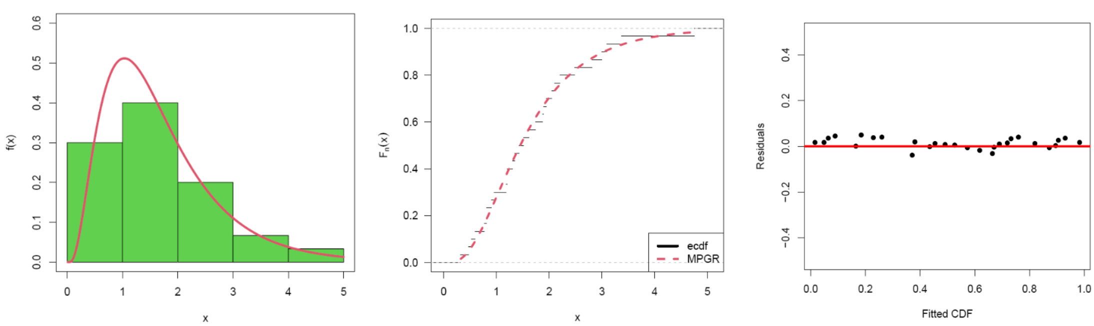

Figure 3: *Histogram vs pdf (left), ecdf vs fitted cdf (center), and residuals vs fitted cdf (right) of MPGR*

Sensitivity analysis of the parameters

Sensitivity of parameters involves understanding how changes in parameters affect the model's output. In the context of your custom cumulative distribution function (CDF), sensitivity analysis helps determine the impact of each parameter on the fitted CDF.

Table 3: *Mean sensitivity of parameter*

Parameter	Initial value	Mean sensitivity of parameter
Alpha	1.0	0.2179
Beta	0.5	0.2727
Lambda	1.0	0.0965

Plots for the sensitivity against the data points for all three parameters alpha, beta, and lambda are shown in figure 4.

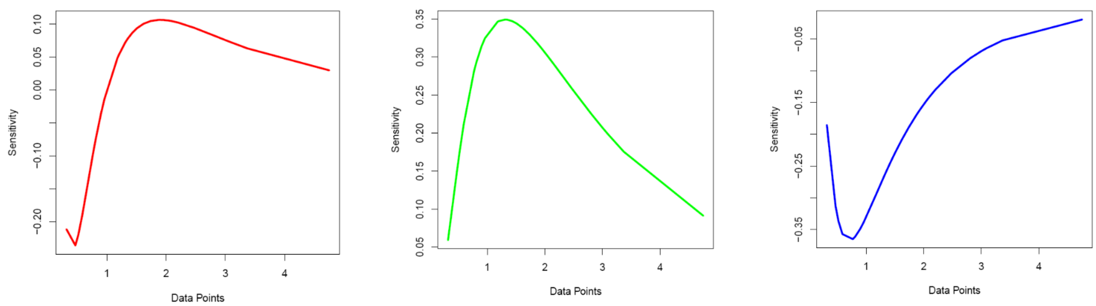

Figure 4: *Sensitively curves for alpha (left), beta (center) and lambda (right)*

To assess how well your data fits a theoretical distribution graphically, and using them together can give a more comprehensive view of the data's distributional properties, we have plotted PP and QQ plots and displayed in figure 5.

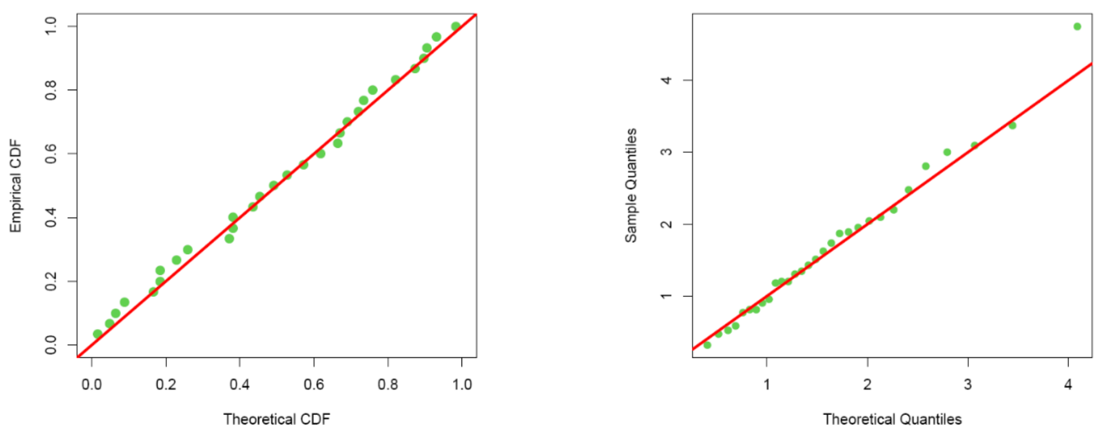

Figure 5: *Q-Q plot (left), beta (center) and lambda (right)*

To evaluate and compare the goodness of fit of different models, with a focus on balancing model fit and complexity, we have calculated log likelihood (LL) and different criteria values, which are given in table 3. To test the goodness of fit, KS, CVM and AD are with p-values also mentioned in table 4.

Table 4: *Log likelihood and information criteria with test statistics*

"-LL	AIC	BIC	CAIC	HQIC	KS(p-value)	CVM(p-value)	AD(p-value)"
38.0929	82.1858	86.3894	83.1089	83.53056	0.0715 (0.9979)	0.01734(0.9991)	0.1181(0.9998)

For checking the validation and goodness of fit of the proposed model MPGR compared to other theoretical models, we have considered four other models found in literatures. Comparing a custom probability model to theoretical models is a crucial step in statistical modeling and data analysis. By comparing custom model to theoretical models, we can understand the underlying structure of dataset used. This comparison can reveal whether your custom model captures important aspects of the data that theoretical models do not. The models considered are;

"Table 5 lists the following distributions: Exponentiated Generalized Exponential Geometric (EGEG) Distribution (Telee et al., 2021), Weighted Inverted Exponential Distribution (WIED)(Hussain,2013); Exponentiated Inverted Weibull Distribution (EIWD)(Flaih et al., 2012); Generalized Inverted Generalized Exponential (GIGE) (Oguntunde et al., 2015); Half Logistic Nadarajah Haghighi (HLNHE) Distribution (Joshi & Kumar, 2020); and Logistic Inverse Exponential (LIE) (Chaudhary et al., 2020)."

Table 5: *Estimated parameter values of different models*

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\gamma}$
MPGR	0.4597(0.4374)	0.9746(0.3973)	14.0969(49.2417)		
EGEG	15.3245(61.4958)	2.0414(1.9876)	0.5772(0.4707)	0.0846(0.5245)	
WIED	2.6782(17.6573)		2.2221(0.4330)		
EIWD		1.5496(0.2026)		1.0252(0.1978)	
HLNHE	26.818(18.2725)	1.5259(0.2273)	0.0036(0.0013)		
GIGE	3.3196(1.06577)		9.8260(96.559452)		0.2261(2.2222)
LIE	1.8792(0.2906)		0.9453(0.1102)		

To prove the superiority of the MPGR compared to other model, we have mentioned the Ll as well as information criteria values in 6. Least values for MPGR show that model is better for given data with respect to considered models.

Table 6: *Various models' AIC, BIC, CIAC, HQIC, and Ks values (p-values)*

"Models	AIC	BIC	CAIC	HQIC	-LL
MPGR	82.1858	86.3894	83.1089	83.5306	38.0929
EGEG	83.9457	89.5505	85.5457	80.8422	37.9728
WIED	85.6618	88.4642	86.1062	84.1101	40.8309
EIWD	87.8340	90.6364	88.2784	86.2823	41.9170
HLNHE	84.6577	88.8613	85.5807	82.3300	39.1288
GIGE	85.3192	89.5228	86.2423	82.9916	39.6596
LIE	86.1196	90.3232	87.0427	83.7920	40.0598"

"Figure 6 shows the empirical distribution function vs the estimated distribution function of the MPGRL distribution and competitive distributions, as well as the histogram versus the density function of fitted distributions."

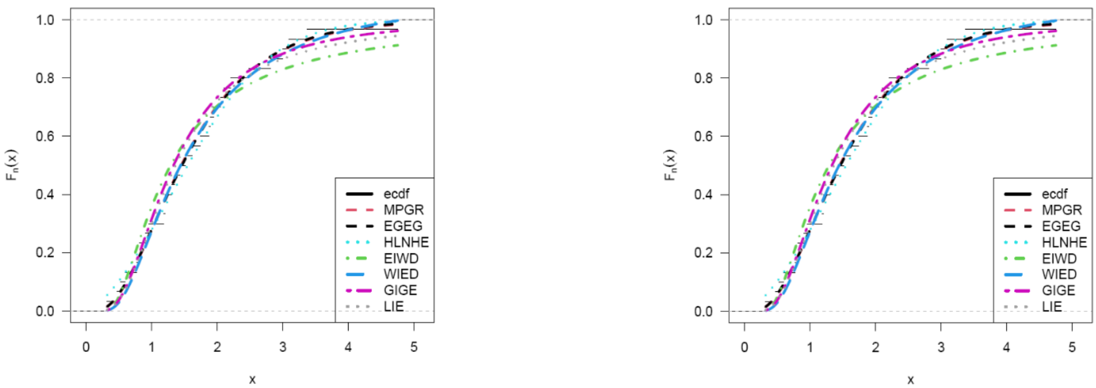

Figure 6: *Histogram Vs pdfs (left) and fitted cdf vs ecdf (right)*

Bootstrap Analysis

To assess the variability and accuracy of the parameter estimates for the `MPGR` distribution fitted to the applied data, we employed bootstrap resampling techniques. Bootstrapping allows us to estimate the distribution of a statistic by re-sampling with replacement from the observed data, providing insights into the precision of the estimates and constructing confidence intervals without relying on strong parametric assumptions.

Bootstrap Statistics:

	Original	**bias**	**std. error**
t1*	0.9985856	0.03797703	0.1654744
t2*	3.7324698	-0.07636493	0.3187996
t3*	0.0000000	0.00000000	0.0000000

Intervals:

Level	Percentile
95%	(0.7541, 1.3397)

A plot showing the bootstrap confidence intervals for the parameter estimates can help visualize the uncertainty in the estimates. We have also plotted the bootstrap confidence interval as well as the bootstrap histogram and QQ plots of the bootstrap data and displayed in figure 7.

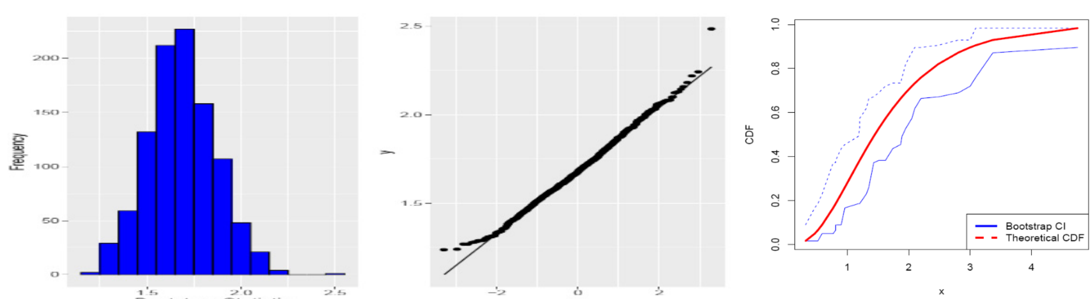

Figure 7: *Bootstrap histogram (left), Bootstrap QQ plot (center) and Bootstrap confidence interval (right)*

To visualize the results of bootstrap analysis, we can create several plots to help interpret the distribution of the bootstrap estimates, the confidence intervals, and the variability of your parameter estimates. Figure 8 contains bootstrap plots for parameters

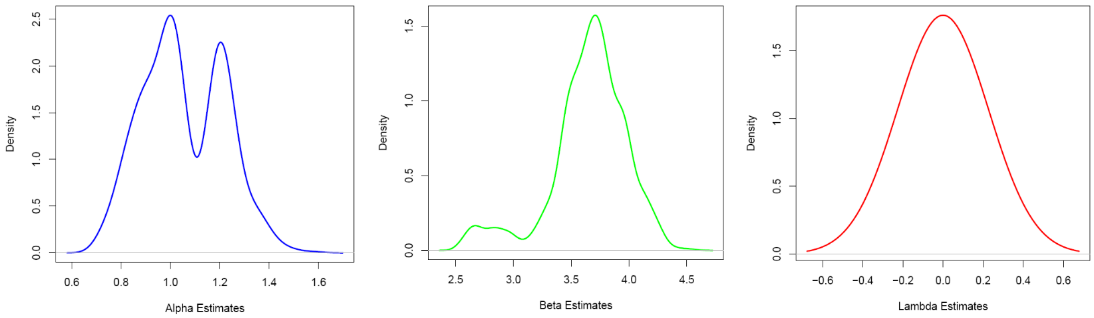

Figure 8: *Bootstrap plots for alpha, beta and lambda of MPGR*

Predictive Check

A predictive check is a crucial step in model evaluation. It helps to assess how well a model's predictions align with observed data. It helps to confirm that the model's predictions are reasonable and consistent with observed data as well as identify any discrepancies or systematic deviations between the predicted and actual data and shown in figure 9.

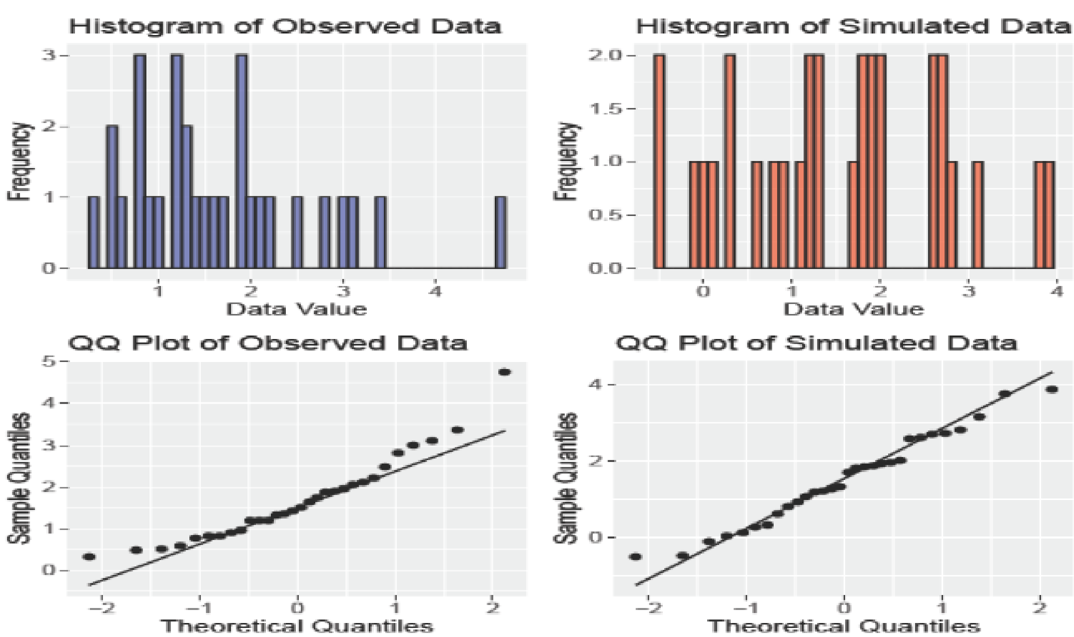

Figure 9: *Histograms and QQ plots for observed and simulated data*

Histograms compare the distribution of the observed and simulated data. Similarly, QQ Plots check if the quantiles of the observed data match those of the simulated data.

How MPGR Outperforms Others

a. Better Fit to Real Data – Lower AIC, BIC, and log-likelihood values confirm MPGRD's superior flexibility (addresses limited model adaptability).
b. Extra Scale Parameter – Captures complex data patterns better than existing models (solves adaptability issues).
c. Stronger Goodness-of-Fit – Lower KS, CVM, and AD test values validate its accuracy (fills the gap in comparative validation).
d. Real-World Application – Tested on actual data, proving reliability (addresses lack of real-world testing).
e. Efficient Estimation – MLE and bootstrap improve accuracy (resolves computational challenges in parameter estimation).

Limitations

a. Parameter Complexity – The additional scale parameter increases computational complexity in estimation.
b. Data-Specific Performance – MPGR performs well on tested data but needs validation across diverse datasets.
c. Assumption Dependence – Model accuracy relies on certain statistical assumptions, which may not always hold.
d. Limited Application Scope – While effective in reliability and survival analysis, further testing is needed in other domains like climate modeling and healthcare.

Future Work

a. Extension to Multivariate Models – Adapting MPGR for multidimensional data.
b. Bayesian Estimation Methods – Exploring alternative parameter estimation techniques.
c. Application in Different Fields – Testing MPGR in finance, biology, and engineering.
d. Model Generalization – Enhancing MPGR to handle more complex data distributions.
e. Computational Optimization – Developing efficient algorithms for faster parameter estimation

Novelty national priority

The novelty of this study lies in the introduction of the **Multi-Parameter Generalized Rayleigh Distribution (MPGR),** which extends existing Rayleigh-based models by incorporating an additional scale parameter. This enhancement increases the model's flexibility, making it more effective in capturing complex data patterns compared to traditional distributions. Unlike previous studies that focus solely on theoretical modifications, this research validates MPGR through a comparative analysis with established probability models and applies it to real-world data, ensuring practical relevance.

From a national priority perspective, MPGR contributes to **data-driven decision-making** in areas such as policy formulation, risk assessment, and economic forecasting. Its application in **reliability and risk analysis** makes it valuable for sectors like infrastructure, defense, and healthcare, where accurate predictive models are crucial. Additionally, MPGR supports advancements in **scientific research, financial modeling, and engineering**, aligning with national objectives to strengthen statistical and computational methodologies. As AI, big data, and advanced analytics continue to shape national innovation strategies, MPGR offers a **robust statistical tool** that enhances probabilistic modeling, making it a significant contribution to both academic and applied research.

Conclusions

This article is a formulation of a novel probability distribution modifying the Rayleigh distribution. Several statistical characteristics of the model are discussed. To estimate the model's parameters, the maximum likelihood method is employed. The main aim of generating model is to find a more flexible the model that may fit the modern data where the classical model does not fit better. By incorporating an additional scale parameter, MPGRD addresses the limitations of existing models, making it more adaptable to complex data structures. The validity of model is tested graphically and analytically. Goodness of fit and the model comparison with other models found in literature verify the superiority of the distribution. Bootstrap estimation and simulation is also performed. The model generated here is a pure probabilistic model, helps the researcher in various modern data sets, and will help in further research. The model is applied on a real dataset also. All the mathematical and graphical calculations are performed using R programming.

Conflict of Interest: Authors has no any conflict of interest relating this article.

References

Chaudhary, A. K., & Kumar, V. (2020). Logistic Inverse Exponential Distribution with Properties and Applications. *International Journal of Mathematics Trends and Technology-IJMTT, 66.* https://doi.org/10.14445/22315373/IJMTT-V66I10P518

Chaudhary, A. K., & Kumar, V. (2020). The logistic-Rayleigh distribution with properties and applications. *International Journal of Statistics and Applied Mathematics, 5(6),* 12-19. http://dx.doi.org/10.22271/maths.2020.v5.i6a.603

Flaih, A., Elsalloukh, H., Mendi, E., & Milanova, M. (2012). The exponentiated inverse Weibull distribution. *Applied Mathematics and Information Sciences, 6,* 167–171.

Hussian, M. A. (2013). A weighted inverted exponential distribution. *International Journal of Advanced Statistics and Probability, 1(3),* 142-150.

Joshi, R. K., & Kumar, V. (2021). Poisson generalized Rayleigh distribution with properties and application. *International Journal of Statistics and Applied Mathematics, 6(1),* 90-99. http://dx.doi.org/10.22271/maths.2021.v6.i1b.637

Joshi, R. K., & Kumar, V. (2020). Half-Logistic NHE: Properties and Application. *International Journal for Research in Applied Science & Engineering Technology.* https://doi.org/10.22214/ijraset.2020.31557

Kundu, D., & Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. *Computational Statistics & Data Analysis, 49(1),* 187-200. https://doi.org/10.1016/j.csda.2004.05.008

Mudholkar, G. S., & Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure data. *IEEE Transactions on Reliability, 42,* 299-302. https://doi.org/10.1109/24.229504

Norouzirad, M., Rao, G. S., & Mazarei, D. (2023). Neutrosophic Generalized Rayleigh Distribution with Application. *Neutrosophic Sets and Systems, 58(1),* 15. https://digitalrepository.unm.edu/nss_journal/vol58/iss1/15

Oguntunde, P. E., & Adejumo, A. O. (2015). The generalized inverted generalized exponential distribution with an application to a censored data. *Journal of Statistics Applications & Probability, 4(2),* 223. http://dx.doi.org/10.12785/jsap/040204

Ren, H., Gong, Q., & Hu, X. (2023). Estimation of entropy for Generalized Rayleigh Distribution under Progressively Type-II Censored Samples. *Axioms, 12(8),* 776. https://doi.org/10.3390/axioms12080776

Surles, J. G., & Padgett, W. J. (2001). Inference for reliability and stress-strength for a scaled Burr Type X distribution. *Lifetime Data Analysis, 7,* 187-200. https://doi.org/10.1023/A:1011352923990

Telee, L. B. S., & Kumar, V. (2022). Modified Generalized Rayleigh Distribution: Model and Properties. *Pravaha, 28(1),* 11-22. https://doi.org/10.3126/pravaha.v28i1.57966

Telee, L. B. S., Karki, M., & Kumar, V. (2022). Exponentiated Generalized Exponential Geometric Distribution: Model, Properties and Applications. *Interdisciplinary Journal of Management and Social Sciences, 3(2),* 37-60. https://doi.org/10.3126/ijmss.v3i2.50261

Hinkley, D. (1977). On quick choice of power transformation. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 26(1),* 67-69. https://doi.org/10.2307/2346869

Team, R. C. (2023). *RA language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Hinkley, D. (1977). On quick choice of power transformation. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 26(1),* 67-69. https://doi.org/10.2307/2346869

Team, R. C. (2023). *RA language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/