THE INVERSE LOMAX DISTRIBUTION WITH APPLICATIONS

Jamilu Yunusa Falgore[1], Adamu Abubakar Umar[1,2], Aliyu Yakubu[1], and Aliyu Ismail Ishaq[1]

[1]Department of Statistics, Ahmadu Bello University, Zaria-Nigeria

[2]Department of Mathematics, University of Manchester, Manchester M13 9PL, UK

Corresponding author: jamiluyf@gmail.com

ABSTRACT

The Inverse Lomax Distribution, a flexible heavy-tailed distribution, has gained attention in statistical modeling, particularly in fields like finance, reliability engineering, and risk analysis. This research delves into the applications of the Inverse Lomax Distribution, offering insights into its robustness in various contexts. The parameters of the ILD were estimated using maximum likelihood method. Key advantages include its ability to model data with high skewness and heavy tails, making it a strong candidate for applications requiring robust outlier sensitivity. Additionally, the distribution's straightforward form enables analytical tractability, facilitating parameter estimation and inference.

Keywords: Inverse Lomax Distribution; Inverse Lomax Log-Logistic Distribution; Skewed Distribution.

1. INTRODUCTION

The Lomax distribution, introduced by Lomax (1954), has been widely used in various fields such as economics, engineering, and survival analysis due to its heavy-tailed characteristics and flexibility in modeling data with decreasing failure rates. Its inverse counterpart, the Inverse Lomax Distribution (ILD), has garnered attention for its applicability in modeling lifetimes and reliability data where the hazard function exhibits specific behaviors. Inverse Lomax Distribution is part of an inverted family of distributions, and it can be used in a variety of situations where the failure rate is non-monotonic (Singh et al. 2013). Inverse Lomax distribution is an alternative to a lot of distributions like Gamma, Weibull, e.t.c. (Sharma and Kumar, 2020).

Recent studies have focused on extending and generalizing the ILD to enhance its flexibility and applicability. For instance, Hassan and Abd-Allah (2019) proposed the Inverse Power Lomax Distribution, introducing an additional shape parameter to better model various data behaviors. Similarly, Almarashi (2021) developed a modified version of the ILD, termed the Modified Logarithmic Transformed Inverse Lomax Distribution, which demonstrated improved fitting for engineering and medical data. Furthermore, Muhammed et al. (2022) introduced the Inverse Lomax Chen Distribution, combining properties of the ILD and the Chen distribution to model data with different hazard rate shapes.

Despite these advancements, there remains a gap in the literature concerning the comprehensive exploration of the ILD's potential applications across diverse fields. Specifically, while various generalizations have been proposed, a unified framework that systematically evaluates the ILD's applicability and performance in real-world scenarios is lacking.

The analysis of extreme events, particularly those related to natural catastrophes, has garnered significant attention in both statistical and actuarial sciences. Understanding the distribution and characteristics of such events is crucial for risk assessment, financial modeling, and disaster preparedness. This study focuses on a unique dataset developed by Hogg and Klugman (2009), which captures 40 losses incurred in 1977 due to wind-related catastrophes. The dataset is characterized by its extreme right skewness, a common feature in loss distributions, where a majority of the observations cluster at lower values, while a few extreme outliers significantly influence the overall distribution. Extreme Value Theory (EVT) is a statistical framework specifically designed to model the tail behavior of distributions, making it highly relevant for analyzing datasets with extreme skewness and outliers. The dataset in question exhibits a pronounced right skew, with a few extreme values (e.g., 32 and 43) that dominate the distribution. EVT has been widely applied in actuarial science to model catastrophic losses, as discussed by Embrechts et al. (2013) in their seminal work, *Modelling Extremal Events for Insurance and Finance*. They emphasize the importance of understanding tail behavior for accurate risk assessment and pricing of insurance products. The presence of extreme skewness in loss data poses significant challenges for traditional statistical methods, which often assume normality or symmetry. McNeil et al. (2005), in *Quantitative Risk Management: Concepts, Techniques, and Tools*, highlight the limitations of conventional models, such as the Gaussian distribution, in capturing the tail behavior of skewed datasets. They advocate for the use of heavy-tailed distributions, such as the Pareto or Generalized Extreme Value (GEV) distributions, which are better suited for modeling extreme events. Wind-related catastrophes, such as hurricanes and tornadoes, have been the subject of numerous studies due to their significant economic impact. Smith (2003), in *Statistics of Extremes: With Applications in Environment, Insurance, and Finance*, provides a comprehensive analysis of wind-related loss data, highlighting the importance of extreme value analysis in understanding the frequency and severity of such events.

The dataset presented by Chhikara and Folks (1977) provides a valuable resource for studying the active repair times (in hours) of an Airborne Communication Transceiver. This dataset has been widely cited in statistical literature, particularly in the context of reliability analysis, survival analysis, and the application of lifetime distributions. Chhikara and Folks (1977) originally used this dataset to demonstrate the applicability of the inverse Gaussian distribution in modeling repair times. The inverse Gaussian distribution is particularly suited for data with positive skewness and is often used in reliability engineering to model lifetimes and failure times. Their work highlighted the flexibility of the inverse Gaussian distribution in capturing the variability and skewness inherent in repair time data. The dataset has been extensively used in reliability studies to estimate the mean time to repair (MTTR) and to analyze the failure rates of systems. For instance, Padgett and Spurrier (1985) utilized this dataset to compare the performance of nonparametric and parametric methods in reliability analysis. They emphasized the importance of choosing appropriate models for skewed data to ensure accurate predictions of system reliability. Several studies have employed this dataset to evaluate the performance of goodness-of-fit tests for skewed distributions. For example, D'Agostino (2017) used this dataset to compare the power of various statistical tests, such as the Kolmogorov-Smirnov and Anderson-Darling tests, in detecting deviations from the inverse Gaussian distribution. The presence of extreme values, such as 22.0 and 24.5 hours, has made this dataset a subject of interest in extreme value theory. Researchers have used it to study the tail behavior of repair times and to develop robust models for predicting rare but critical events in system maintenance.

Aim and Objectives

This paper aims to bridge a gap by providing a thorough investigation of the Inverse Lomax Distribution and its applications. The specific objectives are:

i. To evaluate the performance of the ILD and its variants in modeling real-world data from different domains.

ii. To identify potential areas where the ILD can be effectively applied and suggest directions for future research.

2 Methodology

2.1 The Shapes of the ILD

The probability density function (PDF), cumulative distribution function (CDF), and quantile function of Inverse Lomax distribution are given in equations (1), (2), and (3) as:

$$f(x;\lambda,\gamma) = \lambda\gamma x^{-2}\left(1+\frac{\lambda}{x}\right)^{-1-\gamma};x,\lambda,\gamma > 0, \tag{1}$$

$$F(x;\lambda,\gamma) = \left(1+\frac{\lambda}{x}\right)^{-\gamma};x,\lambda,\gamma > 0, \tag{2}$$

and

$$x_q = \frac{\lambda}{\left(u^{\frac{-1}{\gamma}}-1\right)}, \tag{3}$$

where λ and γ are the scale and shape parameters, and u is uniformly distributed between zero and 1. The PDF plot of ILD is given by:

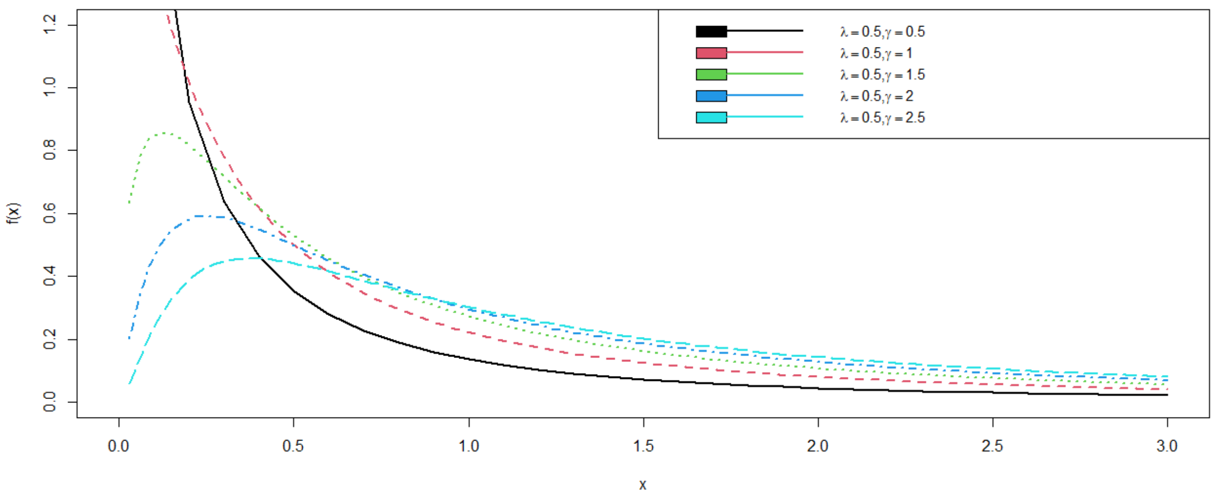

Figure 1: The pdf plot of ILD at various parameter values

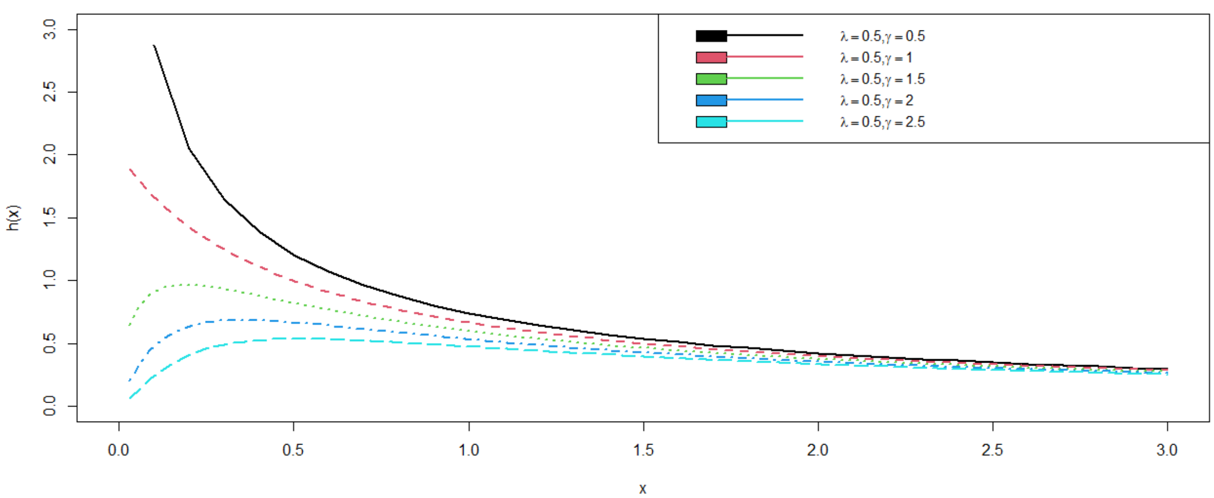

Figure 2: The various hazard shapes of the ILD

From figure 1, each combination of λ and γ would produce a different curve on the plot. As γ increases, the shape of the distribution might change, potentially becoming more skewed or peaked. Also form figure 2, the hazard function can be increasing and decreasing based on the values of the parameters.

2.2 The Maximum Likelihood Estimates of the ILD

The log-likelihood function oof the ILD based on equation (1) is :

$$l(f) = n\log(\lambda\gamma) - 2\sum_{i=1}^{n}\log(x_i) - (1+\gamma)\sum_{i=1}^{n}\log\left(1+\frac{\lambda}{x}\right) \tag{4}$$

The partial derivative of equation (4) with respect to λ is given as:

$$\frac{\partial l(f)}{\lambda} = \frac{n}{\lambda} - \frac{(1+\gamma)}{x\left(1+\frac{\lambda}{x}\right)} \tag{5}$$

The partial derivative of equation (4) with respect to γ is given as:

$$\frac{\partial l(f)}{\gamma} = \frac{n}{\gamma} - \sum_{i=1}^{n} \log\left(1+\frac{\lambda}{x}\right) \tag{6}$$

Equations (5) and (6) can not be solve directly because the parameters are non-linear in parameters. Numerical methods like Newton-Raphson approaches can be used to solve the equations.

3 Simulation and Applications

3.1 The Simulation Studies

Here, a simulation studies was conducted to evaluate the properties of the estimates and reported in Table 1. Equation (3) was used to draw samples from Inverse Lomax Distribution. Also, 1,000 replications were considered. Sample sizes of 50, 100, 200, and 500 were considered. Estimates of the parameters, bias, and root mean squared errors (RMSE) were presented. The values of the parameters (guess) were 0.7 and 0.5 for λ and γ.

Table 1:The Simulation Results

Sample Size	Metrics	$\hat{\lambda}$	$\hat{\gamma}$
50	Means	0.7196	0.5434
	Bias	0.0196	0.0434
	RMSE	0.3769	0.1555
100	Means	0.7037	0.5208
	Bias	0.0037	0.0208
	RMSE	0.2395	0.0862
200	Means	0.7058	0.5088
	Bias	0.0058	0.0088
	RMSE	0.1751	0.0555
500	Means	0.7051	0.5027
	Bias	0.0051	0.0027
	RMSE	0.1092	0.0339

Table 1 suggest that the estimates approach the true parameters as the sample sizes increases. The biases and RMSE approaches zero as the sample size increases.

3.2 The Applications

Goodness-of-fit Statistics used in this study are: Kolmogorov-Smirnov (KS), Cramer Von-Mises, and Anderson Darling statistics. While the Goodness-of-fit Criteria are the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

The first dataset: The following extreme right skewed dataset, developed by Hogg and Klugman (2009), consists of 40 losses that occurred in 1977 due to wind-related catastrophes, and the observations are: 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 6, 8, 8, 9, 15, 17, 22, 23, 24, 24, 25, 27, 3243.

Table2: Goodness-of-Fit Statistics

Statistic	ILD	Weibull	Gamma	Lnorm	Exp
KS	0.17	0.34	0.43	0.21	0.73
CvM	0.23	0.86	1.85	0.32	7.79
AD	1.57	4.97	9.04	2.19	52.63

Table 3: Goodness-of-Fit Criteria

Criterion	ILD	Weibull	Gamma	Lnorm	Exp
AIC	246.74	297.64	328.03	266.28	413.28
BIC	249.96	300.87	331.25	269.50	414.89
-LL	121.37	146.82	162.01	131.14	205.64
Mles	4.62 (0.01), 8.21 (0.04)	0.46 (0.04), 13.89 (5.35)	0.26 (0.05), 0.01 (0.01)	1.81 (0.22), 1.38 (0.16)	0.01 (0.01)

Table 2 presents the Goodness-of-fit Statistics for the first dataset (wind-related catastrophes). ILD is the best candidate with lower values of all the statistics. Moreover, Table 3 presents the Goodness-of-fit Criteria for first dataset. The table suggest that ILD is the best with minimum values of all the criteria.

The second dataset was presented by Chhikara and Folks (1977). It presents the active repair times (in hours) for an Airborne Communication Transceiver. The dataset is: 0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0,3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3,22.0, 24.5.

Table 4: Goodness-of-Fit Statistics

Statistic	ILD	Weibull	Gamma	Exp
KS	0.07	0.12	0.15	0.16
CvM	0.04	0.12	1.18	0.21
AD	0.32	0.89	1.10	1.26

Table 5: Goodness-of-Fit Criteria

Criterion	ILD	Weibull	Gamma	Exp
AIC	205.23	212.94	213.86	212.01
BIC	208.88	216.59	217.52	213.84
-LL	100.61	104.47	104.93	105.01
Mles	0.07 (0.19), 16.74 (41.91)	0.89 (0.09), 3.39 (0.59)	0.93 (0.17), 0.26 (0.06)	0.28 (0.04)

Table 4 presents the Goodness-of-fit Statistics for the second dataset (active repair times). ILD is the best candidate with lower values of all the statistics. Moreover, Table 5 presents the Goodness-of-fit Criteria for second dataset. The table suggest that ILD is the best with minimum values of all the criteria.

4 CONCLUSIONS

In this study, ILD was considered and applied in two different areas. The shapes of the PDF of ILD can be decreasing and skewed to the right. Moreover, the hazard function can be increasing as well as decreasing depending the values of the parameters. These, can be achieved by fixing the scale parameter and then varies the shape parameter. Upon application to an extreme events dataset, the ILD proved its importance over the Exponential, Weibull, Gamma, as well as Log-Normal distributions. As area of further studies, other estimation procedure can be considered. Moreover, different datasets in some areas like medicines, hydrology, and social sciences can be explored.

5 REFERENCES

Almarashi, A. M. (2021). A new modified inverse Lomax distribution: Properties, estimation and applications to engineering and medical data. *Computer Modeling in Engineering & Sciences*, *127*(2), 621-643.

Chhikara, R. S., & Folks, J. L. (1977). The inverse Gaussian distribution as a lifetime model. *Technometrics*, *19*(4), 461-468.

D'Agostino, R. B. (2017). *Goodness-of-fit-techniques*. Routledge.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (2013). *Modelling extremal events: for insurance and finance* (Vol. 33). Springer Science & Business Media.

Hassan, A. S., & Abd-Allah, M. (2019). On the inverse power Lomax distribution. *Annals of Data Science*, *6*, 259-278.

Hogg, R. V., & Klugman, S. A. (2009). *Loss distributions*. John Wiley & Sons.

Lomax, K. S. (1954). The Sources and Nature of Statistical Information in Special Fields of Statistics Regional Economic Statistics. *Journal of the Royal Statistical Society. Series A (General)*, *117*(1), 85-101.

McNeil, A. J. (2005). Quantitative Risk Management Concepts, Techniques and Tools.

Muhammed, S., Dahiru, T., & Yahaya, A. (2022). The Inverse Lomax Chen Distribution: Properties and Applications. *Communication in Physical Sciences*, *8*(3).

Padgett, W. J., & Spurrier, J. D. (1985). On discrete failure models. *IEEE Transactions on Reliability*, *34*(3), 253-256.

Sharma, A., & Kumar, P. (2020). Estimation of Density and Reliability Function of Inverse Lomax Distribution under Type-II Censoring and the Sampling Scheme of Bartholomew.

Singh, S. K., Singh, U., & Kumar, D. (2013). Bayes estimators of the reliability function and parameter of inverted exponential distribution using informative and non-informative priors. *Journal of Statistical computation and simulation*, *83*(12), 2258-2269.

Smith, R. L. (2003). Statistics of extremes, with applications in environment, insurance, and finance. *Extreme values in finance, telecommunications, and the environment*, 20-97.