Modified Logistic Lomax Distribution: Model, Properties, Simulation, and Application to real dataset

Lal Babu Sah Telee
Department of Management Science, Nepal Commerce Campus, Tribhuvan University
Email: lalbabu3131@gmail.com
ORCID: https://orcid.org/0009-0009-0994-6445
Kapil Shah
Department of Management Science, Nepal Commerce Campus, Tribhuvan University
Email: kapil.shah@ncc.edu.np
ORCID: https://orcid.org/0009-0003-8320-3312 (Corresponding Author)
Arun Kumar Chaudhary
Department of Statistics, Nepal Commerce Campus, Tribhuvan University
Email: akchaudhary1@yahoo.com
ORCID: https://orcid.org/0000-0003-4783-3256

https://doi.org/10.3126/pravaha.v31i1.89151

Abstract

We have introduced Modified Logistic Lomax (MLL) Distribution Modifying the Logistic Lomax model. Modification of the model makes it to give different shaped density curve and hazard rate curves enabling model more flexible and suitable for various type of the datasets. Study of various statistical properties and Simulation study is performed. The MLL model is more useful for real datasets where classical probability models do not analyze the dataset appropriately. Maximum likelihood estimation (MLE), Least square estimation (LSE) and Cramer-von Mises methods (CVM) are used for parameter estimation and consistent results of the parameters are seen. Information criterion and graphical approach demonstrate the superiority of the model over some probability models. The MLL model may be helpful for the researcher to know the formulation of the new probability model and its application for real data analysis. All the computation analysis is performed using R programming.

Keywords: *Estimation, Model formulation, Simulation, Statistical properties.*

Cite this paper
Sah Telee, L. B., Shah, K., & Chaudhary, A. K. (2025). Modified Logistic Lomax Distribution: Model, Properties, Simulation, and Application to real dataset. *Pravaha, 31*(1), 52-61.

Introduction

Probability distribution is one of the basic statistical tools for the modern education and inference. In overall the field of life, probability distribution play crucial and decision level role. In literature, there are different significantly powerful probability distributions and has been helping in decision making, forecasting and simulation etc. In modern age, there are many data sets that cannot be fully captured by the classical models and cannot give the accurate analysis. To fulfill this gap, researchers have been formulating different new probability models that capture datasets more precisely. There are various methods of designing the classical model. Modification of the classical model by altering the position of the variables and parameter is one of the methods of getting new probability model. In some cases, we add some extra parameters to the original models to get more flexible model. In literature we can find various modification of the logistic distribution. Windarto et al. (2018) introduced a new modified logistic model by modification of classical logistic differential equation. The Richard model (Richard,

1959) is modification of the logistic growth model. Telee et al. (2023) inversed the variable to get the Modified Inverse Lomax distribution which is very flexible in real data analysis. Lemonte & Cordeiro (2013) modified Lomax to design extended Lomax distribution. Chaudhary et al. (2022) modified Lomax distribution to generate the Inverse exponentiated Odd Lomax Exponential distribution. This study formulates a new probability model by Modifying the Logistic Lomax distribution (Chaudhary & Kumar, 2020) using Lomax distribution as baseline distribution.

The cdf of the Logistic Lomax distribution is given by equation (1)

$$F(x,\alpha,\beta,\lambda)=1-\frac{1}{1+\left(\left(1+\beta x\right)^{\lambda}-1\right)^{\alpha}},x>0,(\alpha,\beta,\lambda)>0 \qquad (1)$$

We have modified this model taking βx = eβx to introduce modified Logistic Lomax distribution. The cdf and pdf of the MLL are given by equation (2) and equation (3)

$$F(x,\alpha,\beta,\lambda)=1-\frac{1}{1+\left[\left(1+\beta e^{\lambda x}\right)-1\right]^{\alpha}};x>0,(\alpha,\beta,\lambda)>0 \qquad (2)$$

$$f(x,\alpha,\beta,\lambda)=2\alpha\lambda\beta e^{\lambda x}\left(1+\beta e^{\lambda x}\right)\left(\left(1+\beta e^{\lambda x}\right)^{2}-1\right)^{\alpha-1}\left(1+\left(\left(1+\beta e^{\lambda x}\right)^{2}-1\right)^{\alpha}\right)^{-2},x>0 \qquad (3)$$

Survival function of the model which gives the probability of the survival beyond the time x is expressed in by equation (4)

$$S(x,\alpha,\beta,\lambda)=\left(1+\left(\left(1+\beta e^{\lambda x}\right)^{2}-1\right)^{\alpha}\right)^{-1},x>0,(\alpha,\beta,\lambda)>0 \qquad (4)$$

Hazard rate which is the instantaneous failure rate of the MLL distribution is expressed by equation (5)

$$h(x)=2\alpha\lambda\beta e^{\lambda x}\left(1+\beta e^{\lambda x}\right)\left(\left(1+\beta e^{\lambda x}\right)^{2}-1\right)^{\alpha-1}\left(1+\left(\left(1+\beta e^{\lambda x}\right)^{2}-1\right)^{\alpha}\right)^{-1},x>0 \qquad (5)$$

The density plots and hazard rate plots for different sets of parameters are displayed in figure 1. Figure on left panel demonstrate different shaped and different natured density plot. Similarly, in right panel of Figure 1, is the hazard rate plots for five sets of parameters. These variation in shape of density and hazard rate curves confirms the flexibility of MLL for capturing different natured datasets.

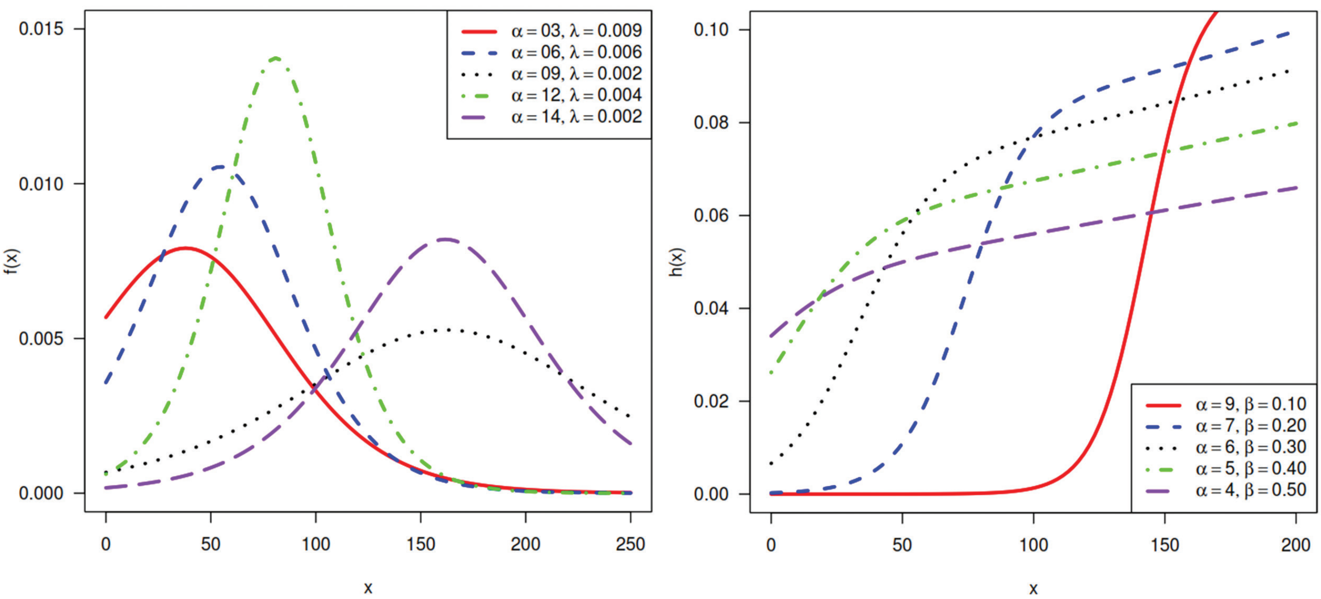

Figure 1: Density plots for β = 0.3 (left) and hazard rate plots for λ = 0.05 (right)

Quantile function

Quantile function generates quantile values which are essential for obtaining the theoretical quantiles. Theoretical quantile is used here to plot the Q-Q plot which helps to test the model validation graphically.

$$Q(p,\alpha,\beta,\lambda)=\frac{1}{\lambda}\ln\left[\frac{1}{\beta}\left(\sqrt{1+\left(\frac{p}{1-p}\right)^{1/\alpha}}-1\right)\right];0<p<1$$

Random Number of Generation

Random Number of Generation which helps to generate data sets following a particular model of the MLL model is given by equation (6)

$$X=Q(U,\alpha,\beta,\lambda)=\frac{1}{\lambda}\ln\left[\frac{1}{\beta}\left(\sqrt{1+\left(\frac{U}{1-U}\right)^{1/\alpha}}-1\right)\right];0<U<1 \qquad (6)$$

Simulation Study

The simulation study demonstrates the performance of maximum likelihood estimation for the proposed distribution. Simulation study for α =1, β =1, and λ = 1 demonstrate the gradual decrease in bias and MSE. As sample size increases from 50, 70, 90, …, 990, bias and MSE tends to zero indicating better fitting of the model.

Table 1: Simulation study for α =1, β =1, and λ = 1

n	α	α_bias	α_MSE	β	β_Bias	λ	λ_bias	λ_MSE
50	3.07	2.07	49.58	4.7	3.7	1.56	0.56	2.25
70	3.67	2.67	58.91	3.03	2.03	1.38	0.38	1.96
90	2.55	1.55	30.34	3.59	2.59	1.51	0.51	2.05
110	1.86	0.86	13.39	2.46	1.46	1.35	0.35	1.28
130	2.05	1.05	20.15	1.82	0.82	1.29	0.29	1.17
150	2.03	1.03	13.91	1.72	0.72	1.21	0.21	0.93
170	1.86	0.86	12.54	1.85	0.85	1.15	0.15	0.75
190	1.47	0.47	5.60	1.51	0.51	1.18	0.18	0.57
210	1.36	0.36	1.86	1.3	0.3	1.11	0.11	0.46
230	1.71	0.71	7.74	1.28	0.28	1.08	0.08	0.44
250	1.46	0.46	6.77	1.84	0.84	1.21	0.21	0.78
270	1.47	0.47	2.93	1.16	0.16	1.04	0.04	0.29
290	1.37	0.37	2.59	1.24	0.24	1.11	0.11	0.34
310	1.32	0.32	2.51	1.11	0.11	1.04	0.04	0.24
330	1.24	0.24	2.3	1.23	0.23	1.15	0.15	0.35
350	1.22	0.22	0.91	1.36	0.36	1.08	0.08	0.37
370	1.10	0.1	0.31	1.12	0.12	1.08	0.08	0.18
390	1.10	0.1	0.32	1.16	0.16	1.09	0.09	0.22
410	1.25	0.25	0.94	1.06	0.06	0.99	-0.01	0.16
430	1.24	0.24	0.64	1.10	0.10	1.01	0.01	0.21
450	1.10	0.10	1.19	1.18	0.18	1.12	0.12	0.21
470	1.05	0.05	0.20	1.17	0.17	1.10	0.10	0.19
490	1.04	0.04	0.16	1.14	0.14	1.10	0.10	0.18
510	1.20	0.20	1.16	1.05	0.05	1.00	0.00	0.13
530	1.06	0.06	0.31	1.12	0.12	1.08	0.08	0.14
550	1.08	0.08	0.17	1.06	0.06	1.04	0.04	0.11
570	1.14	0.14	0.24	1.05	0.05	1.00	0.00	0.13

n	α	α_bias	α_MSE	β	β_Bias	λ	λ_bias	λ_MSE
590	1.05	0.05	0.12	1.08	0.08	1.05	0.05	0.11
610	1.05	0.05	0.15	1.08	0.08	1.07	0.07	0.12
630	1.09	0.09	0.17	1.06	0.06	1.03	0.03	0.12
650	1.08	0.08	0.15	1.07	0.07	1.04	0.04	0.11
670	1.09	0.09	0.18	1.09	0.09	1.04	0.04	0.14
690	1.06	0.06	0.14	1.08	0.08	1.05	0.05	0.12
710	1.07	0.07	0.13	1.07	0.07	1.04	0.04	0.11
730	1.09	0.09	0.18	1.06	0.06	1.03	0.03	0.1
750	1.04	0.04	0.1.0	1.05	0.05	1.03	0.03	0.07
770	1.03	0.03	0.12	1.09	0.09	1.07	0.07	0.11
790	1.04	0.04	0.11	1.07	0.07	1.04	0.04	0.09
810	1.03	0.03	0.09	1.06	0.06	1.04	0.04	0.07
830	1.08	0.08	0.12	1.05	0.05	1.02	0.02	0.1
850	1.09	0.09	0.11	1.02	0.02	0.99	-0.01	0.07
870	1.11	0.11	0.15	1.01	0.01	0.99	-0.01	0.08
890	1.08	0.08	0.13	1.04	0.04	1.02	0.02	0.08
910	1.04	0.04	0.08	1.04	0.04	1.03	0.03	0.07
930	1.06	0.06	0.13	1.05	0.05	1.02	0.02	0.08
950	1.06	0.06	0.1	1.04	0.04	1.01	0.01	0.08
970	1.00	0	0.07	1.07	0.07	1.05	0.05	0.06
990	1.02	0.02	0.10	1.07	0.07	1.05	0.05	0.09

Figure 2-4 demonstrate the bias, MSE and convergence probability of 95% CI interval

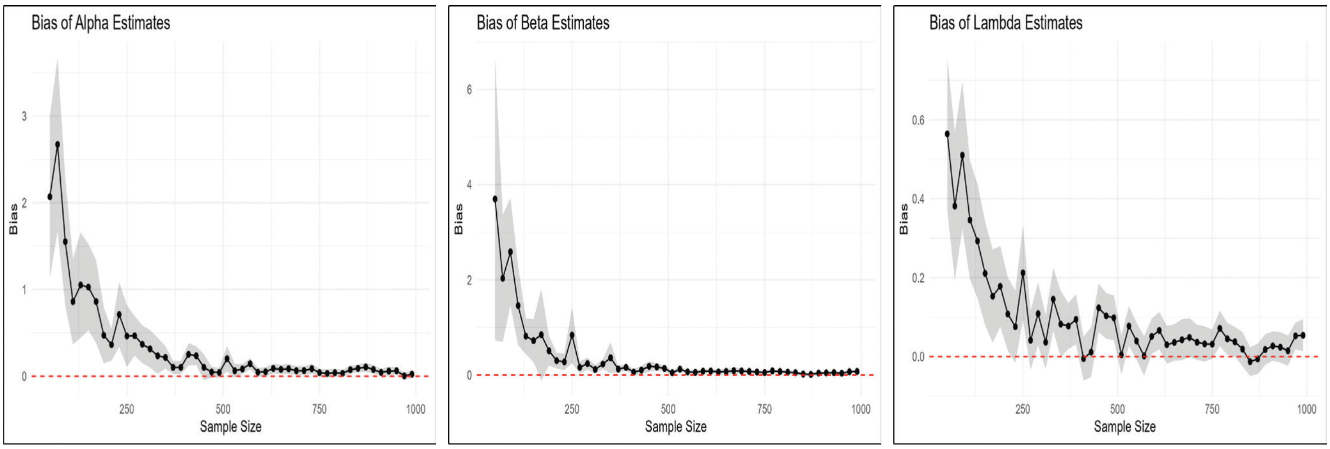

Figure 2: Bias of Parameter Estimates

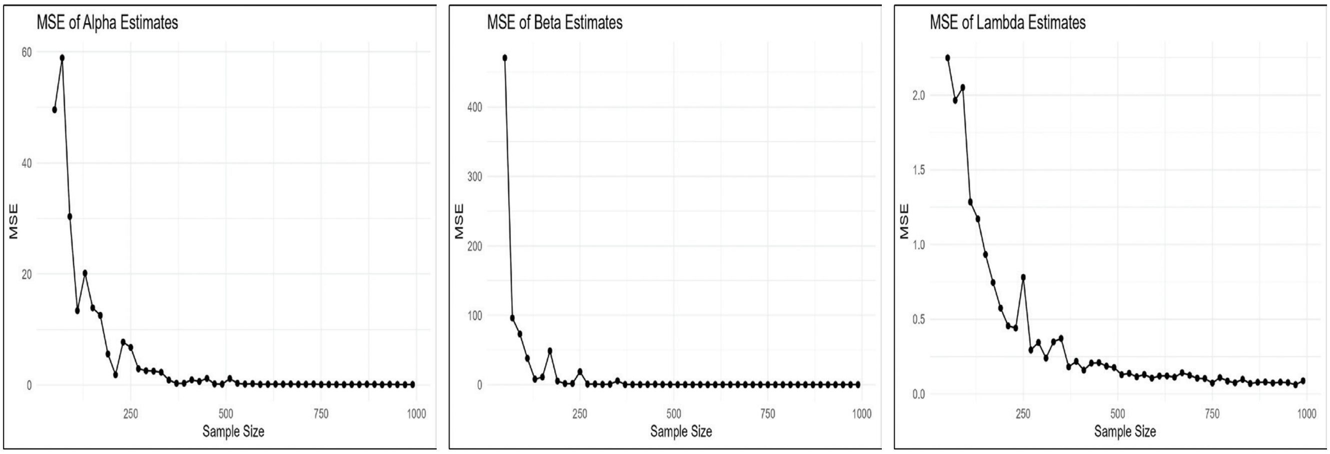

Figure 3: Mean Squared Error of Parameter Estimates

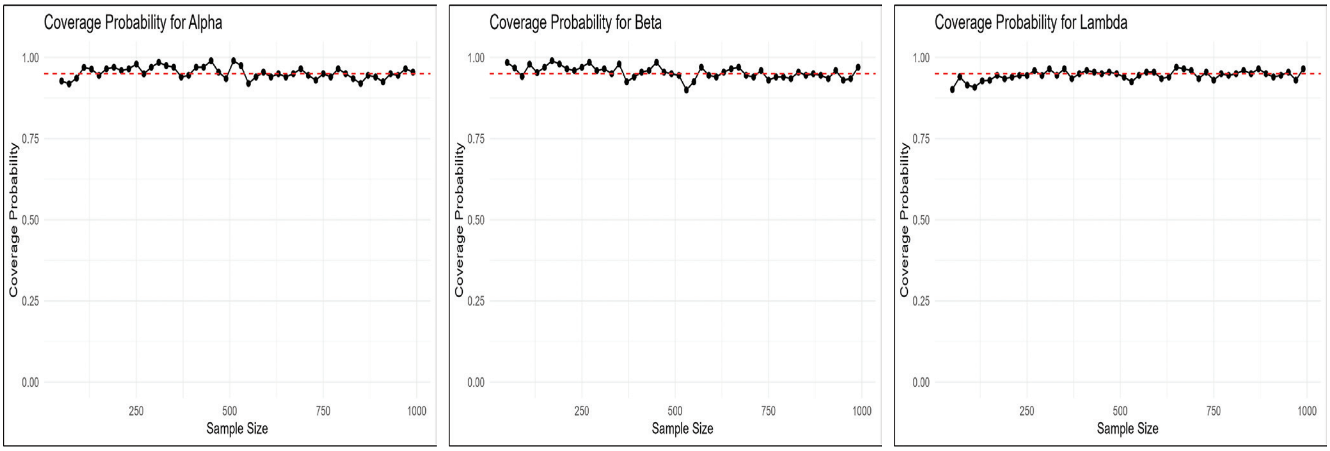

Figure 4: Coverage Probability of 95% Confidence Intervals

Parameter Estimation

Parameter of the model are estimated using maximum likelihood (MLE), Least square methods (LSE) and Cramer-von-Mises methods (CVM).

Maximum Likelihood estimation (MLE)

Maximum likelihood estimation uses the likelihood function to estimate the parameters of the model. The likelihood function of the MLL is given by equation (7)

$$l\left(\alpha,\beta,\lambda \mid x\right) = n\log 2 + n\log\alpha + n\log\beta + n\log\lambda + \lambda\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\ln\left(1+\beta e^{\lambda x_i}\right) + \left(\alpha-1\right)\sum_{i=1}^{n}\ln\left[\left(1+\beta e^{\lambda x_i}\right)^2 - 1\right] - 2\sum_{i=1}^{n}\ln\left[1+\left\{\left(1+\beta e^{\lambda x_i}\right)^2 - 1\right\}^{\alpha}\right] \quad (7)$$

The maximum likelihood estimates (MLEs) of the parameters α, β, and λ can be obtained by maximizing the likelihood function. Analytic methods of optimizing are rigorous so it can be optimized numerically by use of newton-Rapson methods, BFGS or Nelder- mead method.

Least Square Methos (LSE)

Least square method is aa alternative methods for parameter estimation which minimize the sum of squared differences between the empirical and theoretical cumulative distribution function. In these methods of estimation, we minimize the sum of square of the differences between F (x, α, β, λ) and ωi. The LSE function (eq 8) is minimized for a set of parameters that will be the estimated parameter.

$$LSE(\alpha,\ \beta,\ \lambda \mid x) = \sum_{i=1}^{n} \left(F(x_i;\alpha,\beta,\lambda) - \omega_i \right)^2 \text{; where, } \omega_i = \frac{i}{n+1} \tag{8}$$

Differentiating equation (8) with respect to parameters, equating to zero and then simultaneous solution for unknown parameter gives the estimated parameter value. In Case of non-linear equation, it becomes quite impossible to solve the equation we use numerical methods to estimate the parameters.

Cramer-von Mises methods of estimation (CVM)

This is another option for estimation which also minimize the difference the sum of squared differences between the empirical and theoretical cumulative distribution function. In these methods of estimation also, we minimize the sum of square of the differences between F (x, α, β, λ) and ωi. The CVM function (eq 9) is minimized for a set of parameters that will be the estimated parameter.

$$CVM(\alpha,\ \beta,\ \lambda \mid x) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(x_i;\alpha,\beta,\lambda) - \omega_i \right)^2 \text{; where, } \omega_i = \frac{2i-1}{n} \tag{9}$$

Application To real dataset

In this section of study, we have applied the MLL distribution on a real dataset to check the applicability of the model. Birnbaum and Saunders (1969) originally analyzed the data given below which represents the fatigue life of 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second (cps) which consists of 101 observations with maximum stress per cycle 31,000 psi.

70, 96, 212, 90, 97, 99, 100, 103, 104, 104, 105, 107, 108, 108, 108, 109, 109, 112, 112, 113, 114, 114, 114, 116, 119, 120, 120, 120, 121, 121, 123, 124, 124, 124, 124, 124, 128, 128, 129, 129, 130, 130, 130, 131, 131, 132, 131, 131, 131, 132, 132, 133, 134, 134, 134, 134, 134, 136, 136, 137, 138, 138, 138, 139, 139, 141, 141, 142, 142, 142, 142, 142, 142, 144, 144, 145, 146, 148, 148, 149, 151, 151, 152, 155, 156, 157, 157, 157, 157, 158, 159, 162, 166, 163, 164, 166, 163, 168, 170, 174, 196.

Parameters of the model are estimated using optim () function of R software (R Core Team, 2024). Summary statistics which give insight dataset. Summary statistics give descriptive measure of the data. Summary statistics demonstrate that the model in non-normal with positive skewness.

Table 2: Summary Statistics of the dataset.

Min	Q1	Median	Mean	Q3	Max	SK	Kurtosis
70	120	133	133.7	146.0	212.0	0.3304	4.0528

Figure 5 demonstrate the box plot and TTT plots. Box plots confirms that the data set is positively skewed while

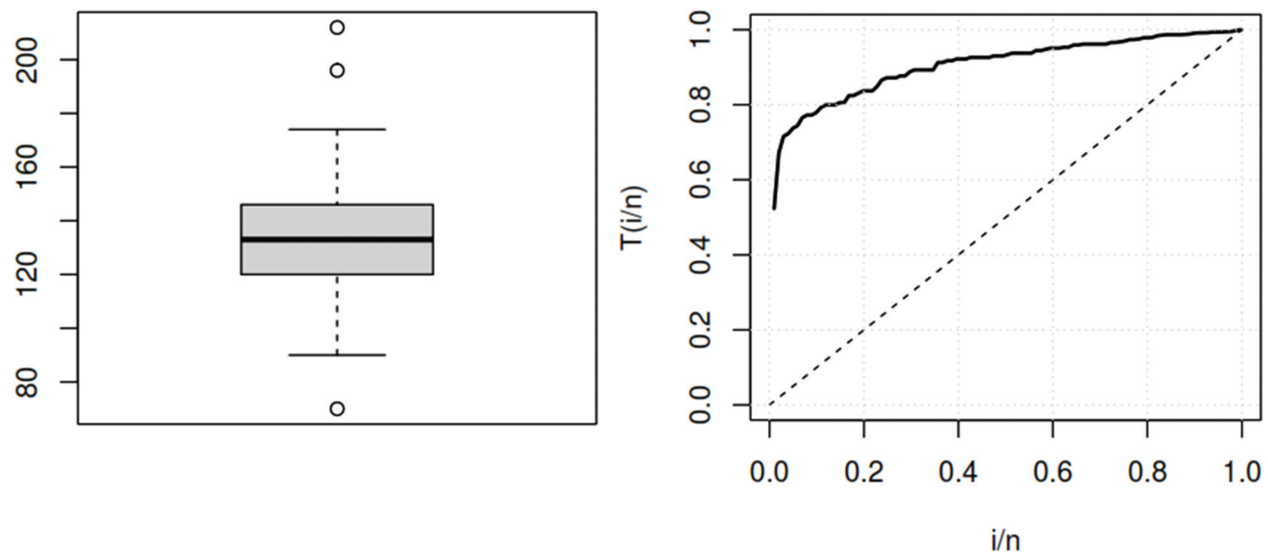

Figure 5: Whisker Box plot (Left) and TTT plot (Right side)

Table 3: Estimated parameters, SE, and 95% Credible Intervals

Method	Parameter	Estimate	SE	CI_lower	CI_upper
LSE	alpha	8.261	2.305	3.035	11.010
	beta	0.137	0.048	0.021	0.185
	lambda	0.008	0.004	0.006	0.022
CVM	alpha	6.382	2.429	3.121	11.351
	beta	0.099	0.049	0.027	0.182
	lambda	0.011	0.004	0.006	0.020
MLE	alpha	8.335	0.974	6.425	10.244
	beta	0.138	0.019	0.101	0.175
	lambda	0.008	0.001	0.006	0.010

Table 4 displays the information criterion values under MLE, LSE and CVM estimates. We have also compared the information criteria, Akaike information (AIC), Bayesian Information Criteria (BIC), Hanna- Quinan Information Criteria (HQIC) and Consistent Akaike Information Criteria (CAIC).

Table 4: Information Criteria under MLE, LSE and CVM

Methods	_-LL	AIC	BIC	CAIC	HQIC
MLE	455.658	917.316	925.161	917.563	920.492
LSE	457.048	920.096	927.941	920.343	923.272
CVM	457.3542	920.708	928.554	920.956	923.885

Table 5: KS, AD and CVM test for MLE, LSE and CVM methods of estimation

Methods	KS(p-value)	AD(p-value)	CVM(p-value)
MLE	0.049(0.969)	0.267(0.961)	0.036(0.953)
LSE	0.112(0.162)	1.573(0.160)	0.278(0.156)
CVM	0.124(0.091)	1.885(0.106)	0.373(0.085)

To describe, how well model fits the dataset, the histogram versus fitted probability density plot is demonstrated in figure 6 (Left) indicating that model fits data better as curve cover most of the are covered by histogram. Similarly,

Figure 6 (Right) demonstrate the empirical cumulative distribution plot versus fitted cdf indicating model fits datasets more adequately.

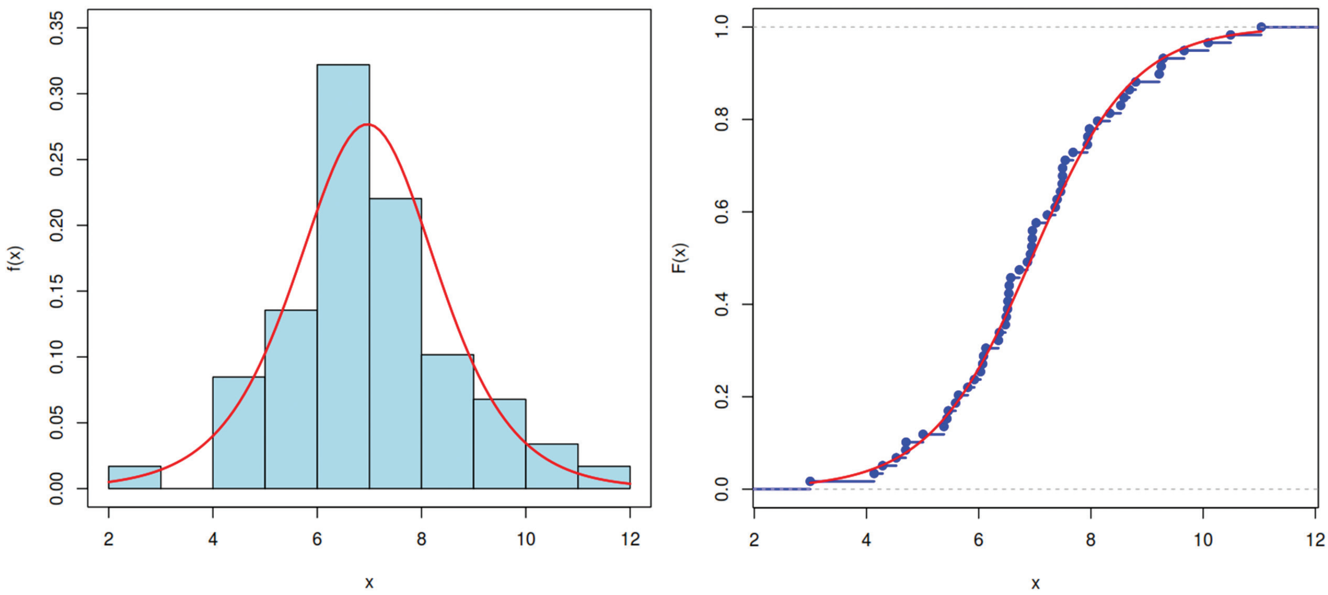

Figure 6: Histogram versus fitted PDF(Left) and ECDF vs Fitted CDF

The P-P and Q-Q plots are the graphical testing of the model whether it fit the data or not. PP plot in figure 7(Left) shows that most of the data point lie in the diagonal indicating that data fits the model well. The QQ plot in figure 7(right) indicate that data fits the proposed model better but a little week fitting at the tails.

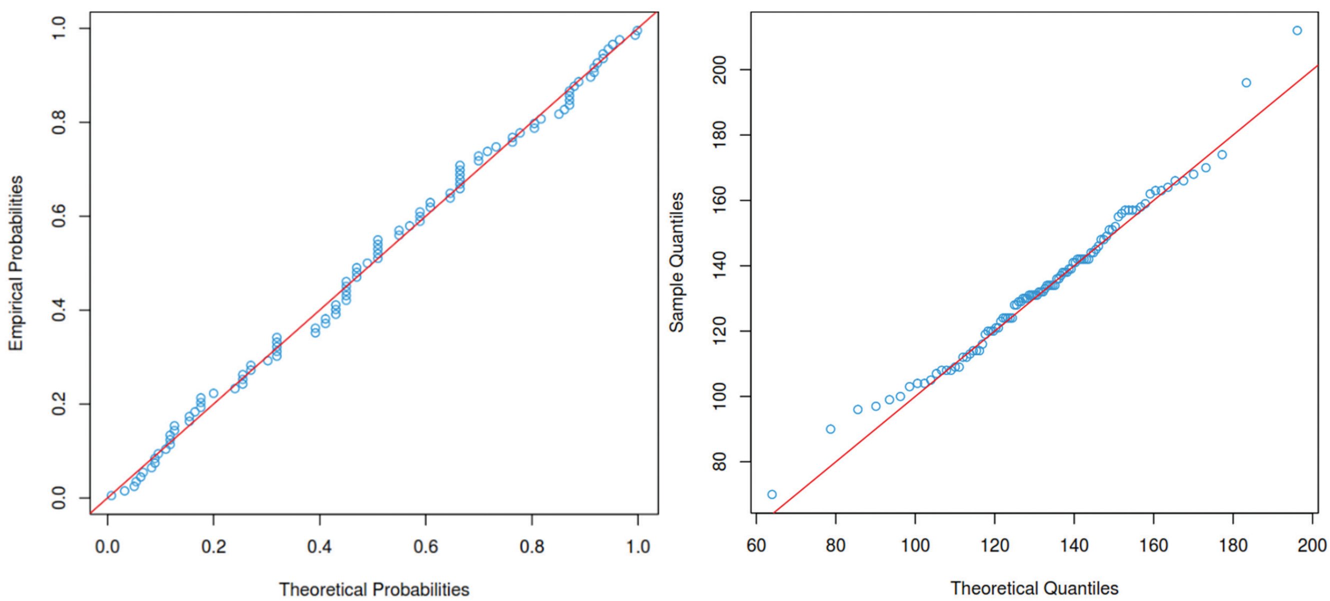

Figure 7: P-P plots (Left) and Q-Q plots (Right) of the MLL

To prove the superiority of the MLL distribution, it is compared with some classical models. The model compared are Generalized Rayleigh (GR) distribution (Kundu & Raqab, 2005), Generalized Exponential (GE) distribution (Gupta & Kundu, 1999) and Exponential power (EP) distribution (Smith & Bain, 1975). Results demonstrate that model MLL fits better compared to classical models (Table 6)

Table 6: Log likelihood and information criteria of competing models

Models	-LL	AIC	BIC	HQIC	CAIC
MLL	455.658	917.316	925.161	917.563	920.492
GR	457.767	918.756	923.980	918.878	920.871
GE	463.732	941.465	936.600	931.875	933.582
EP	476.789	957.579	962.810	957.700	959.697

Also, to test the goodness of fit and comparing as better fit, KS, AD and CVM test statistics and respective p values are mentioned in table 7.

Table 7: Test Statistics and respective p values

Models	KS(p-value)	AD(p-value)	CVM(p-value)
MLL	0.049(0.969)	0.267(0.961)	0.036(0.953)
GR	0.090(0.385)	0.603(0.645)	0.105(0.562)
GE	0.107(0.201)	2.072(0.084)	0.311(0.126)
EP	0.137(0.043)	4.506(0.005)	0.694(0.013)

To verify the superiority of MLL model, Histogram against the fitted pdf of MLL and competing models are displayed in 8(left). Furthermore, empirical CDF versus theoretical CDF of MLL and competing models are plotted in figure 8(Right)

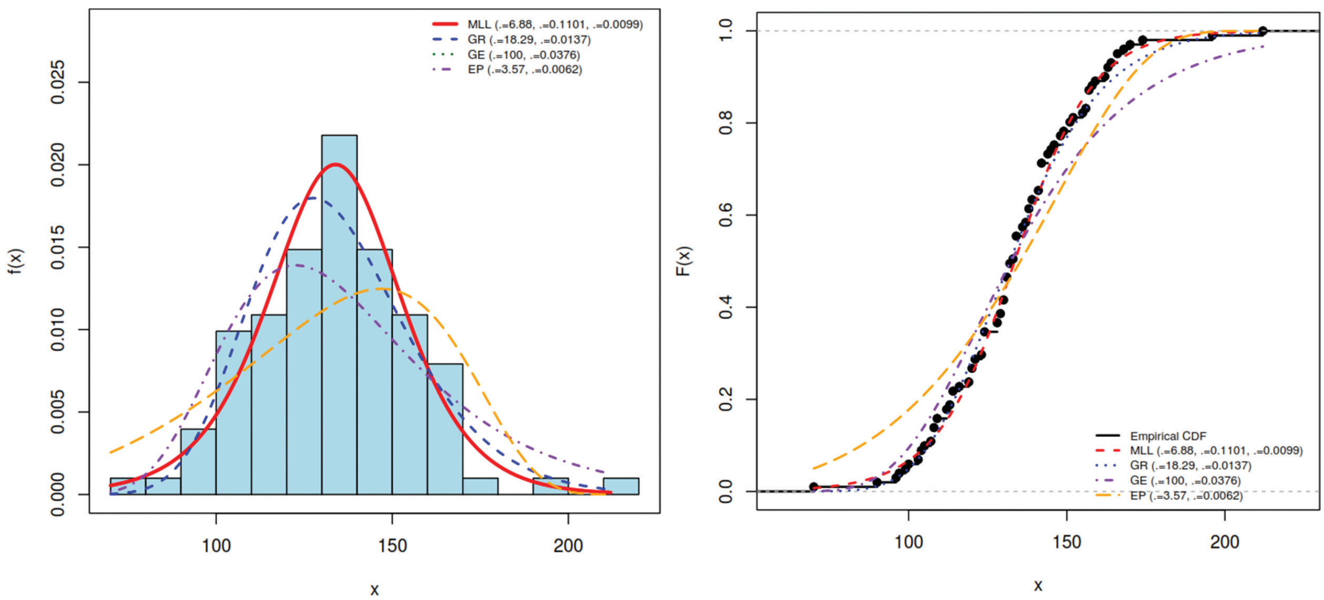

Figure 8: Histogram versus fitted pdfs and ECDF versus theoretical CDF

Conclusion

This study introduces Modified Logistic Lomax Probability distribution formulated using Logistic Lomax (MLL) Distribution. Modification of the model enabled MLL more flexible an applicable for real dataset more precisely. The simulation study demonstrates the performance of maximum likelihood estimation for the proposed distribution. Parameters are estimated using Maximum Likelihood estimation (MLE), Least square estimation (LSE) and Cramer- von Mises Estimation showing consistent parameter estimates. Information criterion and the goodness of fit testing support the MLL as better compared to some classical distributions. The MLL distribution will play a vital role in modern data analysis and will help researchers to know about the formulation of custom probability models and its uses in real life problem solution.

Limitations of study

Although MLL give better results on some of the classical probability model. Comparisons with another model is not performed here. Application on only one set of data may not generalize overall performances of the MLL model.

Future Work of study

Application of the model on different types of datasets, studying basic statistical properties and application of some vital methods of estimation like Bayesian estimation may be subjects of further study.

Conflicts of Interest

Authors has no any conflict of interest

References

F.J. Richards. A flexible growth function for empirical use. *Journal of Experimental Botany, 10*(29): 290–300, 1959.

Windarto, W., Eridani, E., & Purwati, U. D. (2018). A new modified logistic growth model for empirical use. *Communication in Biomathematical Sciences, 1*(2): 122. DOI: 10.5614/cbms.2018.1.2.5

Lemonte, A. J., & Cordeiro, G. M. (2013). An extended Lomax distribution. *Statistics, 47*(4), 800-816. https://doi.org/10.1080/02331888.2011.568119

Telee LBS, Yadav RS, Kumar V(2023). Modified Inverse Lomax Distribution: Model and properties. *Discovery Scientific Society,* Discovery 2023; 59: e110d1352 doi: https://doi.org/10.54905/disssi.v59i333.e110d1352

Chaudhary, A. K., & Kumar, V. (2020). The logistic Lomax distribution with properties and applications. *International Journal of New Technology and Research, 6*(12), 74-80. https://doi.org/10.31871/IJNTR.6.12.21

Chaudhary, A. K., Telee, L. B. S., & Kumar, V. (2022). Inverse Exponentiated Odd Lomax Exponential Distribution: Properties and Applications. *Nepalese Journal of Statistics, 6*(01), 29–50. https://doi.org/10.3126/njs.v6i01.50801

Smith, R.M. and Bain, L.J. (1975). An exponential power life-test distribution, *Communications in Statistics, 4,* 469-481

Gupta, R. D., & Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical planning and inference, 137*(11), 3537-3547. https://doi.org/10.1016/j.jspi.2007.03.030

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

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www. R-project. org/.*

Birnbaum, Z. W., & Saunders, S. C. (1969). Estimation for a family of life distributions with applications to fatigue. *Journal of applied probability, 6*(2), 328-347.