Enhancing the Applicability of Log Uniform Distribution using Weightage Parameter

Abstract

In modeling real-life datasets bounded within a finite interval, several probability distributions characterized by location parameters are available. Among these, the uniform distribution is commonly used due to its assumption of equal probability across the specified range. The log uniform distribution is used in scenarios where varying probabilities are to be assigned across the specified range. However, in scenarios involving length-biased sampling—where larger values receive proportionally greater probabilities—the log uniform distribution becomes inadequate. To address this limitation, the present paper introduces a length-biased (weighted) log uniform distribution. The proposed model provides greater flexibility and thereby broadens the scope of potential applications. Its statistical properties are derived, and parameter estimation is carried out using the method of maximum likelihood. A comprehensive simulation study is also conducted to evaluate the performance of the estimators.

Key words: Length biased, log uniform, parametric estimation, weighted.

1. Introduction

Probability distributions are the foundation of statistical modeling. Classical distributions like Normal, Poisson, Exponential, Binomial, etc., have been extensively used in modeling, but real-world data often show complexities that these traditional models cannot fully capture, which necessitates introduction of new distributions. New distributions are often motivated by fields like finance (modeling extreme risks), epidemiology (disease counts), reliability (failure times), and climatology (rainfall modeling). The need for new distributions arises primarily due to flexibility in modeling over-dispersion and under-dispersion, mixture and heterogeneity in the data as well as bounded data. In order to address over or under-dispersion of the data, many novel distributions are recently proposed. Lord et al.(2008) considered Conway-Maxwell-Poisson distribution for its flexibility in handling both overdispersion and underdispersion in the crash data. Saha et al. (2020) showed that the Poisson-Tweedie distribution can be the better model for modeling crash counts data. Combining two or more distributions is another choice to better capture over-dispersion. Sankaran(1970), Shanker and Mishra (2014) attempted compounding of Poisson distribution with Lindley distribution. In survival or reliability analysis, lifetime distributions are extensively used. For instance, Prodhani and Shanker (2025) applied truncated Sujatha distribution for medical dataset. Patil et al. (2025) used truncated Shambhu distribution for modeling strength of aircraft window glass. But there are many instances where the survival or reliability data is confined in a certain interval. Very few distributions are available for modeling such situations. Log-uniform (LU) distribution, which is bounded between two points, offer decreasing density, unlike constant density offered by uniform distribution. Anu and Sebastian (2025) proposed an idea of exponentiated log-uniform distribution as a generalization of LU distribution, which offers more flexibility due to the added parameter. Shaji and Sebastian (2023) developed transmuted log-uniform distribution and they found it to be a better fit for certain datasets as compared to Transmuted Weibull distribution (Aryal and Tsokos (2011)) and Transmuted Quasi-Akash distribution (Hassan et al. (2022)).

In many fields, primarily in reliability and survival analysis, the data obtained is often not a random data. Longitudinal data may guarantee randomness, but the cross-sectional data often introduces selection bias in the data, as pointed out by Vardi(1982). Length bias is a selection bias in which a probability with which sample unit is drawn is proportional to its size (characteristic value). If this bias is not taken into account while analyzing the sample, erroneous conclusions will be made about the underlying model or distribution. The length biased sample, instead of representing its parent density, represents weighted version of the parent density. Due to this fact researchers are concerned about weighted versions of the known probability densities. This paper attempts the generalization of LU distribution to introduce weighted LU distribution.
The rest of the paper is organized as follows. Section 2 introduces weighted

log-uniform (WLU) distribution. Statistical properties of the proposed WLU distribution are discussed in section 3. The reliability properties of WLU distribution are investigated in section 4. In section 5, parametric estimation is discussed in brief. The performance of the estimation procedure is discussed in section 6. Section 7 highlights the conclusions.

2. Weighted Log Uniform Distribution

A random variable X is said to follow log uniform distribution, if $ln(X)$ follows Uniform distribution,where ln is the natural logarithm function (the logarithm to base e). The probability density function (PDF) of random variable X following log uniform distribution is given as follows.

$$f(x;a,b) = \frac{1}{x[ln(b) - ln(a)]}; a \leq x \leq b, 0 < a < b \in R \tag{1}$$

where a and b are the location parameters of the distributions that define the minimum and maximum values of the distribution. The cumulative distribution function (CDF) is given by:

$$F(x;a,b) = \frac{ln(x) - ln(a)}{ln(b) - ln(a)}; a \leq x \leq b, 0 < a < b \in R \tag{2}$$

The weighted log uniform distribution is defined as following.

$$f_w(x) = \frac{w(x)}{E(w(x))} f(x); a \leq x \leq b, 0 < a < b \in R \tag{3}$$

where $w(x)$ is the non-negative weight function, with expectation $E(w(x))$. We have different choices of weight functions based on the different weight model. In the paper, we use $w(x) = x^r, r = 1, 2, ...$ as a weight function. Then the Probability function of weighted log-uniform distribution ($WLU(a,b,r)$)is given as follows.

$$f_w(x) = \frac{x^r}{E(x^r)} f(x); a \leq x \leq b, 0 < a < b \in R, r = 1, 2, ... \tag{4}$$

Therefore, the density of weighted log-uniform distribution with parameters a, b and r is given as follows.

$$f_w(x) = \frac{r x^{r-1}}{b^r - a^r}; a \leq x \leq b, 0 < a < b \in R, r = 1, 2, ... \tag{5}$$

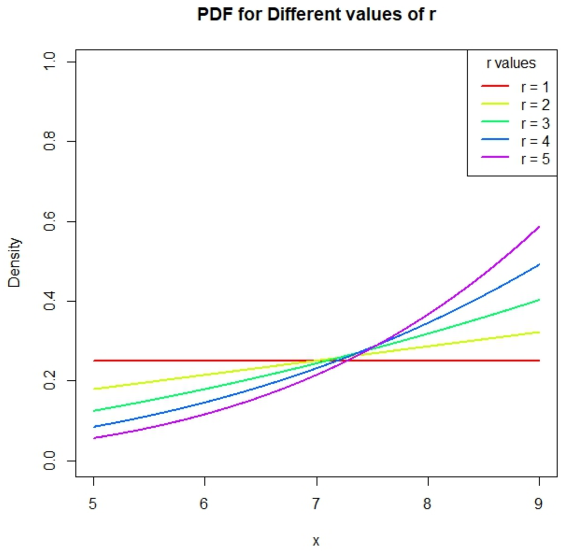

Figure 1: Plot of pdf for different values of r

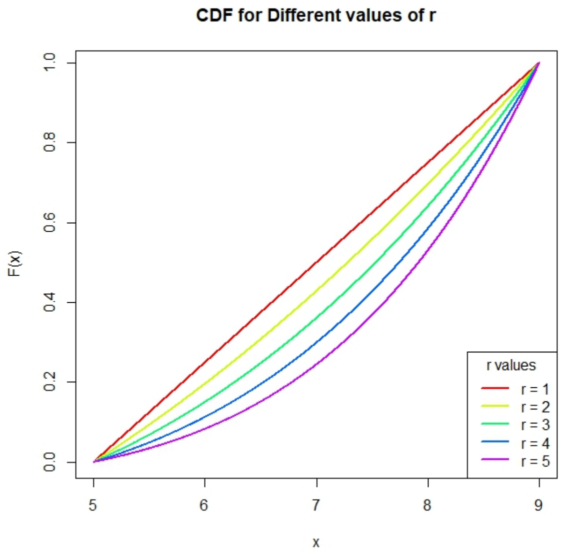

Figure 2: Plot of cdf for different values of r

The cumulative distribution function (cdf) of weighted log-uniform distribution is given by-

$$F_w(x) = \frac{x^r - a^r}{b^r - a^r}; a \leq x \leq b, 0 < a < b \in R, r = 1, 2, ... \tag{6}$$

The Fig. 1 shows the PDF of the Weighted Log-Uniform Distribution for different values of r. As r increases, the density starts to concentrate more towards the upper parameter b of the distribution. That is, higher r values result in steeper rises and more concentrated density near the upper end. This demonstrates how the parameter r influences the skewness of the distribution. The Fig. 2 shows the cdf of the Weighted Log-Uniform Distribution for different values of r. Lower values of r result in faster accumulation of probability at comparatively smaller values of x.

3. **Statistical properties**

In this section, we discuss the moments, coefficient of variation and entropy of the weighted log-uniform distribution.

If X is a random variable that follows a weighted log uniform $WLU(a, b, r)$ distribution with pdf as given in (6), then $k^{th}(k = 1, 2, ...)$ order raw moment about origin μ'_k is obtained as -

$$\mu'_k = E(X^k) = \int_a^b x^k f_w(x) dx \tag{7}$$

On simplification, we get the expression for k^{th} order raw moment about origin μ'_k as follows.

$$\mu'_k = E(X^k) = \frac{r}{b^r - a^r} \frac{1}{k + r}(b^{k+r} - a^{k+r}) \tag{8}$$

Putting k=1,2,3,4, we get first four raw moments of the weighted log-uniform distribution as follows.

$$\mu'_1 = \frac{r}{b^r - a^r} \frac{1}{r + 1}(b^{r+1} - a^{r+1})$$

$$\mu'_2 = \frac{r}{b^r - a^r} \frac{1}{r + 2}(b^{r+2} - a^{r+2})$$

$$\mu'_3 = \frac{r}{b^r - a^r} \frac{1}{r + 3}(b^{r+3} - a^{r+3})$$

$$\mu'_4 = \frac{r}{b^r - a^r} \frac{1}{r + 4}(b^{r+4} - a^{r+4})$$

Using these moments, the basic descriptive measures of weighted log-uniform distribution are obtained.

$$E(X) = \frac{r(b^{r+1} - a^{r+1})}{(r+1)(b^r - a^r)} \tag{9}$$

$$Variance(X) = \frac{r(b^{r+2} - a^{r+2})}{(r+2)(b^r - a^r)} - (\frac{r(b^{r+1} - a^{r+1})}{(r+1)(b^r - a^r)})^2 \tag{10}$$

The coefficient of variation (CV) is as follows.

$$CV = \frac{(r+1)(b^r - a^r))}{r(b^{r+1} - a^{r+1})} \sqrt{\frac{r(b^{r+2} - a^{r+2})}{(r+2)(b^r - a^r)} - \left[\frac{r(b^{r+1} - a^{r+1})}{(r+1)(b^r - a^r)}\right]^2} \tag{11}$$

4. **Reliability properties**

In this section, we study survival function, hazard function, reverse hazard function for the weighted log-uniform distribution with parameters a, b, and r given in (6).

The survival function $S_w(x)$ of the weighted log-uniform distribution is given by-

$$S(x; a, b, r) = 1 - \frac{ln(x) - ln(a)}{ln(b) - ln(a)}; a \leq x \leq b, 0 < a < b \in R$$

$$S(x; a, b, r) = \frac{ln(x) - ln(a)}{ln(b) - ln(a)}; a \leq x \leq b, 0 < a < b \in R \tag{12}$$

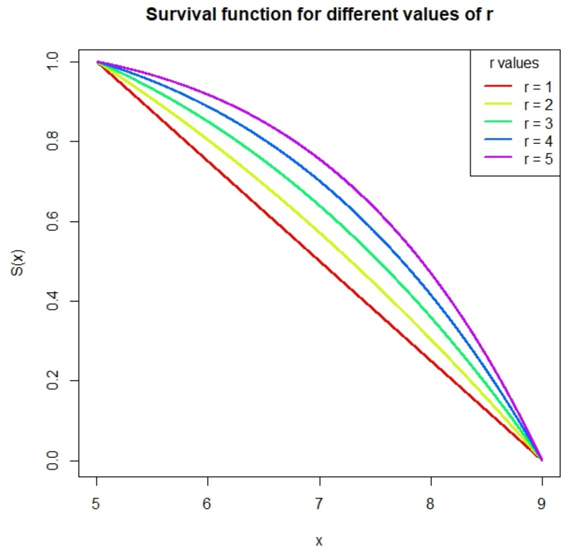

Figure 3: Plot of Survival function for different values of r

The Fig.3 shows survival functions $S(x)$ of the Weighted Log-Uniform distribution for different values of r. As r increases, the rate of decline of the function lowers, and then it decreases rapidly as compared to the case of $r = 1$.
The hazard function $h(x)$ of a weighted log-uniform distribution is given by-

$$h(x) = \frac{f_w(x)}{S_w(x)}$$

$$h(x) = \frac{r}{b^r - x^r} x^{r-1} \tag{13}$$

WLU distribution belongs to IFR family of distributions. The variation in the hazard function w.r.t. r is depicted in the Figure 4.

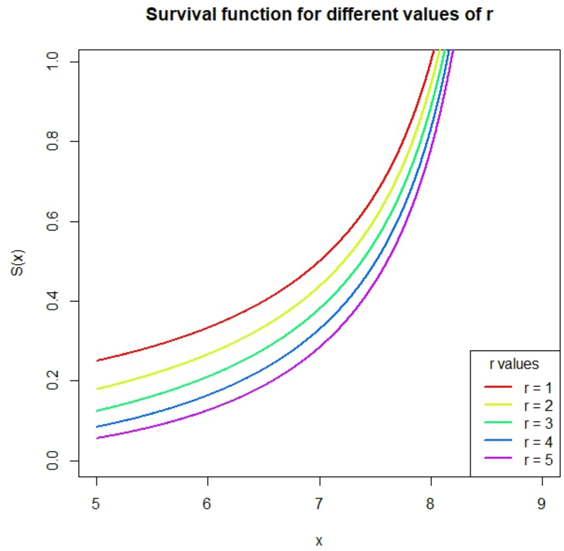

Figure 4: Plot of Hazard function for different values of r

The reverse hazard function of a weighted log-uniform distribution is given by-

$$h(x) = \frac{f_w(x)}{F_w(x)}$$

$$h(x) = \frac{r}{x^r - a^r} x^{(r-1)} \tag{14}$$

The cumulative hazard function of a weighted log-uniform distribution is given by-

$$H(x) = -log\left[\frac{b^r - x^r}{b^r - a^r}\right] \tag{15}$$

Distribution of order statistics:
Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ be the order statistics of a random sample $X_1, X_2, ..., X_n$ of size n, $n \in N$ taken from weighted log-uniform distribution. Then the probability density function of m^{th} order statistics $X_{(m)}$ is defined as

$$f_{X_{(m)}}(x) = \frac{n!}{(m-1)!(n-m)!} \frac{rx^{(r-1)}}{(b^r - a^r)^n} (x^r - a^r)^{m-1} (b^r - x^r)^{n-m} \tag{16}$$

5. **Parameter estimation**

In this section, we discuss the estimation of the unknown parameters a, b, r of WLU distribution using the method of maximum likelihood (ML). Let $x_1, x_2, x_3, \ldots, x_n$ be a random sample of size n drawn from $WLU(a, b, r)$ distribution. Then the likelihood function corresponding to this sample is given as -

$$L(a, b, r; \underline{x}) = \prod_{i=1}^{n} \frac{r}{b^r - a^r} x_i^{r-1} \tag{17}$$

where $\underline{x} = x_1, x_2, x_3, \ldots, x_n$.
Considering the setup of underlying distribution, it is clear that $x_{(1)}$ and $x_{(n)}$ are the ML estimators of a and b respectively. To estimate r, we obtain the log likelihood as follows.

$$ln(L(a, b, r; \underline{x})) = n * ln(r) - n * ln(b^r - a^r) + (r - 1) \sum_{i=1}^{n} ln(x_i) \tag{18}$$

As the two parameters a and b are estimated using sample minimum and sample maximum respectively, therefore only one parameter r is to be estimated using Eq.(18). Moreover, r, the parameter to be estimated, is integer valued, so it can be found out by trial and error values with value incrementation of one in each iteration. Alternatively, one may use numerical methods to arrive at the argument that maximizes Eq.(18).

6. **Simulation study**

This section discusses the algorithm to obtain random sample from the proposed $WLU(a, b, r)$ distribution as well as the parametric estimation based on this sample. In order to know the performance of the estimators, we have drawn sample of size (say) n from $WLU(a, b, r)$. Considering the parameters to be unknown, we estimated those using the sample. This procedure is repeated for 10000 such samples to get 10000 estimates of each parameter.

The algorithm for obtaining a sample of size n from $WLU(a, b, r)$ distribution is given below-

(1) Specify a, b and r.

(2) Specify the sample size n.

(3) Draw a random number, say U, from $U(0, 1)$ distribution.

(4) Obtain random sample X as $X = ((b^r - a^r) * U + a^r)^(1/r)$

(5) Repeat step (3) and (step(4) n times to get n realizations on $WLU(a, b, r)$.

For simulation study, we have taken $a = 5, b = 9, r = 3$. Also, sample sizes (n) are taken to be $5, 10, 20, 40$ and simulation sizes are taken to be $100, 1000, 10000, 100000$. Following are the results obtained using simulation.

Table 1: Estimates and MSEs for the parameter a.

				Sample size		
			5	**10**	**20**	**40**
Simulation size	100	Estimate	5.716	5.4175	5.2452	5.1208
		MSE	0.195	0.1172	0.0474	0.0101
	1000	Estimate	5.7156	5.416	5.2277	5.1175
		MSE	0.2622	0.1264	0.0432	0.0133
	10000	Estimate	5.704	5.4073	5.2253	5.1194
		MSE	0.2722	0.1141	0.0408	0.0125
	100000	Estimate	5.6994	5.4089	5.2255	5.12
		MSE	0.2666	0.114	0.0406	0.0127

Table 2: Estimates and MSEs for the parameter b.

				Sample size		
		5	**10**	**20**	**40**	
Simulation size	100	Estimate	7.6225	7.7726	7.8874	7.9553
		MSE	0.1222	0.0475	0.0119	0.0021
	1000	Estimate	7.6227	7.8037	7.9043	7.9519
		MSE	0.1108	0.0370	0.0083	0.0022
	10000	Estimate	7.6411	7.8057	7.9022	7.9504
		MSE	0.1055	0.0344	0.0093	0.0024
	100000	Estimate	7.6325	7.8076	7.9022	7.9500
		MSE	0.1110	0.0335	0.0091	0.0024

Table 3: Estimates and MSEs for the parameter r.

			Sample size			
			5	**10**	**20**	**40**
Simulation size	100	Estimate	1	1	2	3
		MSE	5.96	5.62	2.63	1.35
	1000	Estimate	1	1	1	3
		MSE	5.742	5.532	5.188	1.431
	10000	Estimate	1	1	1	3
		MSE	5.5492	5.437	5.2513	1.4367
	100000	Estimate	1	1	1	3
		MSE	5.5856	5.5255	5.2885	1.4188

From Table 1, effect of sample size on the performance of estimator of a is clearly observable. As sample size increases, the estimate tends towards the actual value of the parameter a. Same thing is observed for estimators of b and r also (Table 2 and Table 3). The effect of simulation size is also noteworthy. With the increase in the simulation size, the estimators are performing better.

7. Conclusions

For the datasets, naturally constrained between two values, a flexible model, namely weighted log uniform distribution is proposed. The statistical and reliability properties of the distribution are studied in detail. The estimation process is also prescribed for the developed model. Simulation study is done to assess the applicability of the distribution.

References

1. Lord, D., Guikema, S. D., & Geedipally, S. R. (2008). Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. Accident Analysis & Prevention, 40(3), 1123-1134.

2. Saha, D., Alluri, P., Dumbaugh, E., & Gan, A. (2020). Application of the Poisson-Tweedie distribution in analyzing crash frequency data. Accident Analysis & Prevention, 137, 105456.

3. Sankaran, M. (1970). 275. note: The discrete Poisson-lindley distribution. Biometrics, 145-149.

4. Shanker, R., & Mishra, A. (2014). A two-parameter Poisson-Lindley distribution. International journal of Statistics and Systems, 9(1), 79-85.

5. Prodhani, H. R., & Shanker, R. (2025). The Truncated Sujatha Distribution with Properties and Applications in Engineering and Medical Sciences. Journal of Statistical Theory and Practice, 19(4), 85.

6. Dhanashri Patil, Manjushri Chikkalli, Shruti Hipale, & Chandrakant Gardi (2025). Modeling strength of the aircraft window glass by truncated Shambhu distribution. Reliability: Theory & Applications, 20 (4 (89)), 996-1005. doi: 10.24412/1932-2321-2025-489-996-1005

7. Anu Av. & Sebastian, R. (2025). Some applications of exponentiated log-uniform distribution. Reliability: Theory & Applications, 20(1 (82)), 124-134.

8. Shaji, A. K., & Sebastian, R. (2023). Some applications of transmuted log-uniform distribution. Reliability: Theory & Applications, 4(76), 1046-1055.

9. Aryal, G. R., & Tsokos, C. P. (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution. European Journal of pure and applied mathematics, 4(2), 89-102.

10. Hassan, A., Wani, S. A., Shafi, S., & Dar, S. A. (2022). Transmuted quasi akash distribution applicable to survival times data. Thailand Statistician, 20(2), 338-356.

11. Vardi, Y. (1982). Nonparametric estimation in the presence of length bias. The Annals of Statistics, 10(2), 616-620.