STATISTICA, anno LXXXII, n. 4, 2022

ON THE SHIFTED HYBRID LOG-NORMAL DISTRIBUTION

Damodaran Santhamani Shibu
Department of Statistics, University College, Thiruvananthapuram- 695034, Kerala, India

Soman Latha Nitin
Department of Statistics, University College, Thiruvananthapuram- 695034, Kerala, India

Muhammed Rasheed Irshad [1]
Department of Statistics, Cochin University of Science and Technology (CUSAT), Cochin 682022, Kerala, India

SUMMARY

The log-normal distribution is widely used to model positive valued data in many areas of applied research. However, sometimes the log-normal distribution does not completely satisfy the fitting expectations in every real life situations. In this paper, we introduce, investigate, and discuss a more flexible shifted hybrid log-normal distribution for which the log-normal distribution is a special case. Also, various properties, special cases and estimation procedure of the new distribution are discussed. Moreover, the performances of maximum likelihood estimators of the parameters are examined using a brief simulation study. The flexibility and performance of the new distribution is also illustrated through two applications by fitting two real datasets of different situations.

Keywords: Hybrid log-normal distribution; Estimation; Lifetime data; Fracture toughness data; Simulation.

1. INTRODUCTION

Log-normal distribution has a widespread application in many fields of biological sciences and physical sciences and also in Economics and Business. Examples include astrophysics (see Gandhi, 2009), environmental sciences (see Benning and Barnes, 2009), computer science (see Doerr *et al.*, 2013), economics (see Cobb *et al.*, 2013), biomedical (see Feng *et al.*, 2013) and radiology (see Neti and Howell, 2008). Limpert *et al.* (2001)

[1] Corresponding Author. E-mail: irshadmr@cusat.ac.in

compared the use of the log-normal distribution across several different science disciplines. The log-normal distribution has a long-standing and vibrant history. Galton (1879) and McAlister (1879) initiated the study of the distribution in papers published together, relating it to the use of the geometric mean as an estimate of location. Much later, Kapteyn and van Uven (1916) discussed the genesis of the distribution and gave a graphical method for estimating the parameters. Wicksell (1917) used the method of moments for three-parameter estimation, introducing a third parameter, the threshold, to fit the distribution of ages at first marriage. Nydell (1919) obtained asymptotic standard errors for the moment estimates. The distribution appeared in papers of the 1930s that developed probit analysis in bioassay. Later, Yuan (1933) introduced the bivariate version of log-normal distribution. The log-normal distribution has been often applied to analyse data on occupational radiation exposure by many authors since the work made by Gale (1967).

The probability density function for a log-normal random variable W is given by

$$q(w) = \frac{1}{\sqrt{2\pi}\sigma w} \exp\left[-\frac{(\log w - \mu)^2}{2\sigma^2}\right], \quad w > 0, \ \mu \in \mathbb{R}, \ \sigma > 0. \tag{1}$$

Kumazawa and Numakunai (1981) introduced a new distribution called the Hybrid log-normal distribution, which is more suitable than the log-normal distribution for fitting data on occupational radiation exposure. A continuous random variable Y is said to follow a Hybrid log-normal (HLN) distribution if its corresponding cumulative distribution function (cdf) and probability density function (pdf) are respectively given by

$$H_Y(y|\alpha, \sigma, \mu) = \int_0^y \frac{1}{\sqrt{2\pi}} \left(\frac{\alpha}{u} + \sigma\right) \exp\left[-\frac{(\alpha \log\, u + \sigma\, u + \mu)^2}{2}\right] du, \tag{2}$$

and

$$h_Y(y) = \frac{1}{\sqrt{2\pi}} \left(\frac{\alpha}{y} + \sigma\right) \exp\left[-\frac{(\alpha \log\, y + \sigma\, y + \mu)^2}{2}\right], \quad y > 0, \tag{3}$$

where $\alpha \geq 0, \sigma \geq 0, \alpha + \sigma > 0$ and $\mu \in \mathbb{R}$. For $\sigma = 0$, the pdf in Eq.(3) reduces to log-normal distribution ($LN(-\mu\alpha^{-1}, \alpha^{-2})$).

Also, a little investigation has been done on a theoretical basis under this distribution. In this paper, we introduce and investigate the properties of a new extended version of HLN distribution by shifting the location of the HLN random variable. A key advantage of this distribution is that the log-normal distribution, which is well known to be in use in a wide spectrum of disciplines includes a sub-distribution.

The rest of the paper is organized as follows. In Section 2, we introduce the novel distribution and discuss its special cases, moments, reliability measures and order statistics. Section 3 deals with the maximum likelihood estimation procedure of the proposed distribution. Then, to analyse the performance and flexibility of maximum likelihood estimators of the distribution parameters, a simulation study has been conducted in Section 4. Section 5 illustrates the applications of the new distribution due to two real-life datasets of different situations. Finally, Section 6 covers the penultimate concluding remarks.

2. The Shifted Hybrid Log-Normal Distribution

A location parameter shifts the density function to the left or right on the horizontal axis. It is found suitable in modelling where the data points do not fall below some regions. In reliability theory, the location parameter indicates that a failure cannot occur before this time, and it is also referred to as the failure-free time. We here propose to add a location parameter in the pdf of the HLN distribution given in Eq.(3) and discuss the aspects of estimation and statistical properties.

The proposed four-parameter distribution will be referred to as the *Shifted hybrid log-normal* distribution, and hereafter denoted as $SHLN(\lambda, \alpha, \sigma, \mu)$ or simply SHLN.

Definition 1. *If a random variable (RV) Y follows the HLN distribution in Eq.(3), then the RV $X = Y + \lambda$ is said to follow the SHLN distribution. The cdf of the random variable X is given by*

$$F(x) = \int_{\lambda}^{x} \frac{1}{\sqrt{2\pi}} \left(\frac{\alpha}{u-\lambda} + \sigma \right) \exp\left\{ -\frac{1}{2} [\alpha \log(u-\lambda) + \sigma\ (u-\lambda) + \mu]^2 \right\}\ du,$$

which can also be written as

$$F(x) = \Phi(\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu), \tag{4}$$

where $\Phi(\cdot)$ is the cdf of Standard Normal distribution. Then, the pdf of the random variable X is provided by

$$f(x) = \frac{1}{\sqrt{2\pi}} \left(\frac{\alpha}{x-\lambda} + \sigma \right) \exp\left\{ -\frac{1}{2} [\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2 \right\}, \qquad x > \lambda, \tag{5}$$

where $\lambda \geq 0$, $\alpha \geq 0$, $\sigma \geq 0$, $\alpha + \sigma > 0$ and $\mu \in \mathbb{R}$.

Also, plots of Figure 1 display the cdf and Figure 2 display the pdf of the SHLN distribution. Now, the moment generating function of the SHLN distribution is expressed as

$$M(t) = \frac{I^*(t)}{\sqrt{2\pi}},$$

where

$$I^*(t) = \int_{\lambda}^{\infty} \left(\frac{\alpha}{x-\lambda} + \sigma \right) \exp\left\{ tx - \frac{1}{2} [\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2 \right\},$$

which is well-defined for all $t \in \mathbb{R}$.

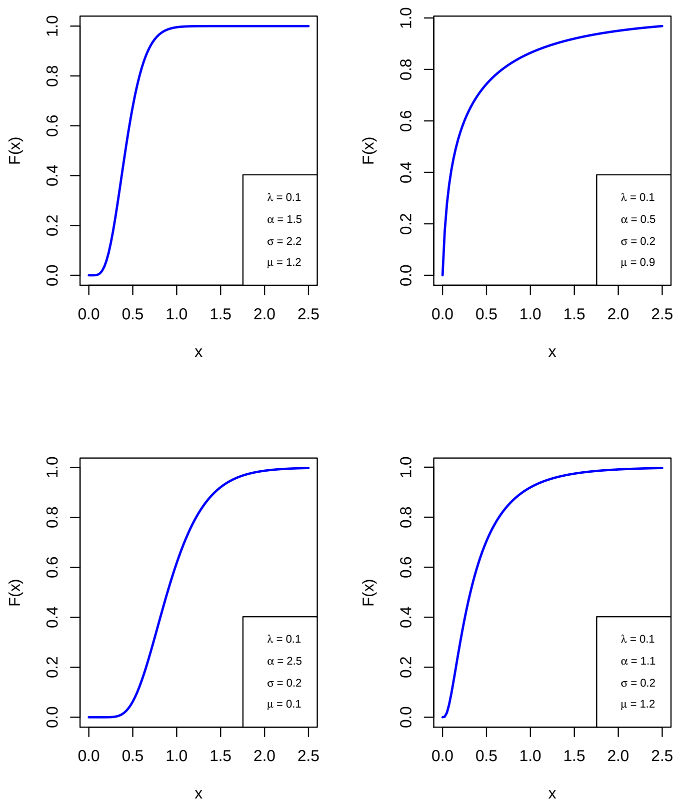

Figure 1 – CDF plot of SHLN distribution

2.1. *Special cases*

The SHLN distribution defined by Eq. (5) has the following sub-models.

1. For $\lambda = 0$, SHLN reduces to the HLN distribution given in Eq. (3).

2. For $\sigma = 0$, SHLN reduces to the 3-parameter log-normal distribution.

3. For $\lambda = 0$ and $\sigma = 0$, SHLN reduces to the 2-parameter log-normal distribution given in Eq. (1).

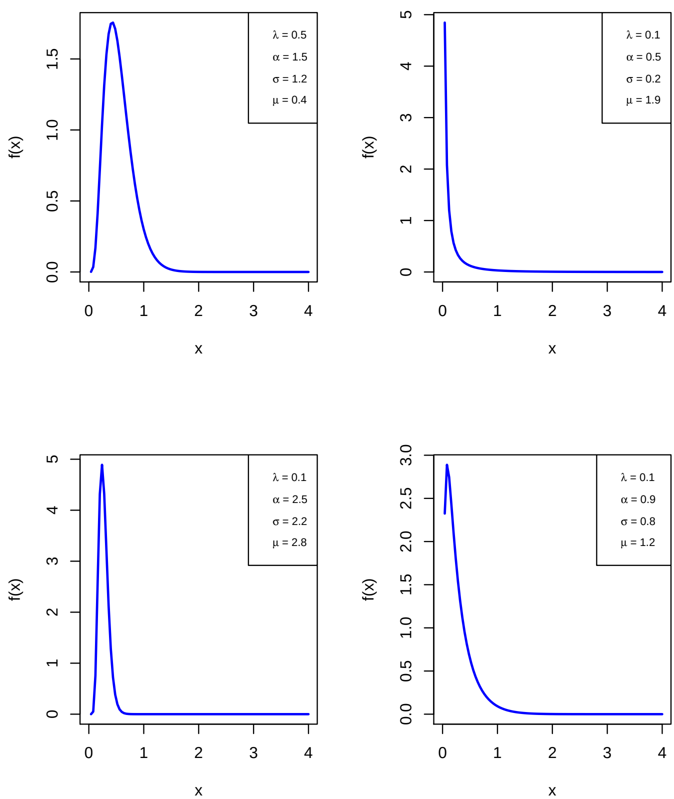

Figure 2 – PDF plot of SHLN

2.2. *Reliability Measures*

The main purpose of system reliability analysis is to identify the critical components in a system and to quantify the impact of component failures. So here, the functions of reliability measures of SHLN distribution are indispensable to derive.

Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$, then the general formulas for the survival function $S(x)$, hazard rate function $h(x)$, cumulative hazard rate function $R(x)$ and the reversed hazard rate function $r(x)$ are respectively given by $S(x) = 1 - F(x)$, $h(x) = \frac{f(x)}{S(x)}$, $R(x) = -\ln(S(x))$, $r(x) = \frac{f(x)}{F(x)}$.

2.2.1. *Functions of Survival, Hazard rate, Cumulative hazard rate and Reversed hazard rate*

The survival function $S(x)$ of SHLN distribution is given by

$$\begin{aligned} S(x) =& 1 - F(x) \\ =& 1 - \int_{\lambda}^{x} \frac{1}{\sqrt{2\pi}} \left(\frac{\alpha}{x-\lambda} - \sigma \right) \exp\left\{ -\frac{1}{2} [\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2 \right\} \mathrm{d}x, \end{aligned}$$

which can be written as

$$S(x) = 1 - \Phi(\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu). \tag{6}$$

The hazard rate function $h(x)$ of SHLN distribution is given by

$$h(x) = \frac{(\alpha - \sigma(x-\lambda))\, e^{-\frac{1}{2}[\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2}}{\sqrt{2\pi}\,(x-\lambda)\,[1 - \Phi(\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu)]}. \tag{7}$$

The cumulative hazard rate function $R(x)$ of SHLN distribution is given by

$$R(x) = -\ln(\,1 - \Phi(\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu)). \tag{8}$$

The reversed hazard rate function $r(x)$ of SHLN distribution is given by

$$r(x) = \frac{(\alpha - \sigma(x-\lambda)) \exp\left\{ -\frac{1}{2}[\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2 \right\}}{\sqrt{2\pi}\,(x-\lambda)\,\Phi(\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu)}, \tag{9}$$

where $\Phi(\cdot)$ is the cdf of the Standard Normal distribution.

Plots in Figure 3 portray the hazard rate function of SHLN distribution. It displays, the hazard function possesses various shapes including increasing, decreasing and bath-tub shapes.

2.2.2. *Conditional moments, Vitality function and Moments of Residual and Reverse residual life*

For lifetime distributions, it is of greater interest to know the conditional moments $E(X^n|X > t)$, $n = 1, 2, \ldots$ which are important in prediction. Thus, conditional moments of SHLN distribution is given by

$$E(X^n|X > t) = \frac{1}{\sqrt{2\pi}\, S(t)} \varkappa(n,t), \tag{10}$$

where

$$\varkappa(n,t) = \int_{t}^{\infty} x^n \left(\frac{\alpha}{x} + \sigma \right) \exp\left\{ -\frac{1}{2}[\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2 \right\} \mathrm{d}x. \tag{11}$$

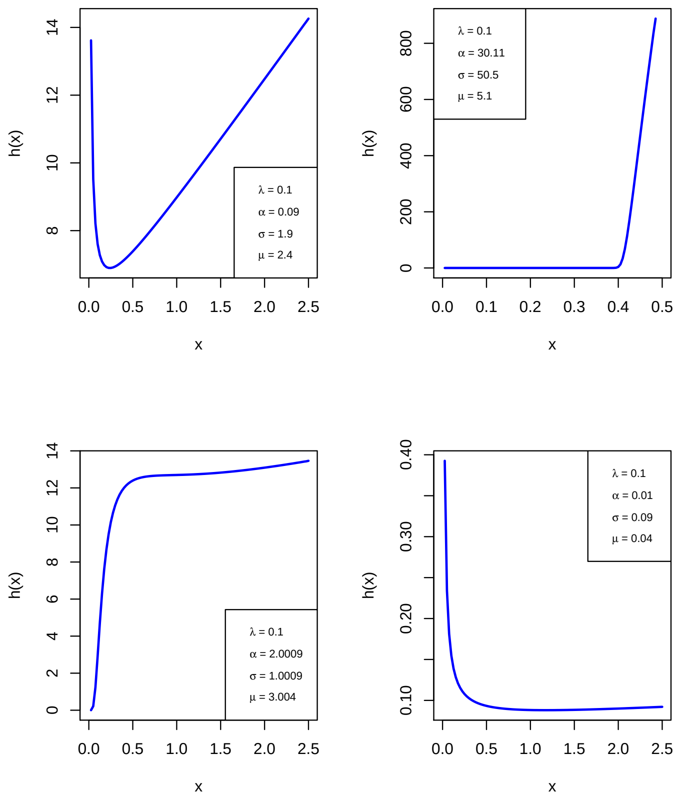

Figure 3 – Hazard rate function plot of SHLN distribution

The vitality function is a very useful tool in modeling life-time data. This function play important roles in reliability engineering, biomedical science, and survival analysis. It is worth mentioning that the rapid ageing of a component needs to low vitality relatively, whereas high vitality implies relatively slow ageing during the given time period. For $n = 1$ in the conditional moments given in Eq.(10), gives the vitality function $V(t)$ of SHLN distribution, that is,

$$
\begin{aligned}
V(t) =& E(X|X > t) \\
=& \frac{\varkappa(1,t)}{\sqrt{2\pi}\, S(t)},
\end{aligned} \tag{12}
$$

where $\varkappa(n,t)$ is given in Eq.(11).

The concept of geometric vitality function based on the geometric mean of the residual lifetime. If X be a random variable which represents the lifetime of a component, then $\log G(t)$ represents the geometric mean of lifetimes of components which have survived up to time t. For a non-negative random variable X follows an absolutely continuous distribution function, with $E(\log X) < 1$, the geometric vitality function is defined as

$$\begin{aligned} \log G(t) &= E(\log X | X > t) \\ &= \frac{1}{S(t)} \int_t^{\infty} \log x f(x) \mathrm{d}x, \end{aligned} \tag{13}$$

where $S(t) = P(X > t)$ denotes the survival function. Now, the geometric vitality function of SHLN random variable is given by

$$\log G(t) = \frac{\Omega(t)}{\sqrt{2\pi}\, S(t)}, \tag{14}$$

where

$$\Omega(t) = \int_t^{\infty} \log x \left(\frac{\alpha}{x} + \sigma\right) \exp\left\{-\frac{1}{2}[\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2\right\} \mathrm{d}x. \tag{15}$$

The r^{th} order moment of the residual life of the SHLN distribution is given by

$$\begin{aligned} \mu_r(t) &= E[(X-t)^r | X > t] \\ &= \frac{1}{S(t)} \int_t^{\infty} \sum_{i=0}^{r} \binom{r}{i} (-t)^{r-i} x^i f(x) \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\, S(t)} \sum_{i=0}^{r} \binom{r}{i} (-t)^{r-i}\, \varkappa(i,t), \end{aligned} \tag{16}$$

where $\varkappa(n,t)$ is given in Eq. (11). Hence, the Mean residual life (MRL) function of the SHLN distribution is given by

$$\begin{aligned} \mu_1(t) &= E[(X-t) | X > t] \\ &= \frac{1}{S(t)} \int_t^{\infty} \sum_{i=0}^{1} \binom{1}{i} (-t)^{1-i} x^i f(x) \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\, S(t)} \sum_{i=0}^{1} \binom{1}{i} (-t)^{1-i}\, \varkappa(i,t) \\ &= V(t) - t. \end{aligned}$$

Hence,

$$\mu_1(t) = \frac{\varkappa(1,t)}{\sqrt{2\pi}\, S(t)} - t. \tag{17}$$

Similarly, the second moment of the residual lifetime of the SHLN distribution is given by

$$\begin{aligned} \mu_2(t) &= t^2 - 2t\ V(t) + \varkappa(2,t) \\ &= t^2 - 2t\ \frac{\varkappa(1,t)}{\sqrt{2\pi}\ S(t)} + \varkappa(2,t). \end{aligned} \tag{18}$$

The variance of the residual life function of the SHLN distribution can be obtained by using $\mu_1(t)$ and $\mu_2(t)$.

The r^{th} order moment of the reversed residual life of the SHLN distribution is given by

$$\begin{aligned} m_r(t) &= E[(t-X)^r | X \le t] \\ &= \frac{1}{F(t)} \int_0^t (t-x)^r f(x) \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\ F(t)} \sum_{i=0}^{r} \binom{r}{i} (-1)^{r-i} t^i\ \Delta(r,i,t), \end{aligned} \tag{19}$$

where

$$\Delta(n,i,t) = \int_0^t x^{(n-i)} \left(\frac{\alpha}{x} + \sigma\right) \exp\left\{-\frac{1}{2}[\alpha \log(x-\lambda) + \sigma(x-\lambda) + \mu]^2\right\} \mathrm{d}x. \tag{20}$$

Now, the mean $(m_1(t))$ and second moment $(m_2(t))$ of the reversed residual life of the SHLN distribution can be obtained by setting $r = 1,2$ respectively in Eq.(19) and is given by

$$m_1(t) = \frac{1}{\sqrt{2\pi}} \left[t - \frac{\Delta(1,0,t)}{F(t)} \right], \tag{21}$$

and

$$m_2(t) = \frac{1}{\sqrt{2\pi}} \left[t^2 + \frac{\Delta(2,0,t) - 2t\ \Delta(2,1,t)}{F(t)} \right], \tag{22}$$

where $\Delta(n,i,t)$ is given in Eq. (20). Also, using $m_1(t)$ and $m_2(t)$, one can obtain the variance of the reversed residual life function of the SHLN distribution.

2.3. *Order Statistics*

Let $X_1, X_2, \dots, X_n$ be a random sample from the SHLN distribution and its order statistics is $X_{1:n}, X_{2:n}, \dots, X_{n:n}$. Let $f_{i:n}(x)$ and $F_{i:n}(x)$ denote the pdf and the cdf of the i^{th} order statistic $X_{i:n}$, respectively. Hence, using the standard expressions of order statistics, we find that

$$f_{i:n}(x)=\frac{1}{\sqrt{2\pi}}\frac{n!}{(i-1)!(n-i)!}\left(\frac{\alpha}{x-\lambda}-\sigma\right)\exp\left\{-\frac{1}{2}[\alpha\log(x-\lambda)+\sigma(x-\lambda)+\mu]^2\right\}$$
$$\times\left[\Phi(\alpha\log(x-\lambda)+\sigma(x-\lambda)+\mu)\right]^{i-1}$$
$$\times\left[1-\Phi(\alpha\log(x-\lambda)+\sigma(x-\lambda)+\mu)\right]^{n-i}, \tag{23}$$

and

$$F_{i:n}(x)=\sum_{j=i}^{n}\binom{n}{j}\left[\Phi(\alpha\log(x-\lambda)+\sigma(x-\lambda)+\mu)\right]^{j}$$
$$\times\left[1-\Phi(\alpha\log(x-\lambda)+\sigma(x-\lambda)+\mu)\right]^{n-j}, \tag{24}$$

where $\Phi(\cdot)$ is the cdf of the Standard Normal distribution.

3. Maximum Likelihood Estimation

Let $X_1,X_2,...,X_n$ be a random sample from SHLN$(\lambda,\alpha,\sigma,\mu)$ distribution. The log-likelihood function for the parameter vector $\theta=(\lambda,\alpha,\sigma,\mu)^T$ is given by

$$\mathscr{L}_n=-n\log\sqrt{2\pi}+\sum_{i=1}^{n}\log\left(\frac{\alpha}{x_i-\lambda}+\sigma\right)-\frac{1}{2}\sum_{i=1}^{n}[\alpha\log(x_i-\lambda)+\sigma(x_i-\lambda)+\mu]^2. \tag{25}$$

The score function associated with the log-likelihood function is

$$\mathbf{U}=\left(\frac{\partial\mathscr{L}_n}{\partial\lambda},\frac{\partial\mathscr{L}_n}{\partial\alpha},\frac{\partial\mathscr{L}_n}{\partial\sigma},\frac{\partial\mathscr{L}_n}{\partial\mu}\right)^T, \tag{26}$$

where

$$\frac{\partial\mathscr{L}_n}{\partial\mu}=0\Rightarrow\hat{\mu}=-\frac{1}{n}\sum_{i=1}^{n}\left[\hat{\alpha}\log(x_i-\hat{\lambda})+\hat{\sigma}(x_i-\hat{\lambda})\right], \tag{27}$$

$$\frac{\partial\mathscr{L}_n}{\partial\alpha}=\sum_{i=1}^{n}\frac{1}{\hat{\alpha}+\hat{\sigma}(x_i-\hat{\lambda})}$$
$$-\sum_{i=1}^{n}\log(x_i-\hat{\lambda})\left[\hat{\alpha}\log(x_i-\hat{\lambda})+\hat{\sigma}\,(x_i-\hat{\lambda})-\hat{\alpha}\,\overline{\log(x-\hat{\lambda})}-\hat{\sigma}\,\overline{(x-\hat{\lambda})}\right], \tag{28}$$

$$\frac{\partial\mathscr{L}_n}{\partial\sigma}=\sum_{i=1}^{n}\frac{x_i}{\hat{\alpha}+\hat{\sigma}(x_i-\hat{\lambda})}$$
$$-\sum_{i=1}^{n}(x_i-\hat{\lambda})\left[\hat{\alpha}\log(x_i-\hat{\lambda})+\hat{\sigma}\,(x_i-\hat{\lambda})-\hat{\alpha}\,\overline{\log(x-\hat{\lambda})}-\hat{\sigma}\,\overline{(x-\hat{\lambda})}\right] \tag{29}$$

and $\overline{\log(x-\hat{\lambda})} = \frac{1}{n}\log(x_i-\hat{\lambda})$ and $\overline{(x-\hat{\lambda})} = \frac{1}{n}(x_i-\hat{\lambda})$.

Now, concerning the SHLN distribution, the likelihood function for estimating the location parameter is represented by multiplying an indicator function and is given by

$$\begin{aligned}\mathbf{L}(\lambda,\alpha,\sigma,\mu|\mathbf{x}) = & \left(\frac{1}{\sqrt{2\pi}}\right)^n \prod_{i=1}^n \left(\frac{\alpha}{x_i-\lambda}+\sigma\right) \\ & \times \exp\left\{-\frac{1}{2}\sum_{i=1}^n [\alpha\log(x_i-\lambda)+\sigma(x_i-\lambda)+\mu]^2\right\} \\ & \times I_{[\lambda,\infty)}(\min\{x_1,\dots,x_n\}).\end{aligned} \tag{30}$$

Here, $I_{[\lambda,\infty)}(\min\{x_1,\dots,x_n\})$ is an indicator function defined as

$$I_{[\lambda,\infty)}(\min\{x_1,\dots,x_n\}) = \begin{cases} 1 & \text{if } \min\{x_1,\dots,x_n\} \geq \lambda \\ 0 & \text{if } \min\{x_1,\dots,x_n\} < \lambda. \end{cases} \tag{31}$$

It can be seen from the above expression of $\mathbf{L}(\lambda,\alpha,\sigma,\mu|\mathbf{x})$ that the maximum likelihood estimator of λ is $X_{(1)} = \min\{x_1,\dots,x_n\}$.

4. Simulation

In this section, we conduct simulation experiments to assess the long-run performance of the MLEs of the SHLN parameters for some finite sample sizes. Thus, we generate samples of sizes $n = 25, 100, 500$ from the SHLN for the parameter values $\lambda = 0.001$, $\alpha = 0.2$, $\sigma = 0.9$ and $\mu = 0.6$ and iterated each sample 500 times. Then, we compute the average bias and MSE for all replications in the relevant sample sizes. That is, the analysis computes the values by the given formulas:

- Average bias of the simulated estimates $= \frac{1}{500}\sum_{i=1}^{500}(\hat{\Theta}_i - \Theta)$,
- Average MSE of the simulated estimates $= \frac{1}{500}\sum_{i=1}^{500}(\hat{\Theta}_i - \Theta)^2$,

where $\hat{\Theta} = (\hat{\lambda},\hat{\alpha},\hat{\sigma},\hat{\mu})$ are estimates of the parameter vector $\Theta = (\lambda,\alpha,\sigma,\mu)$. The simulation results are reported in Table 1.

TABLE 1
Estimates, Average bias and MSE values of MLEs from simulation of the SHLN distribution.

Sample Size	Parameters	Estimates	Bias	M.S.E
25	λ	0.1	0.10	0.01
	α	0.23	0.03	0.004
	σ	0.95	0.05	0.16
	μ	0.71	0.11	0.13
100	λ	0.1	0.10	0.01
	α	0.22	0.02	0.001
	σ	0.88	-0.02	0.04
	μ	0.67	0.07	0.05
500	λ	0.1	0.10	0.01
	α	0.22	0.02	0.001
	σ	0.85	-0.05	0.01
	μ	0.67	0.07	0.02

5. Application

We provide the applications using two real datasets to demonstrate the potentiality of the SHLN distribution over some parent models such as Log-Normal (LN), 3-parameter Log-Normal (3-LN), Gamma, Loglogistic and HLN distributions. In each case, the parameters are estimated by maximum likelihood estimation technique using RStudio software. The first dataset taken from Balakrishnan *et al.* (2009) was provided by the Mexican Institute of Social Security (IMSS) which contains the data on the lifetimes (in years) of retired women with temporary disabilities, which are incorporated in the Mexican insurance public system and who died during 2004. The second dataset taken from Nadarajah and Kotz (2006) is the fracture toughness data from a material namely $Bi_2Sr_2CaCu_2O_{8+x}$. We have fitted the above-mentioned parent distributions to the first and second datasets and computed the Log-likelihood (LL), Kolmogorov-Smirnov (KS), Anderson-Darling (A^*), Cramér-von Misses (W^*) statistics, AIC and BIC values and is presented in Table 2 and Table 3, respectively. It can be obtained that the SHLN fits both datasets better than the existing parent distributions given above.
From Table 2 and Table 3, it can be seen that the proposed SHLN distribution is the most suitable model for the datasets we used. That is, the SHLN distribution's goodness-of-fit statistics values are lower than those of the other distributions we compared. In other words, it is observed that the SHLN distribution has the smallest KS, A^*, W^*, AIC and BIC values. In this context, we can conclude that the proposed SHLN distribution provides a better fit than the compared distributions.

TABLE 2
Maximum-likelihood estimates, goodness-of-fit statistics, AIC and BIC values to dataset 1.

Estimates	LN	3-LN	Gamma	Loglogistic	HLN	SHLN
$\hat{\mu}$	3.84	3.84	-	-	-6.43	-2.57
$\hat{\sigma}$	0.23	0.23	0.42	47.34	0.08	0.09
$\hat{\alpha}$	-	-	19.99	7.61	0.66	0.06
$\hat{\lambda}$	-	0.0001	-	-	-	21.99
LL	-1060.54	-1060.54	-1055.85	-1063.38	-1052.86	-1047.46
KS	0.11	0.11	0.09	0.09	0.08	0.07
A^*	3.53	3.56	2.69	3.44	1.78	1.63
W^*	0.63	0.64	0.48	0.52	0.31	0.28
AIC	2125.08	2127.09	2115.70	2130.76	2111.72	2102.92
BIC	2132.35	2137.99	2122.97	2138.03	2122.62	2117.46

TABLE 3
Maximum-likelihood estimates, goodness-of-fit statistics, AIC and BIC values to dataset 2.

Estimates	LN	3-LN	Gamma	Loglogistic	HLN	SHLN
$\hat{\mu}$	0.83	3.06	-	-	-2.82	-1.27
$\hat{\sigma}$	0.35	1.28	3.70	2.35	0.76	1.07
$\hat{\alpha}$	-	-	8.93	4.82	1.16	0.11
$\hat{\lambda}$	-	0.002	-	-	-	1.20
LL	-14.31	-14.31	-14.10	-14.75	-13.99	-10.51
KS	0.18	0.18	0.17	0.16	0.15	0.14
A^*	0.35	0.35	0.31	0.35	0.28	0.24
W^*	0.06	0.06	0.05	0.06	0.05	0.04
AIC	32.61	34.61	32.21	33.49	33.99	29.02
BIC	33.58	36.07	33.18	34.46	35.45	30.96

6. Concluding Remarks

In this paper, we provide and analyse the features of a new extended version of the hybrid log-normal distribution by moving the location of the random variable. One significant feature of this distribution is that it includes the log-normal distribution, which is widely used in a variety of professions, as a sub-distribution. We offer precise expressions for different reliability metrics for the SHLN distribution. The hazard rate function of the SHLN distribution possesses an increasing, decreasing, and bathtub-shaped graphical representation. In terms of inference, the maximum likelihood estimation approach is utilised to estimate the distribution parameters. The goodness-of-fit tests are applied to two real datasets concerning the data on the lifetimes (in years) of retired women with temporary disabilities provided by the Mexican Institute of Social Security (IMSS), and the fracture toughness data from a material namely $Bi_2Sr_2CaCu_2O_{8+x}$. In terms of fitting, the novel distribution consistently outperforms the compared distribution in the literature. A simulation study is carried out to assess the performances of MLEs of the SHLN parameters, and their consistency is confirmed. We anticipate that the proposed model will find a broader range of applications in positive real-world dataset modelling.

Acknowledgements

The authors would like to express their gratitude to the Editor and the anonymous reviewers for their constructive comments, which significantly improved the current version of our article.

References

N. Balakrishnan, V. Leiva, A. Sanhueza, E. Cabrera (2009). *Mixture inverse Gaussian distributions and its transformations, moments and applications*. Statistics, 43, no. 1, pp. 91–104.

J. L. Benning, D. L. Barnes (2009). *The effects of scale and spatial heterogeneities on diffusion in volcanic breccias and basalts: Amchitka Island, Alaska*. Journal of Contaminant Hydrology, 106, no. 3, pp. 150 – 165.

B. R. Cobb, R. Rumí, A. Salmerón (2013). *Inventory management with log-normal demand per unit time*. Computers & Operations Research, 40, no. 7, pp. 1842 – 1851.

C. Doerr, N. Blenn, P. Mieghem (2013). *Lognormal infection times of online information spread*. PloS ONE, 8, no. 5, p. e64349.

C. Feng, W. Hongyue, N. Lu, X. Tu (2013). *Log transformation: application and interpretation in biomedical research*. Statistics in Medicine, 32, pp. 230–239.

H. J. GALE (1967). *Some examples of the application of the lognormal distribution in radiation protection*. The Annals of Occupational Hygiene, 10, no. 1, pp. 39–45.

F. GALTON (1879). *XII. The geometric mean, in vital and social statistics*. Proceedings of the Royal Society of London, 29, no. 196-199, pp. 365–367.

P. GANDHI (2009). *The flux-dependent rms variability of X-Ray binaries in the optical*. The Astrophysical Journal Letters, 697, p. L167.

J. KAPTEYN, M. VAN UVEN (1916). *Skew Frequency Curves in Biology and Statistics*. v. 2. Hoitsema Bros., Groningen.

S. KUMAZAWA, T. NUMAKUNAI (1981). *A new theoretical analysis of occupational dose distributions indicating the effect of dose limits*. Health physics, 41, pp. 465–475.

E. LIMPERT, W. A. STAHEL, M. ABBT (2001). *Log-normal distributions across the sciences: Keys and clues*. BioScience, 51, pp. 341–352.

D. MCALISTER (1879). *The law of the geometric mean*. Proceedings of the Royal Society of London, 29, pp. 367–376.

S. NADARAJAH, S. KOTZ (2006). *q exponential is a Burr distribution*. Physics Letters A, 359, pp. 577–579.

V. NETI, R. R. HOWELL (2008). *Lognormal distribution of cellular uptake of radioactivity: statistical analysis of alpha-particle track autoradiography*. Journal of Nuclear Medicine: official publication, Society of Nuclear Medicine, 49, pp. 1009–1016.

S. NYDELL (1919). *The mean errors of the characteristics in logarithmic-normal distributions*. Scandinavian Actuarial Journal, 1919, no. 1, pp. 134–144.

S. D. WICKSELL (1917). *On logarithmic correlation with an application to the distribution of ages at first marriage*. Meddelanden fran Lunds Astronomiska Observatorium Serie I, 84, pp. 1–22.

P. T. YUAN (1933). *On the logarithmic frequency distribution and the semi-logarithmic correlation surface*. The Annals of Mathematical Statistics, 4, no. 1, pp. 30–74.