

A Mathematical Analysis of the Dinesh–Umesh–Sanjay Lindley Mixture Mode

S. Ilani and N. Yassine

Keywords and phrases: DUS-Lindley, Weibull, mixed DUSLW distribution, entropy, EM algorithm.

This manuscript is an extended and significantly improved version of the paper entitled "Order Statistics of Mixed Random Variables: Theoretical Approach" presented at ICRAMI 2025 (Tunisia, November 2025).

Abstract Efficient and effective data analysis depends largely on accurate and flexible probability distributions, particularly extended versions of well-established models that provide a systematic framework for analyzing complex real-world data. This study presents and investigates a mixture of the DUS-Lindley and Weibull distributions as a theoretically sound framework for data analysis. The main statistical properties of the proposed mixture, including the PDF, CDF, entropy, ordinary and incomplete moments, and order statistics are explored. Parameter estimation for the proposed mixture is carried out using the Expectation–Maximization (EM) algorithm. The performance of the model is then examined through simulation studies across multiple sample sizes. Finally, to assess its practical relevance, applications to real-world datasets are provided.

1 Introduction

Data analysis, which necessitates the use of suitable and compatible statistical distributions, forms the fundamental means by which studies in various fields are conducted. Efforts to maximize the accuracy and effectiveness of data analysis have increased over the years, aiming to develop new distributions and methods, thus improving the understanding of data and the predictability of results. Consequently, many new distributions have been proposed recently, generating more applicable and realistic methods for data analysis that better capture real life data. For instance, Koki Shimizu et al. produced the distribution of the eigenvalues of a gamma matrix from the product of an extended Gaussian matrix and its transposed matrix [1].

In [2], the authors generated the discrete Lindley distribution, where they studied its statistical properties, estimated its parameters, and simulated the results. In addition, the quantile functions of various distributions were used to generate a new generalized family of Lindley distributions. Through real-world data fitting applications, this article demonstrated its strong performance [3]. Also, a new bounded distribution, the Marshall–Olkin distribution, which can be applied to data exhibiting both negative and positive skewness, was introduced [4]. In their study, the distribution was employed to model and analyze two COVID-19 datasets. Moreover, the Exponentiated Extended Extreme value (EEEV) distribution with three parameters were discussed in [5]. Due to the flexibility of the extended distributions, many studies have been proposed; for instance, the Exponentiated Weibull, New Exponentiated Beta, and Exponentiated Gamma Burr-type X distributions [6, 7, 8]. In the field of trigonometric-based distributions, Mohammed Elgarhy et al applied the Dinesh-Umesh-Sanjay (DUS) transformer to develop an extension of the classical Weibull distribution known as the tangent DUS Weibull distribution (TDUSW) [9]. Utilizing this transform, the three-parameter generalized Dinesh–Umesh–Sanjay generalized exponential (GDUS-GE) distribution was introduced [10].

Distributions arising from the maximum and minimum random variables with known probability distributions represent another approach that has been used in the literature to develop new or generalized distributions. Another efficient method to extend classical distributions is to study their extreme functions [11, 12, 13, 14]. These distributions were used to model the distribution of imperfect-quality items in inventory problems [15, 16].

In another context, the literature has employed mixtures of distributions as an efficient approach for developing novel and more applicable models with greater relevance to real-life data. Mixture distributions are believed to better express complex data than a single-type distribution, as they provide improved prediction for complicated data and capture heavy tails, multimodality, and heterogeneous failure modes more comprehensively. Also, the mixture of exponential distribution with discrete points is studied with its properties in [17]. A prior study presented a finite-mixture model by mixing the Exponential, Gamma, and Weibull distributions as an effective method to approximate the parameters of heterogeneous survival data [18]. Another example of mixed-type distributions is that proposed by [19], where the normal mean-variance mixture of Lindley (NMVL) distribution is studied and its parameters are estimated using the Expectation–Maximization (EM) algorithm. This method has a proven advantage over single-component distributions in the accuracy of estimating parameters in real-world data [20]. Due to the various applications of mixed distributions, recent studies have emerged [21, 22, 23, 24, 25, 26].

Although the aforementioned studies establish strong applicability, they are limited to either a single type of real word application domain, such as failure time behavior or inventory problems, or to multiple applications but still in the same domain. However, this article shows the versatility of the mixed DUS Lindley-Weibull distribution (DUSLW) by applying it to different types of applications, including medical sciences, finance and queuing systems.

This article is arranged as follows; in Section 2, the PDF, CDF, entropy, survival, and hazard rate functions are derived, and graphs are presented. In Section 3, some statistical properties such as moments, skewness, kurtosis, incomplete moments, and order statistics are discussed. In Section 4, the MLEs of the model are obtained using the EM algorithm. A simulation study is then presented in Section 5 to demonstrate the efficiency of this method. Section 6 is devoted to data application to estimate the parameters of the DUSLW model and to discuss of the results. In Section 7, the conclusions and discussions are presented. Finally, future work is displayed in Section 8.

2 Mixture of DUS Lindley and Weibull distribution

In this study, we introduce a mixture distribution called the **DUS Lindley-Weibull distribution (DUSLW)**. This distribution has two continuous components: the DUS Lindley distribution with weight β , and the Weibull distribution with weight $1 - \beta$. The DUS Lindley distribution is obtained by applying the Dinesh-Umesh-Sanjay transformation to the Lindley distribution [27], and its CDF and PDF are defined, respectively, as follows:

$$G(x; \xi) = \frac{1}{e-1} \left(\exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) - 1 \right), \quad (2.1)$$

$$g(x; \xi) = \frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right), \quad (2.2)$$

$$\text{for } x > 0, \xi > 0,$$

where, ξ is the rate parameter.

The CDF and PDF of the two-parameter Weibull distribution are illustrated below in terms of the shape parameter k and scale parameter λ :

$$H(x; k, \lambda) = 1 - e^{-(x/\lambda)^k}, \quad (2.3)$$

$$h(x; k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, \quad (2.4)$$

for $x \geq 0, k > 0, \lambda > 0$.

2.1 CDF and PDF of the DUSLW distribution

The cumulative distribution function for the mixed DUS Lindley–Weibull (DUSLW) distribution is:

$$F(x; \Theta) = \beta \frac{1}{e-1} \left(\exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) - 1 \right) + (1-\beta) \left(1 - e^{-(x/\lambda)^k} \right), \quad (2.5)$$

and the probability density function is:

$$f(x; \Theta) = \beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) + (1-\beta) \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, \quad (2.6)$$

where, $\Theta = (\beta, \xi, \lambda, k)$.

2.2 Rényi entropy of the DUSLW distribution

The Rényi entropy of the mixed distribution is defined as:

$$I(x) = \frac{1}{1-\rho} \log \int_0^\infty f(x)^\rho dx, \quad \rho > 0, \rho \neq 1, \quad (2.7)$$

then,

$$I(x) = \frac{1}{1-\rho} \log \int_0^\infty \underbrace{\left(\beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) \right)}_{A(x)} + \underbrace{\left((1-\beta) \frac{k}{\lambda} (x/\lambda)^{k-1} e^{-(x/\lambda)^k} \right)}_{B(x)} dx. \quad (2.8)$$

The domain of integration is decomposed into two subdomains:

$$S_1 = \{x \geq 0; B(x) \leq A(x)\}, \quad S_2 = \{x \geq 0; B(x) > A(x)\}.$$

therefore,

$$I(x) = \frac{1}{1-\rho} \log \left(\int_{S_1} A(x)^\rho \sum_{m=0}^\infty \binom{\rho}{m} \left(\frac{B(x)}{A(x)} \right)^m dx + \int_{S_2} B(x)^\rho \sum_{m=0}^\infty \binom{\rho}{m} \left(\frac{A(x)}{B(x)} \right)^m dx \right), \quad (2.9)$$

Using Fubini's theorem:

$$\begin{aligned} I(x) &= \frac{1}{1-\rho} \log \left(\sum_{m=0}^\infty \binom{\rho}{m} \int_{S_1} A(x)^\rho \left(\frac{B(x)}{A(x)} \right)^m dx + \sum_{m=0}^\infty \binom{\rho}{m} \int_{S_2} B(x)^\rho \left(\frac{A(x)}{B(x)} \right)^m dx \right), \\ &= \frac{1}{1-\rho} \log \sum_{m=0}^\infty \binom{\rho}{m} \left(\int_{S_1} A(x)^{\rho-m} B(x)^m dx + \int_{S_2} B(x)^{\rho-m} A(x)^m dx \right). \end{aligned} \quad (2.10)$$

By applying the series expansion to the exponential terms, the above integrals can be computed.

2.3 Survival and hazard rate functions

The survival and hazard rate functions of the DUSLW distribution are:

$$SF(x; \Theta) = 1 - \beta \frac{1}{e-1} \left(\exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) - 1 \right) - (1 - \beta) \left(1 - e^{-(x/\lambda)^k} \right), \quad (2.11)$$

$$h(x; \Theta) = \frac{\beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) + (1 - \beta) \frac{k}{\lambda} (x/\lambda)^{k-1} e^{-(x/\lambda)^k}}{1 - \beta \frac{1}{e-1} \left(\exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) - 1 \right) - (1 - \beta) \left(1 - e^{-(x/\lambda)^k} \right)}. \quad (2.12)$$

Initial Behavior of the Hazard Rate Function

The hazard rate function of the DUSLW distribution is $h(x; \Theta)$. Its behavior as $x \rightarrow 0$ is

$$\lim_{x \rightarrow 0} h(x; \Theta) = \frac{\beta \xi^2}{(e-1)(1+\xi)} + (1 - \beta) \frac{k}{\lambda} \lim_{x \rightarrow 0} \left(\frac{x}{\lambda} \right)^{k-1}. \quad (2.13)$$

Depending on the value of the Weibull shape parameter k , we have

$$\lim_{x \rightarrow 0} h(x; \Theta) = \begin{cases} \infty & 0 < k < 1 \\ \frac{\beta \xi^2}{(e-1)(1+\xi)} & k > 1, \\ \frac{\beta \xi^2}{(e-1)(1+\xi)} + (1 - \beta) \frac{k}{\lambda} & k = 1 \end{cases} \quad (2.14)$$

The behavior of the PDF of the DUSLW distribution for various parameter choices is investigated and shown in Figures 1 and 2. It is observed that changing the parameters yields a wide range of behaviors and shapes. For instance, as the mixing weight β increases, the PDF increasingly reflects the DUS-Lindley component, whereas as β decreases, its behavior approaches that of the Weibull distribution. This behavior can also be seen in Table 1, where the DUS-Lindley, Weibull, and exponential distributions are special cases of the DUSLW distribution.

In addition, increasing the values of k and ξ results in a higher peak. Moreover, the PDF curve becomes flatter as the Weibull scale parameter increases. On the other hand, the CDF is shown in Figure 3, and the distinct shapes of the hazard rate function corresponding to different parameter settings are also illustrated.

Table 1. Special cases of the DUSLW distribution.

Distribution	PDF	CDF
DUS-Lindley ($\beta = 1$)	$\frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right)$	$\frac{1}{e-1} \left(\exp \left(1 - \left(1 + \frac{\xi x}{1+\xi} \right) e^{-\xi x} \right) - 1 \right)$
Weibull ($\beta = 0$)	$\frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(\frac{x}{\lambda})^k}$	$1 - e^{-(\frac{x}{\lambda})^k}$
Exponential ($\beta = 0, k = 1$)	$\frac{1}{\lambda} e^{-x/\lambda}$	$1 - e^{-x/\lambda}$

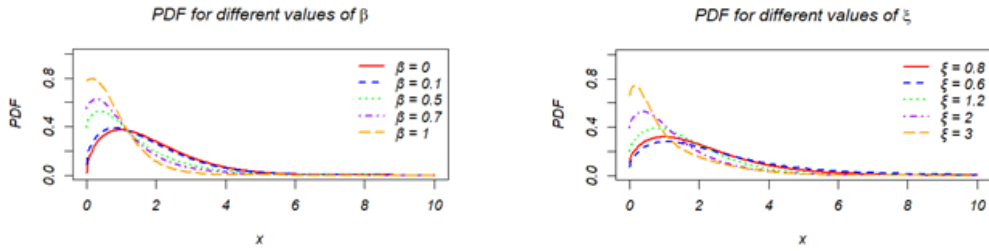


Figure 1. PDF of the mixture of DUS Lindley and Weibull distributions for different values of β and ξ .

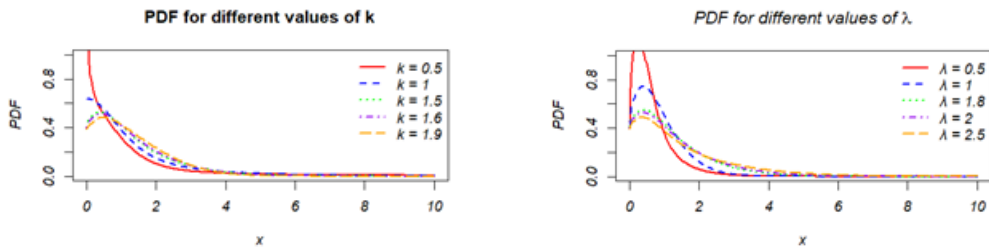


Figure 2. PDF of the mixture of DUS Lindley and Weibull distributions for different values of k and λ .

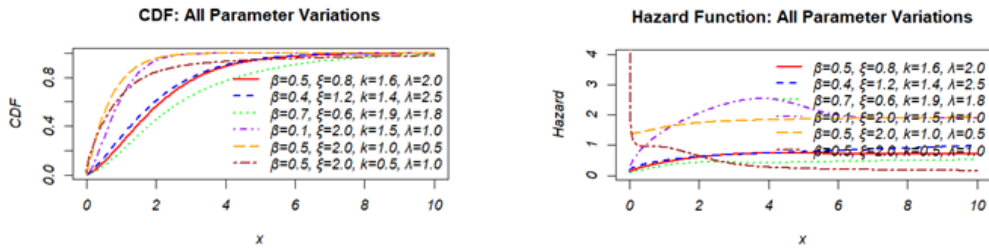


Figure 3. CDF and hazard rate of the mixture of DUS Lindley and Weibull distributions for different parameters.

3 Statistical Properties

3.1 Ordinary Moments

Using (2.6), the r th moment of the DUSLW distribution can be determined as follows:

$$\begin{aligned}
 \mu'_r &= \int_{-\infty}^{+\infty} x^r f(x) dx \\
 &= \int_0^{+\infty} x^r \left[\beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \left(e^{1-(1+\frac{\xi x}{1+\xi})} e^{-\xi x} \right) \right. \\
 &\quad \left. + (1-\beta) \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k} \right] dx \\
 &= \beta \mu'_{1,r} + (1-\beta) \mu'_{2,r}.
 \end{aligned} \tag{3.1}$$

where, $\mu'_{1,r}$ and $\mu'_{2,r}$ are the r th moment functions of the DUS Lindley [27] and Weibull distributions, respectively, given by:

$$\mu'_{1,r} = \sum_{l=0}^{\infty} \sum_{m=0}^l \sum_{n=0}^m \frac{(-1)^l \binom{l}{m} \binom{m}{n} \Gamma(n+r+1) ((m+1)\xi + (n+r+1))}{l!(e-1)\xi^r(1+\xi)^{n+1}} \frac{1}{(m+1)^{n+r+2}}, \quad (3.2)$$

$$\mu'_{2,r} = \lambda^r \Gamma\left(1 + \frac{r}{k}\right). \quad (3.3)$$

Thus, the r th moment of the DUSLW distribution can be written as:

$$\begin{aligned} \mu'_r &= \beta \sum_{l=0}^{\infty} \sum_{m=0}^l \sum_{n=0}^m \frac{(-1)^l \binom{l}{m} \binom{m}{n} \Gamma(n+r+1) ((m+1)\xi + (n+r+1))}{l!(e-1)\xi^r(1+\xi)^{n+1}} \frac{1}{(m+1)^{n+r+2}} \\ &\quad + (1-\beta) \lambda^r \Gamma\left(1 + \frac{r}{k}\right). \end{aligned} \quad (3.4)$$

The first four moments about the origin are obtained by substituting $r = 1, 2, 3, 4$ in (3.4). The numerical values of the first four moments of the DUSLW distribution for various parameter settings are estimated using the R program, where the infinite summation is truncated at a sufficiently large value. Approximate values of the standard deviation, skewness, and kurtosis are then obtained using formulas (3.9), (3.10), and (3.11). The results are presented in Table 2, where one parameter is varied at a time while the others are held fixed. As expected, the coefficients of skewness and kurtosis are affected by changes in the parameters β , ξ , and k . However, unlike the Weibull distribution, variation in the parameter λ also affects skewness and kurtosis, showing an inverse relationship. The standard deviation, skewness and kurtosis are given in terms of the four moments such that:

$$\mu'_1 = \beta \sum_{l=0}^{\infty} \sum_{m=0}^l \sum_{n=0}^m \frac{(-1)^l \binom{l}{m} \binom{m}{n} \Gamma(n+2) ((m+1)\xi + n+2)}{l!(e-1)\xi(1+\xi)^{n+1}} \frac{1}{(m+1)^{n+3}} + (1-\beta) \lambda \Gamma\left(1 + \frac{1}{k}\right), \quad (3.5)$$

$$\mu'_2 = \beta \sum_{l=0}^{\infty} \sum_{m=0}^l \sum_{n=0}^m \frac{(-1)^l \binom{l}{m} \binom{m}{n} \Gamma(n+3) ((m+1)\xi + n+3)}{l!(e-1)\xi^2(1+\xi)^{n+1}} \frac{1}{(m+1)^{n+4}} + (1-\beta) \lambda^2 \Gamma\left(1 + \frac{2}{k}\right), \quad (3.6)$$

$$\mu'_3 = \beta \sum_{l=0}^{\infty} \sum_{m=0}^l \sum_{n=0}^m \frac{(-1)^l \binom{l}{m} \binom{m}{n} \Gamma(n+4) ((m+1)\xi + n+4)}{l!(e-1)\xi^3(1+\xi)^{n+1}} \frac{1}{(m+1)^{n+5}} + (1-\beta) \lambda^3 \Gamma\left(1 + \frac{3}{k}\right), \quad (3.7)$$

$$\mu'_4 = \beta \sum_{l=0}^{\infty} \sum_{m=0}^l \sum_{n=0}^m \frac{(-1)^l \binom{l}{m} \binom{m}{n} \Gamma(n+5) ((m+1)\xi + n+5)}{l!(e-1)\xi^4(1+\xi)^{n+1}} \frac{1}{(m+1)^{n+6}} + (1-\beta) \lambda^4 \Gamma\left(1 + \frac{4}{k}\right). \quad (3.8)$$

$$\sigma = \sqrt{\mu'_2 - (\mu'_1)^2}, \quad (3.9)$$

$$SK = \frac{\mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3}{\sigma^3}, \quad (3.10)$$

$$KU = \frac{\mu'_4 - 4\mu'_1\mu'_3 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4}{\sigma^4}. \quad (3.11)$$

To illustrate the computations, the statistical properties of the DUSLW distribution obtained theoretically are evaluated for the parameter set ($\beta = 0.10$, $\xi = 2.00$, $k = 1.50$, $\lambda = 1.00$) in Table 3. The theoretical results for the mean, standard deviation, and skewness for three different parameter sets are compared with the averages obtained from simulating 10,000 datasets following the DUSLW distribution. The biases are also calculated, and it is observed that the simulated averages of the DUSLW distribution are in close agreement with the theoretical results, which provides validation for the theoretical formulas.

Table 2. Numerical values for the first four moments, standard deviations, skewness, and kurtosis of the DUSLW distribution.

β	ξ	k	λ	μ'_1	μ'_2	μ'_3	μ'_4	σ	SK	KU
0.10	2.00	1.50	1.00	0.824	1.087	1.830	3.683	0.639	1.005	4.184
0.30	2.00	1.50	1.00	0.666	0.881	1.490	3.025	0.661	1.108	4.233
0.50	2.00	1.50	1.00	0.508	0.674	1.150	2.367	0.645	1.435	5.052
0.70	2.00	1.50	1.00	0.350	0.467	0.810	1.709	0.587	2.005	7.358
0.90	2.00	1.50	1.00	0.192	0.261	0.470	1.051	0.473	3.146	14.803
0.50	0.50	1.50	1.00	0.727	2.320	15.177	145.778	1.338	4.540	33.702
0.50	1.00	1.50	1.00	0.577	0.969	2.489	9.398	0.798	2.357	12.992
0.50	2.00	1.50	1.00	0.508	0.674	1.150	2.367	0.645	1.435	5.052
0.50	3.00	1.50	1.00	0.487	0.627	1.039	2.067	0.624	1.454	5.044
0.50	5.00	1.50	1.00	0.471	0.605	1.007	2.013	0.619	1.523	5.268
0.50	2.00	1.50	1.00	0.508	0.674	1.150	2.367	0.645	1.435	5.052
0.50	2.00	2.00	1.00	0.499	0.579	0.815	1.361	0.574	1.038	3.795
0.50	2.00	3.00	1.00	0.503	0.530	0.650	0.956	0.526	0.719	3.398
0.50	2.00	1.50	0.50	0.282	0.227	0.275	0.487	0.384	2.257	12.237
0.50	2.00	1.50	1.00	0.508	0.674	1.150	2.367	0.645	1.435	5.052
0.50	2.00	1.50	1.50	0.733	1.418	3.525	10.517	0.938	1.445	5.011
0.50	2.00	1.50	2.00	0.959	2.460	8.150	32.459	1.241	1.484	5.156
0.50	2.00	1.50	3.00	1.410	5.436	27.150	162.855	1.857	1.525	5.276

Table 3. Comparison of theoretical and simulated statistics for different parameter sets

$\beta = 0.10, \xi = 2.00, k = 1.50, \lambda = 1.00$	
Theoretical mean	0.8240
Average mean	0.8701
Bias	0.0461
Theoretical standard deviation	0.6390
Average standard deviation	0.6113
Bias	-0.0277
Theoretical skewness	1.0050
Average skewness	1.0845
Bias	0.0795

3.2 Incomplete moments

The incomplete moments are defined as follows:

$$m_v(t) = \int_0^t x^v f(x) dx. \quad (3.12)$$

By utilizing the pdf of the DUSLW distribution in (2.6), we get:

$$\begin{aligned}
 m_v(t) = & \beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} \left(\int_0^t x^v e^{-\xi x} \left(e^{1-(1+\frac{\xi x}{1+\xi}) e^{-\xi x}} \right) dx + \int_0^t x^{v+1} e^{-\xi x} \left(e^{1-(1+\frac{\xi x}{1+\xi}) e^{-\xi x}} \right) dx \right) \\
 & + (1-\beta) \frac{k}{\lambda^k} \int_0^t x^{v+k-1} e^{-(x/\lambda)^k} dx.
 \end{aligned} \quad (3.13)$$

After applying the Taylor series expansion on the exponential term and the finite binomial series, we get:

$$\begin{aligned}
 m_v(t) = & \beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} \sum_{n=0}^{\infty} \sum_{j=0}^n (-1)^n \binom{n}{j} \frac{e}{n!} \left(\frac{\xi}{1+\xi} \right)^j \left(\int_0^t x^{v+j} e^{-\xi(n+1)x} dx \right. \\
 & \left. + \int_0^t x^{v+j+1} e^{-\xi(n+1)x} dx \right) \\
 & + (1-\beta) \frac{k}{\lambda^k} \int_0^t x^{v+k-1} e^{-(x/\lambda)^k} dx.
 \end{aligned} \quad (3.14)$$

The above integrals can be solved by substitution and are expressed in terms of incomplete gamma functions as:

$$m_v(t) = \beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} S(\xi, r, n, m) \left(\frac{\gamma_1}{(\xi(n+1))^{v+j+1}} + \frac{\gamma_2}{(\xi(n+1))^{v+j+2}} \right) + (1-\beta) \lambda^v \gamma_3, \quad (3.15)$$

where

$$S(\xi, r, n) = \sum_{n=0}^{\infty} \sum_{j=0}^n (-1)^n \binom{n}{j} \frac{e}{n!} \left(\frac{\xi}{1+\xi} \right)^j, \quad (3.16)$$

$$\gamma_1 = \Gamma[v+j+1] - \Gamma[v+j+1, \xi(n+1)t], \quad (3.17)$$

$$\gamma_2 = \Gamma[v+n+2] - \Gamma[v+j+2, \xi(n+1)t], \quad (3.18)$$

$$\gamma_3 = \Gamma\left[\frac{v+k}{k}\right] - \Gamma\left[\frac{v+k}{k}, (t/\lambda)^k\right]. \quad (3.19)$$

3.3 Order Statistics

Let $X_1, X_2, \dots, X_n$ be a random sample from the DUSLW distribution, and consider $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ the corresponding order statistics. The probability density function of the s th order statistic is given by:

$$f_{s:n}(x) = \frac{f(x)}{B(s, n-s+1)} (F(x))^{s-1} (1-F(x))^{n-s}, \quad (3.20)$$

where

$$B(s, n-s+1) = \frac{\Gamma(s) \cdot \Gamma(n-s+1)}{\Gamma(n+1)} \quad (3.21)$$

is the beta function. Using the binomial expansion of $(1-F(x))^{n-s}$ in (3.20) gives:

$$f_{s:n}(x) = \frac{n!}{(s-1)!(n-s)!} \sum_{w=0}^{n-s} (-1)^w \binom{n-s}{w} f(x) (F(x))^{s+w-1}. \quad (3.22)$$

Applying further expansions on $(F(x))^{s+w-1}$, we obtain:

$$(F(x))^{s+w-1} = \sum_{v=0}^{s+w-1} \sum_{l=0}^{\infty} \sum_{q=0}^{\infty} \sum_{m=0}^q \sum_{j=0}^v C(\Theta, j, l, m, q, s, w, v) x^m e^{-\xi q x - l/\lambda^k x^k}, \quad (3.23)$$

where

$$\begin{aligned}
 & C(\Theta, j, l, m, q, s, w, v) \\
 & = \frac{j^q e^j}{q!} (-1)^{(v-j)+l+q} \left(\frac{\xi}{1+\xi} \right)^m \binom{v}{j} \binom{q}{m} \binom{s+w-v-1}{l} \binom{s+w-1}{v} \\
 & \cdot \left(\frac{\beta}{e-1} \right)^v (1-\beta)^{s+w-v-1}.
 \end{aligned} \quad (3.24)$$

Substituting (3.23) into (3.22), the probability density function of the s th order statistic of the DUSLW distribution is obtained as:

$$f_{s:n}(x) = \frac{n!}{(s-1)!(n-s)!} \sum_{w=0}^{n-s} \sum_{v=0}^{s+w-1} \sum_{l=0}^{\infty} \sum_{q=0}^{\infty} \sum_{m=0}^q \sum_{j=0}^v C(\Theta, j, l, m, q, s, w, v) (-1)^w \binom{n-s}{w} \\ \cdot \left(\beta \frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+x) e^{-\xi x} \left(e^{1-(1+\frac{\xi x}{1+\xi}) e^{-\xi x}} \right) \right. \\ \left. + (1-\beta) \frac{k}{\lambda} (x/\lambda)^{k-1} e^{-(x/\lambda)^k} \right) x^m e^{-\xi q x - l x^k / \lambda^k}. \quad (3.25)$$

4 Parameter estimation for Mixed Lindley-Weibull distribution

4.1 MLE with EM algorithm

Estimation-Maximization (EM) Algorithm, that was first proposed by Dempster et al in 1977 [28], remains to be one of the most robust and valid methods used to evaluate the maximum likelihood estimators of finite mixture models. This method, which primarily aims to solve the incomplete Log-likelihood problem, consists of two steps: the Expectation (E-step) and the Maximization (M-step). In the E-step, the data is considered to be complete, and then an expected value of the complete data log-likelihood is taken. In the subsequent M-step, the log-likelihood is maximized with respect to the parameters, yielding new parameter estimates. Afterwards, alternation between both steps is done repeatedly until the variation in parameters becomes almost negligible. In this approach, it is crucial to construct a complete Log-likelihood function, denoted by $L_c(\Theta | t)$, by introducing the set of missing observations m_1 and m_2 and considering the added constants to be negligible.

$$m_{ki} = m_k(t_i) = \begin{cases} 1, & \text{if observation belongs to } k, \\ 0, & \text{otherwise,} \end{cases} \quad k = 1, 2 \text{ and } i = 1, 2, 3, \dots, n. \quad (4.1)$$

and

$$L_c(\Theta | t, m) = \sum_{i=1}^n m_{1i} \log \left(\frac{1}{e-1} \frac{\xi^2}{1+\xi} (1+t_i) e^{-\xi t_i} \exp \left(1 - \left(1 + \frac{\xi t_i}{1+\xi} \right) e^{-\xi t_i} \right) \right) \\ + \sum_{i=1}^n m_{2i} \log \left(\frac{k}{\lambda} \left(\frac{t_i}{\lambda} \right)^{k-1} e^{-(t_i/\lambda)^k} \right). \quad (4.2)$$

E-Step

The hidden variable vectors $m_i = [m_{1i}, m_{2i}]$ are estimated by evaluating the conditional expectation $E(m_{ki} | t_i)$. The conditional expectations $E(m_{1i} | t_i)$ and $E(m_{2i} | t_i)$ are calculated in the E-step and maximized in the maximization step. These expected variables are necessary conditions to compute the expected complete data Log-likelihood “Q-function” where:

$$\hat{m}_{1i}^{(z)} = E(m_{1i} | t_i) \quad (4.3)$$

$$= \frac{\beta^{(z)} \frac{1}{e-1} \frac{(\xi^{(z)})^2}{1+\xi^{(z)}} (1+t_i) e^{-\xi^{(z)} t_i} \left(e^{1-(1+\frac{\xi^{(z)} t_i}{1+\xi^{(z)}}) e^{-\xi^{(z)} t_i}} \right)}{\beta^{(z)} \frac{1}{e-1} \frac{(\xi^{(z)})^2}{1+\xi^{(z)}} (1+t_i) e^{-\xi^{(z)} t_i} \left(e^{1-(1+\frac{\xi^{(z)} t_i}{1+\xi^{(z)}}) e^{-\xi^{(z)} t_i}} \right) + (1-\beta^{(z)}) \frac{k^{(z)}}{\lambda^{(z)}} \left(\frac{t_i}{\lambda^{(z)}} \right)^{k^{(z)}-1} e^{-(t_i/\lambda^{(z)})^{k^{(z)}}}}, \quad (4.4)$$

$$\hat{m}_{2i}^{(z)} = E(m_{2i} | t_i) \quad (4.5)$$

$$= \frac{(1 - \beta^{(z)})^{\frac{k^{(z)}}{\lambda^{(z)}}} \left(\frac{t_i}{\lambda^{(z)}}\right)^{k^{(z)}-1} e^{-(t_i/\lambda^{(z)})^{k^{(z)}}}}{\beta^{(z)} \frac{1}{e-1} \frac{(\xi^{(z)})^2}{1+\xi^{(z)}} (1+t_i) e^{-\xi^{(z)} t_i} \left(e^{1-(1+\frac{\xi^{(z)} t_i}{1+\xi^{(z)}}) e^{-\xi^{(z)} t_i}} \right) + (1 - \beta^{(z)})^{\frac{k^{(z)}}{\lambda^{(z)}}} \left(\frac{t_i}{\lambda^{(z)}}\right)^{k^{(z)}-1} e^{-(t_i/\lambda^{(z)})^{k^{(z)}}}}. \quad (4.6)$$

The Q-function at the z th iteration can be expressed as:

$$\begin{aligned} Q(\Theta | \hat{\Theta}^{(z)}) &= \sum_{i=1}^n \hat{m}_{1i}^{(z)} \left(-\log(e-1) + 2\log \xi - \log(1+\xi) + \log(1+t_i) \right. \\ &\quad \left. - \xi t_i + 1 - \left(1 + \frac{\xi t_i}{1+\xi}\right) e^{-\xi t_i} \right) \\ &\quad + \sum_{i=1}^n \hat{m}_{2i}^{(z)} \left(\log k - k \log \lambda + (k-1) \log t_i - (t_i/\lambda)^k \right). \end{aligned} \quad (4.7)$$

M-Step

In order to update the parameters at $(z+1)$ iteration, we set the partial derivatives of the Q-function with respect to each of the parameters equal to zero. However, $\hat{\xi}^{(z+1)}$ and $\hat{k}^{(z+1)}$ can be obtained numerically by applying the Newton-Raphson algorithm to their non-linear equations, following similar approaches in the literature, where it has been shown to provide accurate parameter estimates compared to other mixed distributions [29].

$$\hat{\beta}^{(z+1)} = \frac{1}{n} \sum_{i=1}^n \hat{m}_{1i}^{(z)}. \quad (4.8)$$

$$\hat{\lambda}^{(z+1)} = \left(\frac{\sum_{i=1}^n \hat{m}_{1i}^{(z)} t_i^{k^{(z)}}}{\sum_{i=1}^n \hat{m}_{1i}^{(z)}} \right)^{\frac{1}{k^{(z)}}}. \quad (4.9)$$

$$\frac{\sum_{i=1}^n 2\hat{m}_{1i}^{(z)}}{\xi} - \frac{\sum_{i=1}^n \hat{m}_{1i}^{(z)}}{1+\xi} - \sum_{i=1}^n \hat{m}_{1i}^{(z)} t_i - \sum_{i=1}^n \hat{m}_{1i}^{(z)} \frac{t_i}{(1+\xi)^2} e^{-\xi t_i} + \sum_{i=1}^n \hat{m}_{1i}^{(z)} t_i \left(1 + \frac{\xi t_i}{1+\xi}\right) e^{-\xi t_i} = 0. \quad (4.10)$$

$$\frac{\sum_{i=1}^n \hat{m}_{2i}^{(z)}}{k} - \sum_{i=1}^n \hat{m}_{2i}^{(z)} \log \lambda + \sum_{i=1}^n \hat{m}_{2i}^{(z)} \log t_i - \sum_{i=1}^n \hat{m}_{2i}^{(z)} \left(\frac{t_i}{\lambda}\right)^k \log \left(\frac{t_i}{\lambda}\right) = 0. \quad (4.11)$$

5 Simulation

In this part of the study, the performance of the DUSLW distribution is evaluated by the means of simulation studies, applied on different sample sizes ($n = 25, 50, 100, 500$), as well as on diverse sets of parameters ($(\beta, \xi, k, \lambda) = \{(0.55, 0.80, 1.60, 2.00), (0.45, 1.20, 1.40, 2.50), (0.70, 0.60, 1.90, 1.80)\}$). Each observation is created by first determining which distribution it belongs to, using the Bernoulli β . After that, sampling from each variable is done using the appropriate method depending on the distribution type, where the R Weibull (inverse-CDF) is used in the Weibull distribution and the rejection sampling scheme for the DUS Lindley distribution. Simulation studies are performed 100 times for every sample, where the average estimated parameters for each is deduced. The root mean square error (RMSE) and the bias of all estimations are derived, as shown in Table 4. The three sets of parameters studied here can be observed in Table 4, where it is noticeable that the average values and the real values are strongly comparable, with the difference becoming less significant with the increase in the sample size. Correspondingly, the bias reaches closer to zero, and the RMSE decreases, following the increase in the

sample size. Therefore, a satisfactory performance for the maximum likelihood estimation in studying these sample sizes is noted, and the maximum likelihood estimation with EM algorithm is proven to be appropriate for estimating DUSLW distribution’s parameters, being favorable for a larger sample size.

Table 4. Simulation study: DUSLW parameter estimates, RMSE, and Bias for three parameter sets and sample sizes.

<i>n</i>	<i>β</i>			<i>ξ</i>			<i>k</i>			<i>λ</i>		
	Avg	RMSE	Bias	Avg	RMSE	Bias	Avg	RMSE	Bias	Avg	RMSE	Bias
Set 1: <i>β</i> = 0.55, <i>ξ</i> = 0.80, <i>k</i> = 1.60, <i>λ</i> = 2.00												
25	0.4648	0.3514	-0.0852	1.5400	1.7870	0.7400	5.4083	6.8490	3.8083	3.0080	2.4608	1.0080
50	0.5170	0.3161	-0.0330	1.0702	0.8260	0.2702	4.6460	5.4828	3.0460	2.2894	1.2585	0.2894
100	0.5054	0.3004	-0.0446	1.0031	0.5372	0.2031	3.0013	2.6163	1.4013	2.2040	0.7705	0.2040
500	0.5402	0.2503	-0.0098	0.8791	0.2302	0.0791	2.1196	1.3888	2.1196	2.1612	0.4449	0.1612
Set 2: <i>β</i> = 0.45, <i>ξ</i> = 1.20, <i>k</i> = 1.40, <i>λ</i> = 2.50												
25	0.5966	0.3401	0.1466	1.3589	1.0235	0.1589	7.0909	8.5731	5.6909	2.9209	1.9895	0.4209
50	0.6088	0.3271	0.1588	1.0989	0.6928	-0.1011	4.1836	5.0411	2.7836	2.5058	1.9403	0.0058
100	0.5836	0.3268	0.1336	1.3809	1.0320	0.1809	3.3777	3.9394	1.9777	2.5197	1.4235	0.0197
500	0.5333	0.2689	0.0833	1.1874	0.4592	-0.0126	1.7058	0.8437	0.3058	2.1934	0.7799	-0.3066
Set 3: <i>β</i> = 0.70, <i>ξ</i> = 0.60, <i>k</i> = 1.90, <i>λ</i> = 1.80												
25	0.5873	0.3027	-0.1127	1.0540	2.6866	0.4540	6.6304	7.6242	4.7304	3.5169	3.6196	1.7169
50	0.6079	0.3072	-0.0921	0.9408	1.7204	0.3408	4.6778	5.0913	2.7778	3.1452	2.8892	1.3452
100	0.6162	0.2831	-0.0838	0.9639	2.7678	0.3639	3.3458	2.9851	1.4458	2.7741	2.7574	0.9741
500	0.6291	0.1856	-0.0709	0.6290	0.1448	0.0290	2.0787	0.7578	0.1787	2.0282	0.7029	0.2282

6 Applications

To assess the efficacy of the DUSLW distribution, it was applied to four real-world datasets: 1, 2, 3, and 4, as shown below. Dataset 1 represents the survival time of patients diagnosed with leukemia in a hospital located in Saudi Arabia [30]. Dataset 2 depicts the time of remission in patients with bladder cancer [30]. Dataset 3 illustrates the number of shares traded per week, within the period between June 11, 2021, and June 11, 2024, in the BTC-USD market [9]. Dataset 4 includes the waiting times (in minutes) before serving 100 bank customers [31].

The results attained by means of the DUSLW distribution are evaluated relative to the following distributions: DUS Lindley, Lindley, Weibull, Gamma, and TDUSW.

The Lindley distribution has the PDF given by

$$f(x; \xi) = \frac{\xi^2}{1 + \xi} (1 + x)e^{-\xi x}, \quad x > 0, \xi > 0. \tag{6.1}$$

The Gamma distribution has the PDF

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0. \tag{6.2}$$

The TDUSW distribution has the PDF

$$f(x; \alpha, \beta) = \frac{\pi \alpha \beta}{4(e - 1)} x^{\alpha-1} e^{1-\beta x^\alpha - e^{-\beta x^\alpha}} \sec^2 \left(\frac{\pi}{4} \frac{e^{1-e^{-\beta x^\alpha}} - 1}{e - 1} \right), \quad x, \alpha, \beta > 0. \tag{6.3}$$

The DUS Lindley and Weibull distributions have the PDFs shown in (2.2) and (2.4), respectively.

A number of statistical tools can be used to evaluate competing probability models for a given dataset in order to identify the most appropriate fit. In this study, model fit assessment is mainly conducted using the Akaike Information Criterion (AIC), the Kolmogorov–Smirnov (KS) statistic, the Bayesian Information Criterion (BIC), the Hannan–Quinn Information Criterion (HQIC), along with the corresponding p-value. A model is considered superior if it yields lower AIC, KS,

BIC, and HQIC values or a higher KS p-value.

To further validate the results, the empirical cumulative distribution function (CDF) is compared with the estimated CDFs of the proposed models. Moreover, the histogram derived from the observed frequencies is examined relative to the expected frequencies from each fitted distribution. In addition, the quantile–quantile (Q–Q) plot, one of the most significant visual representations of model fitting, is also obtained for each dataset, with data points falling along the straight line with minimal deviation indicating a better fit. Also, The Survival and hazard functions were plotted, and they showed an overall satisfactory fit in the four cases.

6.1 Dataset 1: Blood cancer data

The first dataset represents the survival times (in months) of patients with leukemia, collected from a hospital in Saudi Arabia [30]. The observed values are as follows:

0.315	0.496	0.616	1.145	1.208	1.263	1.414	2.025	2.036	2.162
2.211	2.370	2.532	2.693	2.805	2.910	2.912	3.192	3.263	3.348
3.348	3.427	3.499	3.534	3.751	3.767	3.858	3.986	4.049	4.244
4.323	4.381	4.392	4.397	4.647	4.753	4.929	4.973	5.074	5.381

This dataset, collected from a hospital in Saudi Arabia, consists of the survival times (in years) of 40 patients detected with leukemia. It represents the duration from diagnosis until the occurrence of the last follow-up or death. Survival datasets are commonly studied in reliability studies and medical statistics, as they provide insight into disease development and patient diagnosis. The observations are continuous and positive, and they show inconsistency and skewness due to differences in patient situations and treatments. Thus, this dataset is suitable for evaluating the performance of lifetime distributions and assessing their ability to model real-world survival behavior. The maximum likelihood estimates (MLEs) of the model parameters, together with the corresponding goodness-of-fit measures for this dataset and their competing distributions, are summarized in Table 5. The results show lower AIC, KS, BIC, and HQIC values and a noticeable higher p-value of the DUSLW distribution relative to the other distributions. Additionally, Figures 4, 5, 6, and 7 demonstrate the adequacy of the proposed DUSLW distribution due to its close fit to the given blood cancer dataset.

Table 5. MLEs for the parameters and associated goodness of fit statistics for the blood cancer data.

Distributions	LL	MLE	AIC	HQIC	BIC	KS-Test	P-value
DUSLW	-65.6938	$\hat{\beta} = 0.2566$	139.3875	141.8301	146.1431	0.0542	0.9998
		$\hat{\xi} = 1.1541$					
		$\hat{k} = 4.5052$					
		$\hat{\lambda} = 4.0324$					
DUS Lindley	-77.2856	$\hat{\xi} = 0.6459$	156.5713	157.1819	158.2602	0.2168	0.0465
Lindley	-80.2506	$\hat{\xi} = 0.5269$	162.5012	163.1118	164.1900	0.2406	0.0195
Weibull	-69.5580	$\hat{k} = 2.4995$	143.1159	144.3372	146.4937	0.1184	0.6291
		$\hat{\lambda} = 3.5185$					
Gamma	-73.5487	$\hat{\alpha} = 3.4648$	151.0974	152.3187	154.4752	0.1582	0.2693
		$\hat{\gamma} = 1.1032$					
TDUSW	-68.1676	$\hat{\alpha} = 4.5052$	140.3352	141.5565	143.7130	0.0930	0.8793
		$\hat{\beta} = 4.0324$					

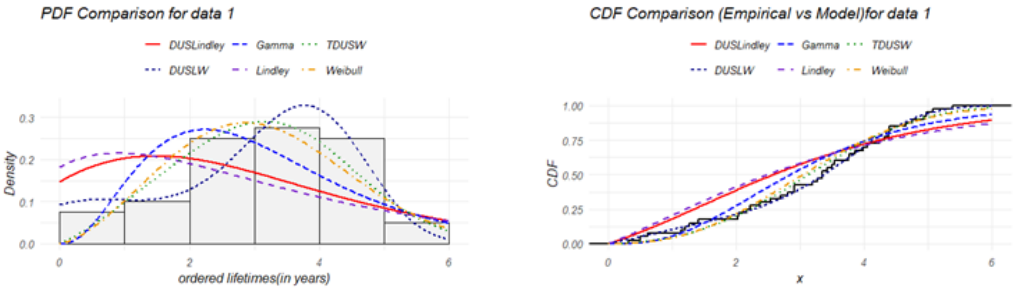


Figure 4. Theoretical and empirical PDF and CDF comparison of the blood cancer data.

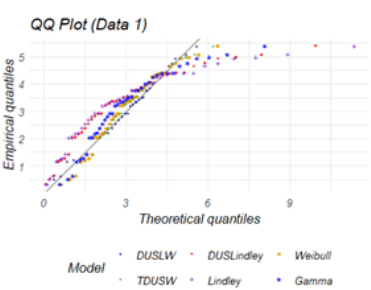


Figure 5. QQ plot comparison of dataset 1.

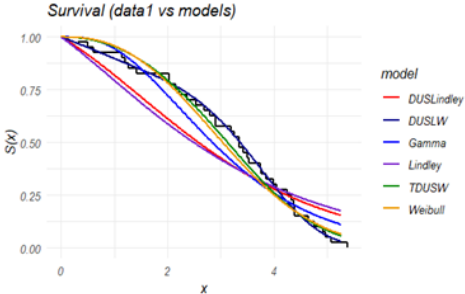


Figure 6. Survival plot comparison of dataset 1.

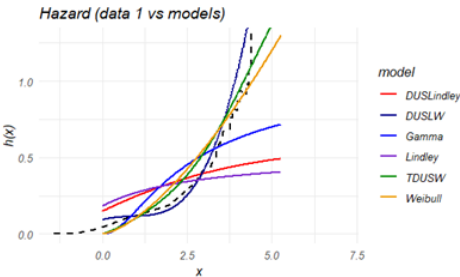


Figure 7. Hazard plot comparison of dataset 1.

6.2 Dataset 2: Bladder cancer data

0.08	0.20	0.40	0.50	0.51	0.81	0.87	0.90	1.05	1.19
1.26	1.35	1.40	1.46	1.76	2.02	2.02	2.07	2.09	2.23
2.26	2.46	2.54	2.62	2.64	2.69	2.69	2.75	2.83	2.87
3.02	3.02	3.25	3.31	3.36	3.36				

It represents the remission times of bladder cancer of 36 patients. Due to the heterogeneity in patient responses, the bladder cancer dataset is appropriate for mixed distributions in particular the DUSLW distribution. The proposed distribution provides a flexible fit to this data, it captures both the early reduction behavior and the heavier tail. Table 6 shows that the values of AIC, KS, BIC, and HQIC are low comparing with other distributions with high p-value. Also, the close-fitting curves of the probability and cumulative density functions along with the QQ plots can be seen in Figures 8 and 9. Moreover, to ensure the validity of this model, survival and hazard rates are also plotted in Figures 10 and 11, where both curves display a close agreement between the fitted and empirical functions. These results show that DUSLW distribution not only provides a close fit to the empirical distribution functions, but also represents medical data more accurately than the alternative distributions.

Table 6. MLEs for the parameters and associated goodness of fit statistics for the bladder cancer data.

Distributions	LL	MLE	AIC	HQIC	BIC	KS-Test	P-value
		$\hat{\beta} = 0.4639$					
DUSLW	-45.3077	$\hat{\xi} = 1.5513$ $\hat{k} = 6.3392$ $\hat{\lambda} = 2.8398$	98.6153	100.8261	104.9494	0.0712	0.9932
DUSL	-54.4870	$\hat{\xi} = 0.9768$	110.9740	111.5267	112.5575	0.1997	0.1133
Lindley	-56.6988	$\hat{\xi} = 0.8016$	115.3976	115.9503	116.9811	0.2073	0.0907
Weibull	-51.3870	$\hat{k} = 1.9570$ $\hat{\lambda} = 2.1645$	106.7741	107.8795	109.9411	0.1659	0.2752
Gamma	-54.0924	$\hat{\alpha} = 2.3096$ $\hat{\gamma} = 1.1905$	112.1848	113.2902	115.3518	0.1945	0.1314
TDUSW	-50.0027	$\hat{\alpha} = 1.7245$ $\hat{\beta} = 0.4074$	104.0054	105.1108	107.1725	0.1381	0.4981

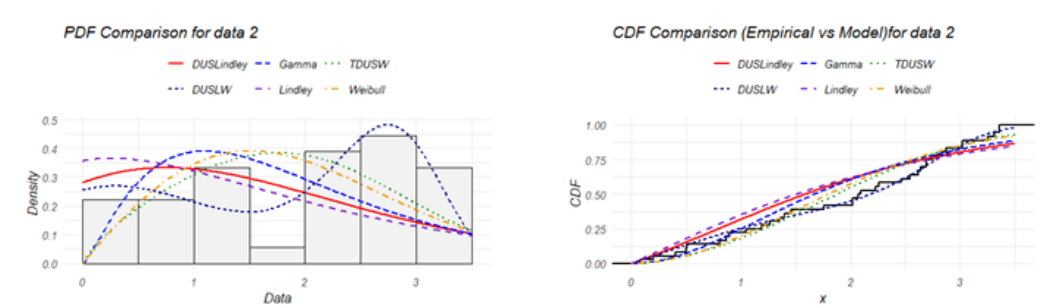


Figure 8. Theoretical and empirical PDF and CDF comparison of the bladder cancer data.

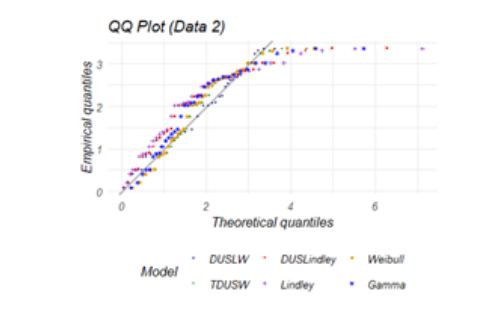


Figure 9. QQ plot comparison of dataset 2.

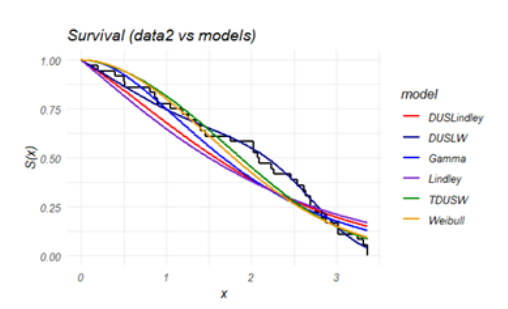


Figure 10. Survival plot comparison of dataset 2.

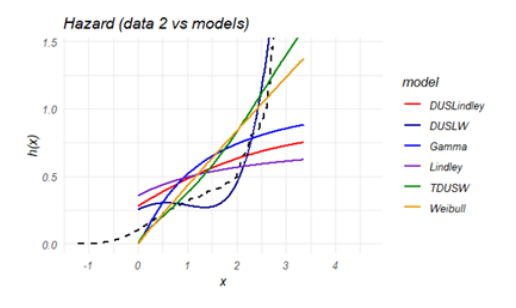


Figure 11. Hazard plot comparison of dataset 2.

6.3 Dataset 3: Total weekly traded volume of BTC-USD.

1.17	2.70	3.06	2.32	1.78	1.48	1.56	2.38	2.27	2.34
2.36	2.24	2.56	2.93	2.26	2.70	2.23	2.62	2.77	2.58
2.56	2.27	2.60	2.53	2.44	2.77	2.08	2.15	1.75	2.04
2.76	1.99	1.96	1.79	1.44	1.77	1.45	1.90	1.88	1.76
1.84	1.89	2.23	1.89	1.75	1.84	2.17	2.39	3.58	1.96
2.13	2.03	2.16	3.10	1.76	1.57	2.07	1.87	2.44	2.24
1.66	1.95	2.09	2.23	2.09	2.60	2.62	2.53	2.80	1.97
1.98	1.76	3.02	3.21	4.71	2.12	1.84	1.56	1.29	1.42
1.08	0.955	0.900	1.75	1.89	1.77	1.71	1.70	2.09	1.80
1.39	1.92	3.06	1.92	1.21	0.981	1.23	1.30	1.37	1.15
1.10	0.845	0.901	0.914	1.16	0.929	1.38	1.07	0.905	1.03
0.812	0.763	0.878	0.852	1.16	0.831	1.13	0.698	0.858	0.823
0.730	0.822	0.727	1.13	1.66	1.16	1.35	1.44	1.30	1.34
1.89	1.68	1.52	1.58	2.02	2.60	1.39	1.55	1.35	1.58
2.14	1.61	3.23	4.06	4.06	3.28	2.16	2.29	2.88	2.58
1.76	2.18	1.62	1.91	2.26	1.71	1.88	0.206	0.325	

The weekly traded volume of BTC-USD displays noteworthy variability, with remarks oscillating from very small values to extreme points. The data designate heavy-tailed behavior, right-skewness, and gathering around moderate values, these are typical the features of financial time series. These characteristics show heterogeneity making mixed distributions suitable.

Regarding this Finance dataset, the performance of the DUSLW distribution remains superior to the other models, as it still exhibits a higher p-value, although it has slightly higher AIC, KS, BIC, and HQIC values compared with certain distributions as illustrated in Table 7. along with Figures 12, 13, 14, and 15 show the best fit of this dataset.

Table 7. MLEs for the parameters and associated goodness of fit statistics for the total weekly traded volume of BTC-USD data.

Distributions	LL	MLE	AIC	HQIC	BIC	KS-Test	P-value
DUSLW	-172.2643	$\hat{\beta} = 0.0783$	352.5285	357.5135	364.8042	0.0448	0.8000
		$\hat{\xi} = 0.8491$					
		$\hat{k} = 3.1463$					
		$\hat{\lambda} = 2.0565$					
DUSL	-225.3166	$\hat{\xi} = 1.0304$	452.6333	453.8795	455.7022	0.2385	0.0000
Lindley	-239.8829	$\hat{\xi} = 0.8264$	481.7657	483.0120	484.8346	0.2633	0.0000
Weibull	-174.2106	$\hat{k} = 2.7002$	352.4212	354.9137	358.5591	0.0536	0.7509
		$\hat{\lambda} = 2.1031$					
Gamma	-174.8650	$\hat{\alpha} = 5.9006$	353.7299	356.2224	359.8678	0.0879	0.1715
		$\hat{\gamma} = 3.1511$					
TDUSW	-175.0497	$\hat{\alpha} = 2.2584$ $\hat{\beta} = 0.3004$	354.0993	356.5918	360.2371	0.0548	0.7262

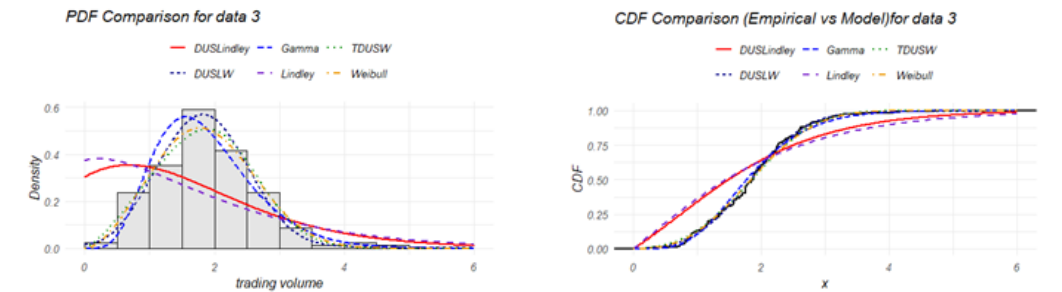


Figure 12. Theoretical and empirical PDF and CDF comparison of the the total weekly traded volume of BTC-USD.

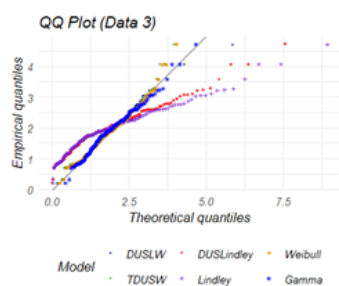


Figure 13. QQ plot comparison of dataset 3.

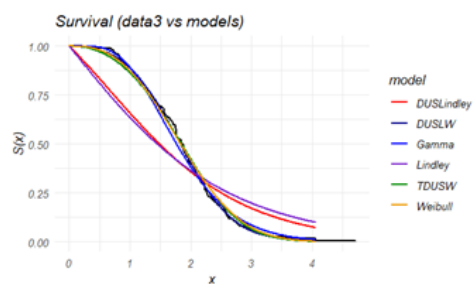


Figure 14. Survival plot comparison of dataset 3.

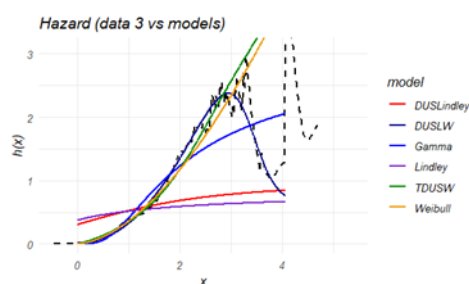


Figure 15. Hazard plot comparison of dataset 3.

6.4 Dataset 4: Waiting times before service.

0.8	2.9	4.3	5.0	6.7	8.2	9.7	11.9	14.1	19.9
0.8	3.1	4.3	5.3	6.9	8.6	9.8	12.4	15.4	20.6
1.3	3.2	4.4	5.5	7.1	8.6	10.7	12.5	15.4	21.3
1.5	3.3	4.4	5.7	7.1	8.6	10.9	12.9	17.3	21.4
1.8	3.5	4.6	5.7	7.1	8.8	11.0	13.0	17.3	21.9
1.9	3.6	4.7	6.1	7.1	8.8	11.0	13.1	18.1	23.0
1.9	4.0	4.7	6.2	7.4	8.9	11.1	13.3	18.2	27.0
2.1	4.1	4.8	6.2	7.6	8.9	11.2	13.6	18.4	31.6
2.6	4.2	4.9	6.2	7.7	9.5	11.2	13.7	18.9	33.1
2.7	4.2	4.9	6.3	8.0	9.6	11.5	13.9	19.0	38.5

This data set is taken from Ghitany et al. (2008) and it describes the waiting times before serving 100 bank customers (in minutes). Such data are often suitable for lifetime distribution analysis. It is continuous, positive exhibiting skewness and variability. Table 8, and Figures 16, 17, 18, and 19 show high fitting of the DUSLW distribution with this data set.

Table 8. MLEs for the parameters and associated goodness of fit statistics for the waiting times of bank customers.

Distributions	LL	MLE	AIC	HQIC	BIC	KS-Test	P-value
DUSLW	-317.4416	$\hat{\beta} = 0.6323$	642.8832	647.1007	653.3039	0.0393	0.9978
		$\hat{\xi} = 0.1847$					
		$\hat{k} = 2.0480$					
		$\hat{\lambda} = 6.9001$					
DUS Lindley	-319.0394	$\hat{\xi} = 0.2230$	640.0788	641.1332	642.6840	0.0589	0.8790
Lindley	-319.0374	$\hat{\xi} = 0.1866$	640.0748	641.1292	642.6800	0.0677	0.7495
Weibull	-318.7307	$\hat{k} = 1.4585$	641.4614	643.5701	646.6717	0.0578	0.8921
		$\hat{\lambda} = 10.9553$					
Gamma	-317.3001	$\hat{\alpha} = 2.0088$	638.6002	640.7090	640.7090	0.0425	0.9936
		$\hat{\gamma} = 0.2034$					
TDUSW	-320.3154	$\hat{\alpha} = 1.2076$	644.6307	646.7394	649.8411	0.0663	0.7712
		$\hat{\beta} = 0.0889$					

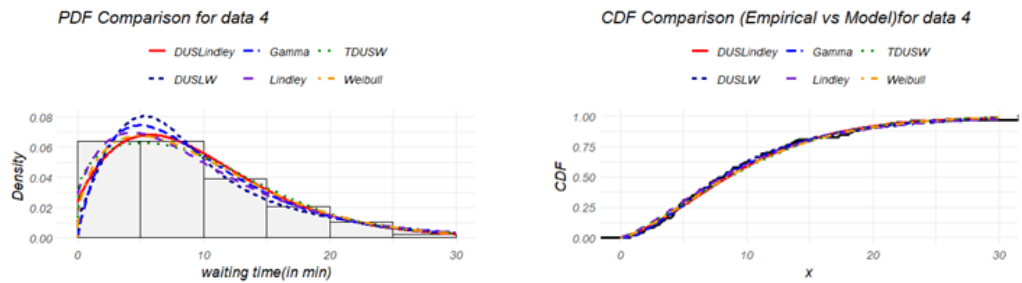


Figure 16. Theoretical and empirical PDF and CDF comparison of the waiting times before serving customers data.

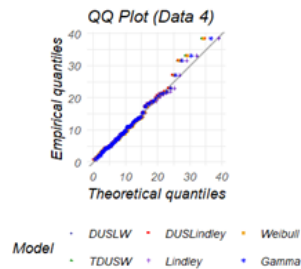


Figure 17. QQ plot comparison of dataset 4.

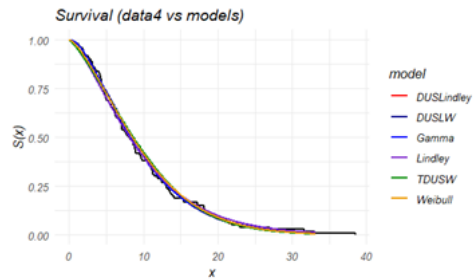


Figure 18. Survival plot comparison of dataset 4.

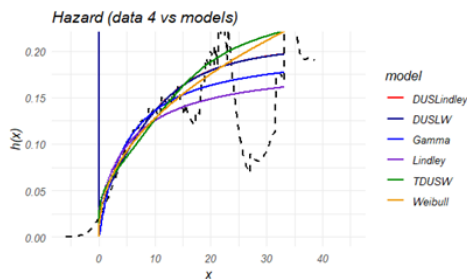


Figure 19. Hazard plot comparison of dataset 4.

7 Conclusion remarks

Data analysis is a complex and fundamental process that mandates the use of distributions that are reliable and applicable in interpreting heterogeneous data. This study assessed the efficacy and applicability of a mixed-type distribution, a combination of the DUS-Lindley and Weibull distributions and tested its reliability in analyzing diverse real-world datasets. The PDF, CDF, ordinary and incomplete moments, and order statistics of this distribution were determined and presented, along with graphs for the PDF and hazard rate function. The expectation–maximization algorithm was used to approximate the parameters of this distribution. Simulations using different sample sizes were also included to evaluate the practicality of this distribution.

The results highlight both the theoretical significance and real-world effectiveness of the proposed DUSLW distribution. Simulation studies were conducted based on the theoretical results to confirm the accuracy of the formulas, showing negligible bias values. The MLE method, along with the EM algorithm, proved reliable, with estimated values highly consistent with the true parameters in the simulations. Additionally, the distribution demonstrated strong performance in modeling real-world datasets, providing a good fit to observed behaviors. These results emphasize the contribution of the DUSLW distribution in advancing theoretical insights in statistical modeling while offering an effective and flexible tool for applied data analysis across various domains, unlike many distributions in the literature that are typically suited to a single type of dataset within a specific domain.

8 Future work

The DUSLW distribution suggests several directions for future studies. For example, adding a new weighted distribution to the proposed model could further improve its flexibility. Focusing on studying functions of the DUSLW distribution, such as minimum and maximum functions, is also a valuable approach. Moreover, future research may explore deriving additional theoretical properties. Using alternative estimation methods could also help improve parameter estimation.

References

- [1] K. Shimizu and H. Hashiguchi, Approximate distribution of eigenvalues of a generalized Wishart matrix under an extended Gaussian model, *Stat. Probab. Lett.*, (2025).
- [2] E. Gómez-Déniz and E. Calderín-Ojeda, The discrete Lindley distribution: properties and applications, *J. Stat. Comput. Simul.*, **81**(11), 1405–1416, (2011).
- [3] D. Hamed and A. Alzaghal, New class of Lindley distributions: properties and applications, *J. Stat. Distrib. Appl.*, **8**, 11, (2021).
- [4] A. Z. Afify, M. Nassar, D. Kumar and G. M. Cordeiro, A new unit distribution: properties, inference, and applications, *Electron. J. Appl. Stat. Anal.*, **15**(2), 460–484, (2022).
- [5] M. G. M. Ghazal, Y.O.U. Tashkand, O. S. Balogun and M. E. Bakr, Exponentiated extended extreme value distribution: properties, estimation, and applications in applied fields, *AIMS Math.*, **9**(7), 17634–17656, (2024).
- [6] Y. L. Oh, F. P. Lim, C. Y. Chen, W. S. Y. Ling and Y. F. Loh, Exponentiated Weibull Burr type X distribution: properties and applications, *Electron. J. Appl. Stat. Anal.*, **15**(3), 553–573, (2022).
- [7] Y. L. Oh, F. P. Lim, C. Y. Chen, W. S. Ling and Y. F. Loh, A new exponentiated beta Burr type X distribution: model, theory and applications, *Sains Malays.*, **52**(1), 281–294, (2023).
- [8] Y. L. Oh, F. P. Lim, C. Y. Chen, W. S. Ling, Y. F. Loh and H. K. Yap, Exponentiated gamma Burr-type X distribution: model, theory and applications, *J. Inequal. Appl.*, **2024**, 142, (2024).
- [9] M. Elgarhy, G. S. S. Abdalla, O. J. Obulezi and E. M. Almetwally, A tangent DUS family of distributions with applications to the Weibull model, *Sci. Afr.*, e02843, (2025).
- [10] A. Alrashidi and I. E. Ragab, Generalized Dinesh–Umesh–Sanjay exponential distribution with application engineering data, *AIP Adv.*, **13**, 115228, (2023).
- [11] N. Yassine, Joint probability distribution and the minimum of normalized random variables, *Procedia Soc. Behav. Sci.*, **230**, 235–240, (2016).
- [12] N. Yassine, A bundling strategy for items with different quality based on functions involving the minimum of two random variables, *Int. J. Eng. Bus. Manag.*, **10**, 1–9, (2018), doi:10.1177/1847979018778919.

- [13] N. Yassine and S. El-Rabih, Assembling components with probabilistic lead times, *IOP Conf. Ser. Mater. Sci. Eng.*, **521**(1), 012013, (2019).
- [14] C. Jneid, N. Yassine and A. El-Zaart, maximum of Exponential-Pareto random variables: theoretical perspective with applications in skin image modeling and production control, *Proc. Int. Conf. Recent Advances in Mathematics and Informatics (ICRAMI)*, IEEE, (2025), doi:10.1109/ICRAMI64946.2025.11472645.
- [15] N. Yassine, The EOQ model with items of imperfect quality and replenishment from different suppliers, *Int. J. Oper. Res.*, **40**(1), 117–136, (2020), doi:10.1504/IJOR.2021.111955.
- [16] N. Yassine, C. Markarian and R. El Khalil, An economic production quantity model with imperfect quality raw material and backorders, In: *Proc. 9th Int. Conf. Oper. Res. Enterp. Syst. (ICORES)*, 202–211, (2020).
- [17] S. Ilani and N. Yassine, Order statistics of mixed random variables: theoretical approach, *Proc. Int. Conf. Recent Advances in Mathematics and Informatics (ICRAMI)*, IEEE, (2025), doi:10.1109/ICRAMI64946.2025.11472778.
- [18] Y. A. Mohammed, B. Yatim and S. Ismail, Mixture model of exponential, gamma and Weibull distributions to analyse heterogeneous survival data, *J. Sci. Res. Rep.*, **5**(2), 132–139, (2015).
- [19] M. Naderi, A. Arabpour and A. Jamalizadeh, Mixture of normal mean-variance Lindley distributions, *J. Stat. Res. Iran*, **13**(2), 197–214, (2017).
- [20] E. E. Elmahdy and A. W. Aboutahoun, A new approach for parameter estimation of finite Weibull mixture distributions for reliability modeling, *Appl. Math. Model.*, **37**(4), 1800–1810, (2013).
- [21] F. Huang, A. Sava, K. H. Adjallah and D. Zhang, A mixture distributions analysis based feature selection approach for bearing remaining useful life estimation, *SN Appl. Sci.*, **5**, 302, (2023).
- [22] R. H. L. Ip and K. Y. K. Wu, A mixture distribution for modelling bivariate ordinal data, *Stat. Pap.*, **65**, 4453–4488, (2024).
- [23] T. Zhan, M. Yi, A. R. Peck, H. Rui and I. Chervoneva, Estimation and model selection for finite mixtures of Tukey's g- & h distributions, *Stat. Comput.*, **35**, 67, (2025).
- [24] N. Yassine and A. AlSagheer, Optimizing a collaborative two-layer supply chain with probabilistic quality, *WSEAS Trans. Bus. Econ.*, **14**, (2017).
- [25] N. Yassine, Inventory planning under supplier uncertainty in a two-level supply chain, *Int. J. Logist. Manag.*, **34**(2), 497–516, (2023).
- [26] C. Carvajal Muquillaza, R. Manríquez and E. Cabrera, θ -weighted mixture distribution: the Weibull-Lomax case, *Front. Appl. Math. Stat.*, **10**, (2024), doi:10.3389/fams.2024.1418589.
- [27] D. Qayoom, A. A. Rather, N. Alsadat, E. Hussam and A. M. Gemeay, A new class of Lindley distribution: System reliability, simulation and applications, *Heliyon*, **10**(19), e38335, (2024).
- [28] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. B*, **39**(1), 1–38, (1977).
- [29] A. S. Al-Moisheer, A. F. Daghestani and K. S. Sultan, Mixture of lindley and inverse Weibull distributions: properties and estimation, *WSEAS Trans. Math.*, **20**, (2021).
- [30] H. S. Klakattawi, Survival analysis of cancer patients using a new extended Weibull distribution, Ph.D. thesis, King Abdulaziz University, (2025).
- [31] M. M. E. Abd El Monsef, W. A. Hassanein and N. M. Kilany, Erlang-Lindley distribution, *Commun. Stat. Theory Methods*, **46**(19), 9494–9506, (2017).

Author information

S. Ilani, Department of Mathematics and Computer Science, Beirut Arab University, Lebanon.
E-mail: ssi186@student.bau.edu.lb

N. Yassine, Department of Mathematics and Computer Science, Beirut Arab University, Lebanon.
E-mail: noura.yassein@bau.edu.lb