Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2025 (43) 3** : 1–13.
ISSN-0037-8712
doi:10.5269/bspm.78451

On Xbeta Distribution: Properties, Estimation and Applications *

Hamouda Messaadia and Halim Zeghdoudi†

ABSTRACT: **Background:** Probability models are essential for evaluating randomness in real-world phenomena and play a crucial role in estimation and prediction. Although many statistical distributions exist, new and more adaptable models are often required to capture modern datasets effectively. By extending existing distributions or introducing new ones, researchers continue to enhance modeling flexibility and applicability. This study follows this line of research by proposing a new probability distribution.

Methods: We introduce the *Xbeta distribution*, a flexible two-parameter family for analyzing bounded data. Its key properties—including shape, survival and hazard functions, quantile function, moments, mode, stress–strength reliability, and entropy measures—are derived in closed form.

Results: Parameters are estimated using maximum likelihood, Anderson–Darling, Cramér–von Mises, and least squares methods. A comprehensive simulation study evaluates estimator performance, while applications to real datasets illustrate the practical value of the model.

Conclusion: The Xbeta distribution demonstrates superior fit and interpretability when applied to flood-level and polyester fiber tensile strength data, outperforming several competing models.
Keywords: Beta distribution, Anderson–Darling method, simulation, modeling, goodness-of-fit.

Contents

* This work was partially supported by the Mohammed-Chérif Messaadia-Souk Ahras University, Algeria.
† Corresponding author
2010 *Mathematics Subject Classification:* 62E10, 62F10, 60E05.
Submitted August 16, 2025. Published October 02, 2025

1. Introduction

Probability models are useful tools for assessing the unpredictability of real-world phenomena. While numerous statistical distributions exist in the literature, there remains an ongoing need to derive more robust and flexible models that can capture the complexity of emerging datasets. These improved models are invaluable for both prediction and estimation purposes. As the availability of large and diverse datasets continues to increase, many traditional distributions often fail to provide an adequate fit. To address this, researchers have contributed significantly by generalizing existing models or developing entirely new probability distributions that are more adaptable to modern data [1,2,3,4,5].

The development of new probability distributions is typically achieved through methods such as parameter induction [6,7], distributional combinations [8], and generator functions [9]. One prominent approach is the *transmutation map*, introduced by Shaw and Buckley [10], which enables the addition of extra parameters to a baseline distribution to capture asymmetry. Given a continuous random variable with cumulative distribution function (CDF) $G(x)$, the transmuted family is defined as

$$F(x) = G(x)\big[1 + \theta - \theta G(x)\big], \qquad -1 < \theta < 1, \tag{1}$$

with the corresponding probability density function (PDF)

$$f(x) = g(x)\big[1 + \theta - 2\theta G(x)\big], \tag{2}$$

where $g(x)$ is the PDF of $G(x)$. This method has been applied in various studies, such as the transmuted-Lomax distribution [17], the transmuted power function distribution [18], and the transmuted new Weibull–Pareto distribution [19]. A comprehensive overview of transmuted families is provided by Tahir and Cordeiro [20].

Another significant contribution to flexible modeling is the Lehmann-type II (L-II) probability model, originally introduced by Lehmann [21]. The closed-form structure of the L-II distribution allows for a thorough investigation of its statistical properties. Various extensions and modifications of the L-II distribution have been proposed, including the modified L-II model [22], the L-II inverse Weibull [23], and the L-II–Teissier distribution [24]. A notable example is the one-parameter modified L-II distribution proposed by Iqbal et al. [22], defined on the unit interval with PDF and CDF

$$f(x) = \frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}}, \quad 0 < x < 1,\ \beta > 0, \tag{3}$$

and

$$F(x) = 1 - \left(\frac{1-x}{1+x}\right)^{\beta}. \tag{4}$$

This distribution, known as the *Xbeta distribution*, exhibits a wide range of hazard rate behaviors, making it highly suitable for modeling diverse data types.

The primary motivation for this work is to develop a new two-parameter probability model on the unit interval, called the *Xbeta distribution*. The key reasons are as follows:

- **Flexibility in shape:** The proposed distribution has probability density and hazard rate functions capable of exhibiting diverse shapes, making it suitable for complex datasets.
- **Analytical tractability:** Several mathematical properties are derived, including moments, quantile function, mode, order statistics, mean residual life, and Rényi entropy.
- **Efficient estimation:** Parameters are estimated using multiple estimation techniques, with a comprehensive simulation study to assess estimator performance.
- **Real-world applicability:** The distribution is applied to a dataset on milk production to demonstrate its practical usefulness.

The remainder of the study is organized as follows: Section 2 introduces the proposed Xbeta distribution in detail, analyzes the shapes of its probability density and hazard rate functions, and derives several of its fundamental mathematical properties, including moments, quantile function, mode, order statistics, mean residual life, and Rényi entropy.

Section 3 focuses on parameter estimation techniques. Four different estimation methods are presented and compared, followed by a comprehensive Monte Carlo simulation study to evaluate the performance of these estimators. In addition, a real-world data application is provided to demonstrate the practical usefulness of the proposed distribution.

Section 4 presents a discussion of the results, including comparisons with other well-known probability models, goodness-of-fit criteria, and likelihood ratio testing. Graphical illustrations and descriptive analyses are also provided to strengthen the empirical findings.

Finally, Section 5 concludes the study with a summary of the main contributions, highlighting the flexibility and applicability of the Xbeta distribution, and outlines potential directions for future research.

2. Material and Method

2.1. Derivation of new model

A random variable X is said to have an Xbeta distribution if its cumulative distribution function (CDF) is

$$F(x) = \left(1 - \left(\frac{1-x}{1+x}\right)^{\beta}\right)\left[1+\theta-\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right], \quad 0<x<1, \tag{5}$$

and the corresponding probability density function (PDF) is

$$f(x) = \frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}}\left[1+\theta-2\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right], \quad 0<x<1, \tag{6}$$

where $\beta > 0$ and $\theta \in [-1, 1]$.

The limiting behavior of the density function of the Xbeta distribution is

$$\lim_{x\to 0} f(x) = 2\beta(1+\theta),$$

and

$$\lim_{x\to 1} f(x) = \begin{cases} \infty, & \beta < 1, \\ 2(1-\theta), & \beta = 1, \\ 0, & \beta > 1. \end{cases}$$

The PDF curves show diverse shapes: increasing, J-shaped, decreasing, reverse-J, bathtub-shaped, and positively skewed.

The survival and hazard functions of the Xbeta distribution are

$$S(x) = 1 - \left(1 - \left(\frac{1-x}{1+x}\right)^{\beta}\right)\left[1+\theta-\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right], \tag{7}$$

$$h(x) = \frac{\frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}}\left[1+\theta-2\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right]}{1-\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\left[1+\theta-\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right]}, \tag{8}$$

with limiting behavior

$$\lim_{x\to 0} h(x) = 2\beta(1+\theta), \qquad \lim_{x\to 1} h(x) = \infty.$$

The hazard and cumulative hazard functions are

$$h(x)=\frac{2\beta}{1-x^2}\left(\frac{(1-\theta)+2\theta\left(\frac{1-x}{1+x}\right)^{\beta}}{(1-\theta)+\theta\left(\frac{1-x}{1+x}\right)^{\beta}}\right), \tag{10}$$

$$H(x)=-\ln\left|\frac{2\beta}{1-x^2}\left(\frac{(1-\theta)+2\theta\left(\frac{1-x}{1+x}\right)^{\beta}}{(1-\theta)+\theta\left(\frac{1-x}{1+x}\right)^{\beta}}\right)\right|. \tag{11}$$

The reverse hazard rate is

$$r_h(x)=\frac{f(x)}{F(x)}=\frac{\frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}}\left[1+\theta-2\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right]}{\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\left[1+\theta-\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right]}. \tag{12}$$

The Mills ratio is

$$M_0(x)=\frac{(1-x^2)\left[(1-\theta)+\theta\left(\frac{1-x}{1+x}\right)^{\beta}\right]}{2\beta\left[(1-\theta)+2\theta\left(\frac{1-x}{1+x}\right)^{\beta}\right]}. \tag{13}$$

2.2. Mathematical properties

2.2.1. Mixture representation. The PDF can be expressed as

$$f(x)=\frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}}\left[1-\theta+2\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right].$$

We use the binomial expansions

$$(1-\eta)^{m-1}=\sum_{i=0}^{\infty}(-1)^i\binom{m-1}{i}\eta^i, \tag{14}$$

$$(1+\eta)^{-m-1}=\sum_{j=0}^{\infty}(-1)^j\binom{m+1}{j}\eta^j. \tag{15}$$

Substituting in Eq. (6), we obtain

$$f(x)=2\beta(1-\theta)\sum_{i=0}^{\infty}(-1)^i\binom{\beta+1}{i}x^i(1-x)^{\beta-1}+4\theta\beta\sum_{i=0}^{\infty}(-1)^i\binom{2\beta+1}{i}x^i(1-x)^{2\beta-1}. \tag{16}$$

2.2.2. Ordinary moments. The r-th raw moment is

$$\mu_r'=\mathbb{E}[X^r]=\int_0^1 x^r f(x)\,dx.$$

Using Eq. (16), we obtain

$$\mu_r'=2\beta(1-\theta)\sum_{i=0}^{\infty}(-1)^i\binom{\beta+1}{i}B(r+i+1,\beta)+4\theta\beta\sum_{i=0}^{\infty}(-1)^i\binom{2\beta+1}{i}B(r+i+1,2\beta). \tag{17}$$

2.2.3. Quantile function. The quantile function of the Xbeta distribution can be derived using the inverse transformation method. Let R be a uniform random variable on $(0,1)$. Then, the quantile function is obtained by solving

$$x = G^{-1}(R),$$

where $G(\cdot)$ denotes the CDF of the Xbeta distribution.

After algebraic simplification, the final expression for the quantile function is given by

$$x_R = \frac{1 - \left\{ \frac{1}{\sqrt{\theta}} \left[\sqrt{1 + \left(\frac{1-\theta}{2\sqrt{\theta}}\right)^2 - R} - \frac{1-\theta}{2\sqrt{\theta}} \right] \right\}^{1/\beta}}{1 + \left\{ \frac{1}{\sqrt{\theta}} \left[\sqrt{1 + \left(\frac{1-\theta}{2\sqrt{\theta}}\right)^2 - R} - \frac{1-\theta}{2\sqrt{\theta}} \right] \right\}^{1/\beta}}. \tag{18}$$

This expression of the quantile function can be employed to generate random numbers from the Xbeta distribution.

2.2.4. Mode. The mode is obtained by solving $\frac{d}{dx} \ln f(x) = 0$. It reduces to

$$x = \frac{3\beta + 2}{2}. \tag{19}$$

2.2.5. Order statistics of Xbeta distribution. Order statistics play an important role in many disciplines, particularly in life testing and reliability analysis. Let a random sample of size n be $x_1, x_2, \ldots, x_n$ drawn from the Xbeta distribution. Denote by $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ the corresponding order statistics. In particular, $X_{(i)}$, $X_{(1)}$, and $X_{(n)}$ represent the i-th, minimum, and maximum order statistics, respectively.

The PDF of the i-th order statistic has the general form

$$f_{(i:n)}(x) = \frac{n!}{(i-1)!(n-i)!} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x), \tag{20}$$

where $f(x)$ and $F(x)$ denote the PDF and CDF of the Xbeta distribution, respectively.

Substituting the expressions of $f(x)$ and $F(x)$ into Eq. (20), the PDF of the i-th order statistic becomes

$$\begin{aligned} f_{(i:n)}(x) = \frac{n!}{(i-1)!(n-i)!} & \left\{ \left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right) \left[1 + \theta - \theta\left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right)\right] \right\}^{i-1} \\ & \times \left\{ 1 - \left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right) \left[1 + \theta - \theta\left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right)\right] \right\}^{n-i} \\ & \times \frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}} \left[1 + \theta - 2\theta\left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right)\right]. \end{aligned} \tag{21}$$

The PDF of the smallest order statistic ($X_{(1)}$) is

$$f_{(1:n)}(x) = n \left\{1 - F(x)\right\}^{n-1} f(x),$$

or equivalently, in expanded form:

$$\begin{aligned} f_{(1:n)}(x) = n & \left\{ 1 - \left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right) \left[1 + \theta - \theta\left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right)\right] \right\}^{n-1} \\ & \times \frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}} \left[1 + \theta - 2\theta\left(1 - \left(\frac{1-x}{1+x}\right)^\beta\right)\right]. \end{aligned} \tag{22}$$

The PDF of the largest order statistic ($X_{(n)}$) is

$$f_{(n:n)}(x) = n\left\{\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\left[1+\theta-\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right]\right\}^{n-1} \times \frac{2\beta(1-x)^{\beta-1}}{(1+x)^{\beta+1}}\left[1+\theta-2\theta\left(1-\left(\frac{1-x}{1+x}\right)^{\beta}\right)\right]. \tag{23}$$

2.2.6. Mean Residual Life. The mean residual life (MRL) function is defined as

$$m_n(t) = \mathbb{E}\left[(X-t) \mid X>t\right] = \frac{1}{S(t)}\int_t^1 xg(x)\,dx - t, \tag{24}$$

where $S(t)$ is the survival function and $g(x)$ is the PDF of the Xbeta distribution.

Now, consider the term

$$\int_t^1 xg(x)\,dx = 2\beta(1-\theta)\sum_{i=0}^{\infty}(-1)^i\binom{\beta+1}{i}\int_t^1 x^{i+1}(1-x)^{\beta-1}\,dx + 4\theta\beta\sum_{i=0}^{\infty}(-1)^i\binom{2\beta+1}{i}\int_t^1 x^{i+1}(1-x)^{\beta-1}\,dx.$$

By evaluating the integrals in terms of Beta functions, we obtain

$$\int_t^1 xg(x)\,dx = 2\beta(1-\theta)\sum_{i=0}^{\infty}(-1)^i\binom{\beta+1}{i}\left[B(2+i,\beta)-B(t;2+i,\beta)\right] + 4\theta\beta\sum_{i=0}^{\infty}(-1)^i\binom{2\beta+1}{i}\left[B(2+i,2\beta)-B(t;2+i,2\beta)\right],$$

where $B(\cdot,\cdot)$ is the complete Beta function and $B(t;\cdot,\cdot)$ is the incomplete Beta function.

Finally, substituting into Eq. (24), the MRL function is expressed as

$$m_n(t) = \frac{1}{S(t)}\left\{2\beta(1-\theta)\sum_{i=0}^{\infty}(-1)^i\binom{\beta+1}{i}\left[B(2+i,\beta)-B(t;2+i,\beta)\right] + 4\theta\beta\sum_{i=0}^{\infty}(-1)^i\binom{2\beta+1}{i}\left[B(2+i,2\beta)-B(t;2+i,2\beta)\right]\right\} - t. \tag{25}$$

2.2.7. Rényi Entropy. The Rényi entropy is defined, for any probability model, as

$$I_R(v) = \frac{1}{1-v}\log\big[I(v)\big], \qquad v>0,\ v\neq 1,$$

where

$$I(v) = \int_{-\infty}^{+\infty} g^{v}(x)\,dx,$$

and $g(x)$ denotes the PDF of the distribution under consideration.

Using Eq. (6), the Rényi entropy for the Xbeta distribution is obtained as

$$I_R(v) = \frac{1}{1-v}\left\{v\log\beta + v\log 2 + \log\left[\sum_{i=0}^{\infty}\sum_{j=0}^{i}\binom{v}{i}\binom{i}{j}\theta^i 2^j \times \sum_{t=0}^{\infty}\frac{\Gamma(v(\beta+1)+\beta j+t)}{\Gamma(t)\,\Gamma(v(\beta+1)+\beta j)}\,B(v+1,\ \beta j+v(\beta-1)+1)\right]\right\}.$$

Here, $\Gamma(\cdot)$ represents the Gamma function and $B(\cdot,\cdot)$ denotes the Beta function.

3. Results

3.1. Estimation of Xbeta Parameters

3.1.1. Maximum Likelihood Estimation. Let X be a random variable such that $X \sim \text{Xbeta}(\beta, \theta)$ with sample size n. By using Eq. (6), the likelihood function is defined as

$$L(\theta) = \prod_{i=1}^{n} \left(\frac{2\beta(1-x_i)^{\beta-1}}{(1+x_i)^{\beta+1}} \left[1 - \theta + 2\theta \Big(\frac{1-x_i}{1+x_i}\Big)^{\beta} \right] \right).$$

The log-likelihood function is

$$\begin{aligned} \ell(\theta) &= n \log \beta + n \log 2 + (\beta - 1) \sum_{i=1}^{n} \log(1 - x_i) \\ &\quad - (\beta + 1) \sum_{i=1}^{n} \log(1 + x_i) + \sum_{i=1}^{n} \log \left(1 - \theta + 2\theta \Big(\tfrac{1-x_i}{1+x_i}\Big)^{\beta} \right). \end{aligned} \tag{26}$$

Taking partial derivatives of Eq. (26), we obtain

$$\begin{aligned} \frac{\partial \ell(\theta)}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^{n} \log(1 - x_i) - \sum_{i=1}^{n} \log(1 + x_i) \\ &\quad + 2\theta \sum_{i=1}^{n} \frac{\left(\frac{1-x_i}{1+x_i}\right)^{\beta} \log\left(\frac{1-x_i}{1+x_i}\right)}{1 - \theta + 2\theta \left(\frac{1-x_i}{1+x_i}\right)^{\beta}}, \end{aligned} \tag{27}$$

and

$$\frac{\partial \ell(\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{2\left(\frac{1-x_i}{1+x_i}\right)^{\beta} - 1}{1 - \theta + 2\theta \left(\frac{1-x_i}{1+x_i}\right)^{\beta}}. \tag{28}$$

Since no closed-form solution exists for Eqs. (27)–(28), numerical methods are required. A suitable method for solving these likelihood equations is the *Newton–Raphson algorithm.*

—

3.1.2. Anderson–Darling Estimation. The Anderson–Darling (AD) estimation method minimizes

$$AD(\theta) = -n - \frac{1}{2} - \sum_{i=1}^{n} (2i - 1) \left[\log F(x_i; \theta) + \log \bar{F}(x_{n+1-i}; \theta) \right],$$

where $\bar{F}(x, \theta) = 1 - F(x, \theta)$.

3.1.3. Cramér–von Mises Estimation. The Cramér–von Mises (CVM) estimation method minimizes

$$C(\theta) = \frac{1}{2n} + \sum_{i=1}^{n} \left[F(x_i; \theta) - \frac{2i - 1}{2n} \right]^2.$$

For the Xbeta distribution,

$$\begin{aligned} C(\theta) = \frac{1}{2n} + \sum_{i=1}^{n} \Bigg[1 - \left(\frac{1-x_i}{1+x_i} \right)^{\beta} \left(1 + \theta - \theta \Big(\tfrac{1-x_i}{1+x_i}\Big)^{\beta} \right) \\ - \frac{2i-1}{2n} \Bigg]^2. \end{aligned}$$

3.1.4. Least Squares Estimation. The Least Squares (LS) estimation minimizes

$$LS(\theta) = \sum_{i=1}^{n} \left[F(x_i; \theta) - \frac{i}{n+1} \right]^2 .$$

For the Xbeta distribution,

$$LS(\theta) = \sum_{i=1}^{n} \left[1 - \left(\frac{1 - x_i}{1 + x_i} \right)^{\beta} \left(1 + \theta - \theta \left(\frac{1-x_i}{1+x_i} \right)^{\beta} \right) - \frac{i}{n+1} \right]^2 .$$

—

3.2. Simulation Study

This section compares the outcomes of four different estimation methods using a comprehensive simulation study.

Using different choices of Xbeta parameters, we generated 5000 random samples of sizes $n = 20, 60, 100,$ and 250 from the Xbeta distribution. The absolute bias (AB), mean relative error (MRE), and mean square error (MSE) values of each estimation method were computed. The simulations were carried out in R, and results are presented in Tables 1 and 2.

The simulation results revealed the following:

- The majority of Xbeta estimates displayed a decrease in AB and MSE as the sample size increased.
- All estimators exhibited excellent performance, producing very low MSEs, with AB values typically close to zero relative to the true parameter values.
- Minimal variations were observed in the MSE and AB values across the four estimation methods.
- Among the methods, maximum likelihood estimation consistently demonstrated the highest accuracy.

Table 1: AB and MSE for ML, AD, CVM, and OLS estimates at $\beta = 0.5$.

		ML		AD		CVM		OLS	
θ	n	AB	MSE	AB	MSE	AB	MSE	AB	MSE
20	θ	0.485	0.585	0.129	0.321	0.105	0.335	0.144	0.286
	β	0.138	0.078	0.004	0.025	0.000	0.026	0.045	0.023
60	θ	0.346	0.391	0.122	0.273	0.149	0.275	0.165	0.275
	β	0.075	0.035	0.002	0.014	0.009	0.012	0.007	0.014
100	θ	0.280	0.311	0.135	0.229	0.126	0.216	0.145	0.232
	β	0.044	0.017	0.005	0.010	0.011	0.009	0.002	0.009
250	θ	0.222	0.231	0.126	0.178	0.144	0.197	0.182	0.206
	β	0.036	0.012	0.013	0.009	0.025	0.009	0.020	0.008

4. Applications of the Xbeta Model

In this section, two datasets are used to evaluate the applicability and flexibility of the Xbeta distribution in comparison with the Beta, Unit-Lindley [11], New Unit-Lindley [12], Kumaraswamy [13], and McDonald [14] distributions.

The analysis demonstrates the superior performance of the Xbeta model across various real datasets.

Table 2: AB and MSE for ML, AD, CVM, and OLS estimates at $\theta = 0.5$.

		ML		AD		CVM		OLS	
β	n	AB	MSE	AB	MSE	AB	MSE	AB	MSE
20	θ	0.4721	0.4789	0.1141	0.1205	0.1438	0.2322	0.0858	0.1532
	β	0.0885	0.0182	0.0342	0.0078	0.0486	0.0141	0.0218	0.0089
60	θ	0.3111	0.2432	0.0983	0.1366	0.0558	0.1549	0.0418	0.1461
	β	0.0495	0.0052	0.0194	0.0027	0.0165	0.0039	0.0068	0.0035
100	θ	0.2051	0.1551	0.0698	0.0807	0.0392	0.1378	0.0618	0.1376
	β	0.3112	0.0033	0.0121	0.0013	0.0109	0.0025	0.0059	0.0023
250	θ	0.0762	0.0532	0.0355	0.0281	0.0625	0.1434	0.0846	0.1421
	β	0.0111	0.0011	0.0053	0.0005	0.0076	0.0022	0.0082	0.0020

4.1. Dataset I

The first dataset represents 20 observations of the maximum flood level (in millions of cubic feet per second) for the Susquehanna River at Harrisburg, Pennsylvania, reported in [16].

$$\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75,$$
$$0.85, 0.95, 0.12, 0.22, 0.33, 0.44, 0.55, 0.66, 0.77, 0.88, 0.99, 0.854, 0.562, 0.245\}.$$

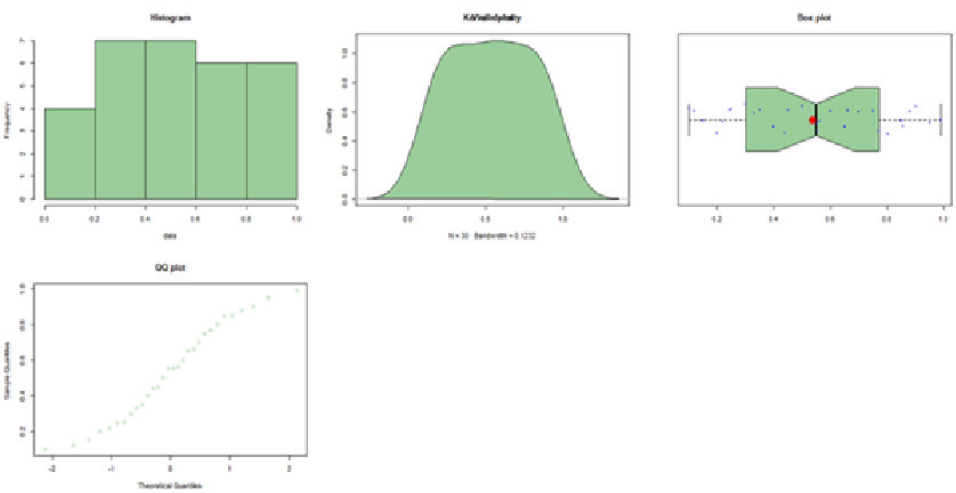

Figure 1: Box plot, QQ-plot, histogram, and kernel density estimation of Dataset I.

4.2. Dataset II

The second dataset corresponds to 30 measurements of tensile strength of polyester fibers, reported in [15].

$$\{0.023, 0.032, 0.054, 0.069, 0.081, 0.094, 0.105, 0.127, 0.148, 0.169, 0.188, 0.216, 0.255, 0.277, 0.311,$$
$$0.361, 0.376, 0.395, 0.432, 0.463, 0.481, 0.519, 0.529, 0.567, 0.642, 0.674, 0.752, 0.823, 0.887, 0.926\}.$$

Table 3: Descriptive Statistics for Dataset I.

Statistic	Value
Count	30.0000
Mean	0.5357
Std	0.2703
Min	0.1000
25%	0.3075
50%	0.5500
75%	0.7650
Max	0.9900

Table 4: Goodness-of-fit measures (p-values) and likelihood-based statistics for Dataset I.

Model	AIC	BIC	-2L	ACIC
Beta	1.4266	1.1831	1.9804	4.7828
New-Unit Lindley	0.8697	/	6.1926	7.5938
Unit-Lindley	0.3047	/	74.6197	76.0209
Kumaraswamy	3.3600	11.7900	2.9313	4.9227
McDonald	117.0500	1.4500	4.1629	0.1544
XBeta	-0.9989	1.0100	1.0775	3.8799

Table 5: Descriptive Statistics for Dataset II.

Statistic	Value
Count	30.0000
Mean	0.3659
Std	0.2685
Min	0.0230
25%	0.1323
50%	0.3360
75%	0.5265
Max	0.9260

Table 6: Goodness-of-fit measures (p-values) and likelihood-based statistics for Dataset II.

Model	AIC	BIC	-2L	ACIC
Beta	1.6205	0.9666	-2.6101	0.1923
New-Unit Lindley	0.2963	/	21.7154	23.1167
Unit-Lindley	1.0504	/	20.1704	21.5716
Kumaraswamy	0.9600	1.6100	-2.6010	0.2010
McDonald	103.0800	1.2000	-0.0410	5.5710
XBeta	-0.2408	1.2508	-2.8758	-0.0734

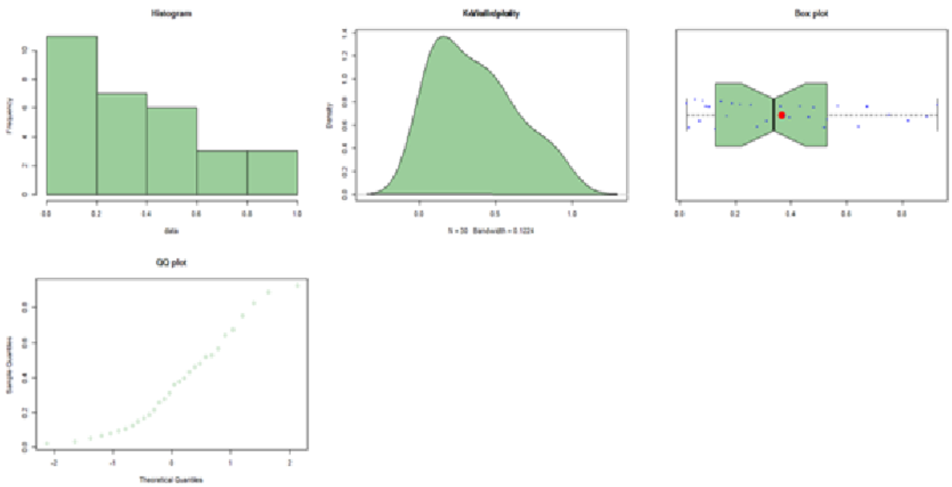

Figure 2: Box plot, QQ-plot, histogram, and kernel density estimation of Dataset II.

4.3. Discussions

To demonstrate the superior performance and practical applicability of the Xbeta distribution, certain well-known model selection statistics were employed, including AIC, BIC, log-likelihood, and ACIC (see Tables 4 and 6). The proposed Xbeta distribution consistently displayed lower values of these goodness-of-fit measures compared to the competing models. Further descriptive summaries of the datasets are given in Tables 3 and 5, while Figures 1–2 illustrate the empirical and fitted distributions, providing clear visual evidence of the superior fit achieved by the Xbeta distribution.

4.4. Likelihood Ratio Test

The likelihood ratio (LR) test was also conducted to further assess model fit. This test compares two different maximum likelihood estimates (MLEs) of a parameter under the null and alternative hypotheses. Let $\ell(\theta_0)$ and $\ell(\theta_1)$ denote the log-likelihoods under the null and alternative hypotheses, respectively. The LR statistic is defined as

$$\Lambda = -2\Big[\ell(\theta_0) - \ell(\theta_1)\Big].$$

For Dataset I, the LR test value was $\Lambda = 5.71$ with a p-value of 0.01704, and for Dataset II, $\Lambda = 13.22$ with a p-value of 0.021. Both results indicate statistical significance, thus identifying the Xbeta distribution as the best-fitting model among the alternatives considered.

5. Conclusion and Perspectives

In this study, a new two-parameter unit distribution, named the *Xbeta distribution*, has been proposed and thoroughly investigated. The new generalization provides greater modeling flexibility for real-world datasets compared to several existing competitors.

Several important mathematical properties of the distribution were derived, including the moments, quantile function, mode, mean residual life function, and entropy. Furthermore, four different estimation methods were employed for parameter estimation, and their performance was evaluated through a comprehensive simulation study.

Applications to two real datasets, one related to flood levels and the other to tensile strength of polyester fibers, were presented to demonstrate the usefulness of the proposed distribution. Based on goodness-of-fit measures, it was concluded that the Xbeta distribution provides superior performance compared to several competitive distributions.

Although the findings are encouraging, numerous opportunities for further investigation persist. The Xbeta distribution can be adapted to a regression framework, allowing for the direct modeling of covariate effects on the unit interval, which is particularly pertinent for proportional and bounded data. Secondly, Bayesian estimation methods and Markov Chain Monte Carlo (MCMC) techniques could be devised to enhance the classical estimation approaches examined herein. Third, multivariate and dependent structures derived from the Xbeta distribution may be examined, for instance, via copula constructions or autoregressive frameworks, to enhance the modeling of intricate real-world phenomena. Ultimately, additional applications in domains such as finance, reliability, survival analysis, and environmental sciences may yield further insights into the practical utility of the proposed model.

Potential Conflict of Interest

All authors declare that there is no conflict of interest related to this article.

References

1. Ahsan-ul-Haq, M., Elgarhy, M., Chesneau, C., & Jamal, F. (2019). A new generator of distributions with applications. *Communications in Statistics – Theory and Methods*, 48(19), 4750–4769.
2. Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1), 63–79.
3. Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull–G family of probability distributions. *Journal of Data Science*, 12(1), 53–68.
4. Rahman, M. M., Ahsan-ul-Haq, M., Chesneau, C., & Jamal, F. (2023). Flexible generator distributions and applications. *Journal of Probability and Statistics*, Article ID 123456.
5. Sapkota, K., Shakya, S., & Paudel, R. (2023). New generated families of distributions for modeling lifetime data. *Mathematics*, 11(2), 215.
6. Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2), 171–178.
7. Marshall, A. W., & Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3), 641–652.
8. Lemonte, A. J., Silva, G. O., & Cordeiro, G. M. (2014). A new class of Weibull models. *Computational Statistics & Data Analysis*, 69, 114–128.
9. Ahsan-ul-Haq, M., & Elgarhy, M. (2018). On the transmuted new distributions. *Journal of Statistical Distributions and Applications*, 5(1), 1–17.
10. Shaw, W. T., & Buckley, I. R. C. (2009). The alchemy of probability distributions: Beyond Gram–Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Research Report.*
11. Mazucheli, J., Menezes, A. F., & Dey, S. (2018). Unit-Lindley distribution: Properties and applications. *Communications in Statistics – Theory and Methods*, 47(3), 568–588.
12. Mazucheli, J., Menezes, A. F., Ghitany, M. E., & Louzada, F. (2020). The new unit-Lindley distribution: Properties, regression model and applications. *Brazilian Journal of Probability and Statistics*, 34(3), 428–453.
13. Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1–2), 79–88.
14. McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3), 647–663.
15. Quesenberry, C. P., & Hales, D. S. (1980). Concentration of strengths from stress–strain data. *Journal of Applied Statistics*, 29(3), 221–227.
16. Dumonceaux, R., & Antle, C. E. (1973). Discriminating between the lognormal and Weibull distributions. *Technometrics*, 15(4), 923–926.
17. Ashour, S. K., and Eltehiwy, M. A. (2013). Transmuted Lomax distribution. *Journal of Advanced Research*, **4**(5), 393–399.
18. Ahsan-ul-Haq, M., and Elgarhy, M. (2016). Transmuted power function distribution: properties and applications. *Pakistan Journal of Statistics and Operation Research*, **12**(3), 443–451.
19. Tahir, M. H., Elgarhy, M., Cordeiro, G. M., and Mansoor, M. (2018). Transmuted new Weibull–Pareto distribution: properties and applications. *Communications in Statistics – Theory and Methods*, **47**(21), 5197–5215.
20. Tahir, M. H., and Cordeiro, G. M. (2015). Transmuted families of distributions: a review and new results. *Brazilian Journal of Probability and Statistics*, **29**(2), 175–195.
21. Lehmann, E. L. (2012). *Elements of Large-Sample Theory.* Springer Science & Business Media.
22. Iqbal, Z., Tahir, M. H., Cordeiro, G. M., and Alizadeh, M. (2021). A one-parameter modified Lehmann type II distribution: properties and applications. *Journal of Statistical Computation and Simulation*, **91**(13), 2681–2702.

23. Tomazella, V. L., Lemonte, A. J., and Moreno-Arenas, G. (2022). The Lehmann type II inverse Weibull distribution: properties and applications. *Communications in Statistics – Simulation and Computation*, **51**(7), 3845–3863.

24. Kumaran, R., and Jha, M. K. (2023). A new Lehmann type II–Teissier distribution with applications. *Journal of Applied Statistics*, **50**(6), 1312–1330.

Mohammed-Chérif Messaadia-Souk Ahras University, Algeria

LaPS laboratory, Badji Mokhtar University–Annaba,

B.P.12, Annaba, 23000, Algeria.

E-mail address: h.messaadia@univ-soukahras.dz, halim.zeghdoudi@univ-annaba.dz