AIMS Mathematics, 10(8): 17740–17778.
DOI: 10.3934/math.2025792
Received: 03 June 2025
Revised: 30 July 2025
Accepted: 31 July 2025
Published: 05 August 2025

http://www.aimspress.com/journal/Math

Research article

A weighted inverse exponential distribution: Properties, estimation and applications

Haiping Ren[1], Jiajie Shi[2], Lianwu Yang[3] and Laijun Luo[4,*]

[1] Teaching Department of Basic Subjects, Jiangxi University of Science and Technology, Nanchang 330013, China
[2] College of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China
[3] School of Mathematics and Computer Sciences, Yichun University, Yichun 336000, China
[4] School of Software Engineering, Jiangxi University of Science and Technology, Nanchang 330013, China

* **Correspondence:** Email: luolaijun@jxust.edu.cn; Tel: +86-137-6718-0932.

Abstract: In this paper, a new unimodal right-skewed weighted inverse exponential distribution is proposed to further solve the problems related to equipment life and survival analyses. Various mathematical properties are analyzed, including the survival and hazard function, tail area property, quantile function, and order statistics. Moreover, the common entropy of the proposed distribution is discussed and compared to accurately measure the information distribution and uncertainty degree of the proposed distribution. The Bayesian estimation and some common classical estimations, such as the maximum likelihood, Anderson Darling, Cramer-von mises, and Ordinary least squares estimation, are used to estimate and analyze the parameters of the proposed distribution. The Lindley approximation and Markov Chain Monte Carlo method with the Metropolis-Hastings algorithm are used to address the complexity of the Bayesian estimation. Additionally, four numerical evaluation criteria are used to compare and analyze the estimated parameters. Finally, by selecting two real datasets for fitting, the proposed distribution is proven to be more flexible and practical in comparison with other distributions. The analysis clearly shows that the proposed distribution efficiently handles these datasets.

Keywords: entropy; tail area property; classical estimation; Bayesian estimation; the Lindley

approximation; Markov Chain Monte Carlo method
Mathematics Subject Classification: 62E10, 62F10

1. Introduction

A statistical distribution model holds significant importance in a diverse range of practical and theoretical applications, such as transportation and medical treatments. Although there are various statistical distribution models in the literature, there are always broader and more flexible distributions that have been proposed and were applied. To solve increasingly complex real-world problems, it is very important to have an appropriate probability distribution to be fitted. In many studies, the inverse transformation technique was used to get the corresponding probability distribution and carry on the analysis and research. For example, Alsadat et al. [1] studied an inverse unit exponential probability distribution using classical and Bayesian methods. They used the Hamilitonian Monte Carlo with no U-turn sampling algorithm to estimate the relevant parameters using twelve parameter estimation methods, including the maximum likelihood estimation. Mahmoud et al. [2] used the transport-of-intensity equation (T-IE) [Y] family to generate the exponential-inverse exponential [Weibull] distribution and explored the relevant important properties, including the Shannon entropy. Pizon [3] investigated inverse Pareto and exponential distributions to create the distribution of the product and derived its properties, such as survival and hazard functions. Chiodo et al. [4] proposed the compound inverse Rayleigh distribution model to accurately described the probability distribution of wind speed extremes. The article traced the origin of the compound inverse Rayleigh distribution and validated its effectiveness through actual wind speed data.

Because of its memoryless properties, the exponential distribution has been one of the most important distributions in probability theory and statistics. It has been widely used in queuing theory and system reliability analyses and has been studied extensively. For example, Ahsan-ul-Haq [5] studied the Poisson moment exponential distribution and showed the flexibility of the distribution with four sets of number characters. Barahona et al. [6] proposed an extended exponential distribution and found that it had great flexibility in terms of the kurtosis coefficient. However, due to the constant failure rate nature of the exponential distribution, it is not applicable to scenarios with inverted bathtub failure rates. To address this limitation, the inverse exponential distribution was proposed by Keller et al. [7], and its properties were studied by Lin et al. [8]. Assuming T follows the exponential distribution with a parameter $\lambda\,(\lambda>0)$, then the inverse exponential distribution can be obtained by transforming T to $T=1/X$, which results in the following probability density function (PDF) of X:

$$f(x)=\frac{\lambda}{x^2}e^{-\frac{\lambda}{x}},x>0. \tag{1}$$

The inverse exponential distribution has been deemed suitable for modeling datasets characterized by inverted bathtub failure rates. However, it possesses a notable limitation: Its inability to effectively analyze highly skewed datasets [9]. Some scholars sought to extend this distribution, and some new distributions based on the inverse exponential distribution, such as the generalized inverse exponential distribution and the inverse unit exponential distribution, have been proposed [10,11]. These distributions have received great attention, and they offer numerous practical applications,

including horse racing, the queuing theory, and modeling wind speeds. Based on the important role of the inverse exponential distribution in stochastic phenomena, the paper proposes an extension of the inverse exponential distribution, which is essentially a weighted inverse exponential distribution. Similar to Khan et al. [12], this distribution is proposed to precisely improve the original inverse exponential distribution by using a weighted distribution.

The rest of this article is structured as follows: In Section 2, the probability density function (PDF) of the weighted inverse exponential distribution and its basic properties are discussed; in Section 3, its mathematical properties are deduced; in section 4, the various entropies of the weighted inverse exponential distribution are inferred; in Section 5, the parameter estimation of the proposed distribution is performed using the maximum likelihood, Anderson Darling, Cramer-von Misses, Ordinary least squares, and a Bayesian estimation; and in Section 6, two real-life datasets are given to test the utility of the distribution.

2. Derivation of the new model and graphical presentation

This section introduces a new extension of the inverse exponential distribution whose cumulative distribution function (CDF) is as follows:

$$F(x;\alpha,\lambda)=(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}},\ x>0,\lambda\geq\alpha>0. \tag{2}$$

Additionally, the PDF of this new distribution is as follows:

$$f(x;\alpha,\lambda)=\frac{e^{-\frac{\lambda}{x}}[\alpha\lambda+x(\lambda-\alpha)]}{x^3}\ ,x>0,\lambda\geq\alpha>0. \tag{3}$$

Here, α is the shape parameter, and λ is the scale parameter.

Obviously, we can prove that the PDF defined in Eq (3) satisfies the following equation:

$$f(x;\alpha,\lambda)=\frac{W(x)f(x;\lambda)}{E(W(X))},$$

where X follows the inverse exponential distribution defined in Eq (1), and $E(W(X))$ is the mathematical expectation of $W(X)=\dfrac{\alpha+X-\dfrac{X\alpha}{\lambda}}{X}=\dfrac{\alpha}{X}+1-\dfrac{\alpha}{\lambda}$. According to Patil and Rao [13], we call the new distribution weighted inverse exponential distribution and briefly denoted it by the weighted inverted exponential (WIE) distribution. The WIE distribution has more flexible tails compared to the original inverse exponential distribution.

- When $\alpha\to 0^+$, the CDF of the inverse exponential distribution can be obtained as follows (Dey [14]):

$$F(x;\lambda)=e^{-\frac{\lambda}{x}},x>0,\lambda>0. \tag{4}$$

- When $\alpha = \lambda$, the single-parameter form of the weighted inverse exponential (SWIE) distribution can be obtained as follows:

$$F(x;\lambda) = (1+\frac{\lambda}{x})e^{-\frac{\lambda}{x}}, x > 0, \lambda > 0. \tag{5}$$

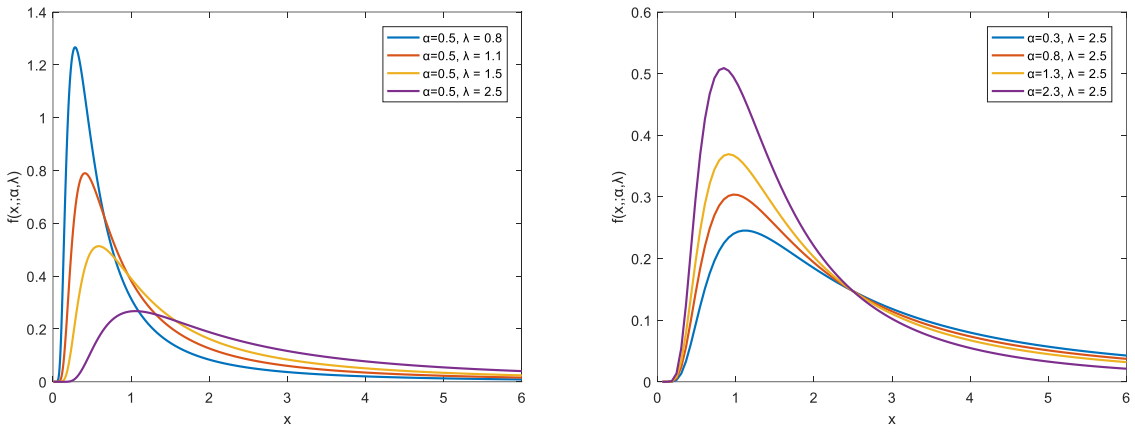
The limiting behavior of the PDF of the WIE distribution when $x \to +\infty$, $x \to 0^+$ and $x \to 1$ is given by,

$$\lim_{x\to+\infty} f(x) = \lim_{x\to 0^+} f(x) = 0,$$

and

$$\lim_{x\to 1} f(x) = e^{-\lambda}(\alpha\lambda + \lambda - \alpha), \text{ respectively.}$$

Graphical presentations of the PDF under different values of parameters α and λ are shown in Figure 1.



(a) **(b)**

Figure 1. Probability density curves of the WIE distribution.

In Figure 1, it is obvious that the probability density plot of the WIE distribution is unimodal and right-skewed, and significantly changes as the two parameters change. When the value of λ remains constant, the peak value of the WIE distribution increases with an increase of α. The PDF plot of the WIE distribution shows a right-leaning trend. However, when the value of α remains unchanged and the value of λ slowly increases, the PDF plot of this distribution gradually tilts to the left and the peak value decreases gradually. Although the plot significantly changes with the parameter changes, the monotonicity remains stable, increasing first and then decreasing.

The expressions of the survival and hazard functions are

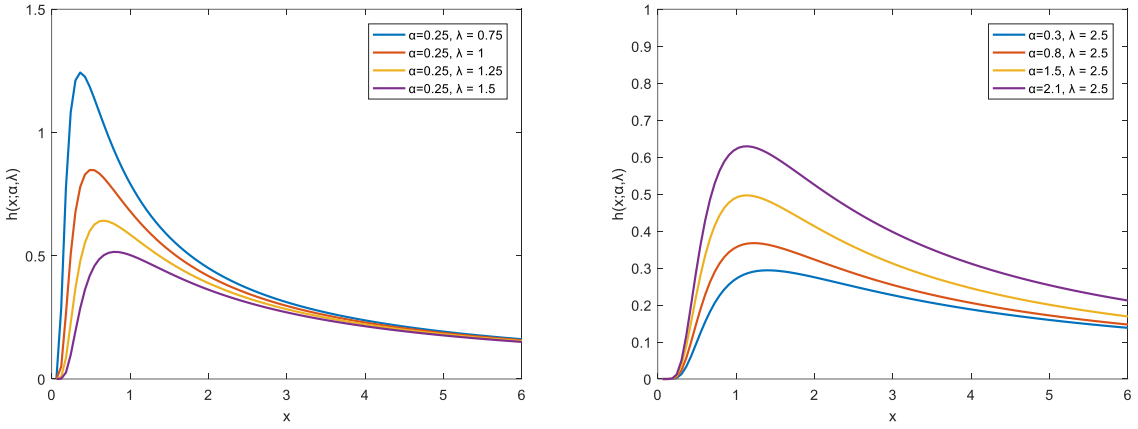
$$S(x;\alpha,\lambda) = 1 - F(x;\alpha,\lambda) = 1 - (1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}, \tag{6}$$

and

$$h(x;\alpha,\lambda)=\frac{f(x;\alpha,\lambda)}{1-F(x;\alpha,\lambda)}=\frac{f(x;\alpha,\lambda)}{S(x;\alpha,\lambda)}=\frac{\alpha\lambda+x(\lambda-\alpha)}{x^2\left(xe^{\frac{\lambda}{x}}-x-\alpha\right)}, \tag{7}$$

respectively.

Similarly, Figure 2 shows the shapes of the hazard function of the WIE distribution under different values of α and λ. The hazard function is unimodal, and it increases first and then decreases until it becomes flat. The smaller the value of λ, the higher the peak of the hazard function is and the more pronounced the right-skewed trend is. α affects the rate of decay of the hazard function. In the case of smaller values of α, the post-peak decay is accelerated.



(a) **(b)**

Figure 2. Hazard-function curves of the WIE distribution.

The cumulative hazard function represents the risk up to a point in time and is generated by integrating the hazard function from 0 to x. The expression of the cumulative hazard function is as follows:

$$H(x)=\int_0^x h(t)dt=-\ln[S(x;\lambda)]=-\ln[1-(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}]. \tag{8}$$

The reverse hazard rate function gives the conditional probability that occurred some events before a given point in time, and is given as follows:

$$r_h(x)=\frac{f(x)}{F(x)}=\frac{\alpha\lambda+x(\lambda-\alpha)}{x^2(x+\alpha)}. \tag{9}$$

The Mills ratio of the WIE distribution is as follows:

$$M_o(x) = \frac{1-F(x)}{f(x)} = \frac{x^2\left(xe^{\frac{\lambda}{x}} - x - \alpha\right)}{\alpha\lambda + x(\lambda - \alpha)}. \tag{10}$$

3. Mathematical properties of the WIE distribution

In this section, we will introduce the mathematical properties of the WIE distribution in detail, such as the mixture representation, quantile function, skewness, kurtosis, and tail area property. This will help us understand the nature of the WIE distribution and evaluate its performance.

3.1. Mixture representation of the WIE distribution

The mixture representation of distributions holds paramount importance in statistics and machine learning, as it facilitates the construction of sophisticated models capable of capturing the multimodal and heterogeneous attributes of data. By amalgamating multiple simple distributions, we can achieve a more precise estimation of the data probability density, thereby facilitating more nuanced cluster analyses and risk assessments. Moreover, the mixture representation augments the flexibility and adaptability of models, thus empowering them to manage data distributions that exhibit diverse shapes and characteristics.

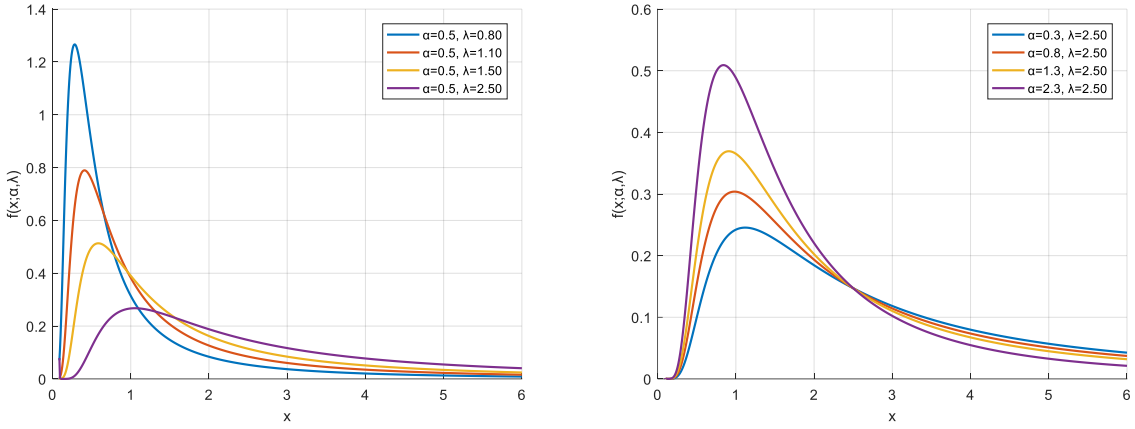
According to the power series expansion formula of exponential functions, the following holds:

$$e^{-\frac{\lambda}{x}} = \sum_{i=0}^{\infty} \frac{(-\frac{\lambda}{x})^i}{i!}, x \in R \setminus \{0\}. \tag{11}$$

By inserting Eq (11) into Eq (3), we obtain the following:

$$f(x) = \frac{\sum_{i=0}^{\infty} \frac{(-\frac{\lambda}{x})^i}{i!}[\alpha\lambda + x(\lambda - \alpha)]}{x^3} = -\frac{\alpha}{x^2} + \sum_{i=1}^{\infty} \frac{\lambda^i(-x)^{-2-i}(-\alpha(-1+i)! - (x+\alpha)i!)}{(-1+i)!i!}. \tag{12}$$

Figure 3 shows the mixture representation of the PDF. The plots of the mixture expression of the PDF agree with that of the original PDF, thus proving that the mixture expression is computationally equivalent to the original expression by substitution, and has important applications in subsequent entropy calculations.



(a) **(b)**

Figure 3. Mixture representation of the WIE distribution's PDF.

3.2. Tail area property of the WIE distribution

As one of the most important statistical features, the k-th moments play an important role in both application and theory. However, according to Lugosi and Mendelson [15] and Nair et al. [16], not all distributions have k-th moments, and the most famous of these are distributions with heavy tails, which are characterized by the absence of one or more moments. If a distribution has k-th moments, then we have the following:

$$\mu_k^{'} = E(x^k) = \int_0^{+\infty} x^k f(x)dx .$$

Using Eq (3), we can derive the following:

$$\begin{aligned}\mu_k^{'} &= \int_0^{+\infty} x^k \frac{e^{-\frac{\lambda}{x}}[\alpha\lambda + x(\lambda-\alpha)]}{x^3}dx \\ &= \int_0^{+\infty} \frac{\alpha\lambda}{x^{3-k}} e^{-\frac{\lambda}{x}}dx + \int_0^{+\infty} \frac{\lambda-\alpha}{x^{2-k}} e^{-\frac{\lambda}{x}}dx \\ &= -(\lambda^{k-1}\int_0^{+\infty} \alpha \frac{\lambda^{1-k}}{x^{1-k}} e^{-\frac{\lambda}{x}} d\frac{\lambda}{x} + (\lambda-\alpha)\lambda^{k-1}\int_0^{+\infty} \frac{\lambda^{-k}}{x^{-k}} e^{-\frac{\lambda}{x}} d\frac{\lambda}{x})\end{aligned} .$$

Let $z = \frac{\lambda}{x}$; thus,

$$\begin{aligned}\mu_k^{'} &= \lambda^{k-1}\int_0^{+\infty} \alpha z^{1-k} e^{-\frac{\lambda}{x}} dz + (\lambda-\alpha)\lambda^{k-1}\int_0^{+\infty} z^{-k} e^{-\frac{\lambda}{x}} dz \\ &= \alpha\lambda^{k-1}\Gamma(2-k;z) + (\lambda-\alpha)\lambda^{k-1}\Gamma(1-k;z)\end{aligned} . \quad (13)$$

According to Eq (13), we find that the moment of the WIE distribution does not exist.

In addition, the existence of the moment can be judged based on the heavy-tailed feature. Leipus et al. [17] analyzed and summarized the methods of judging whether the distribution has a heavy tail feature or a light tail feature. One of them is to judge whether the ratio of the hazard rate to x approaches 0 when x tends to $+\infty$. If the ratio is 0, then the distribution has a heavy tail feature.

Using Eq (7), we can obtain the following:

$$\lim_{x\to\infty}\frac{h(x)}{x}=\frac{\alpha\lambda+x(\lambda-\alpha)}{x^3\left(xe^{\frac{\lambda}{x}}-x-\alpha\right)}=0.$$

It follows that the WIE distribution is a heavy-tailed distribution.

To study its tail decay rate, we substitute Eq (11) into Eq (6), and denote it as follows:

$$S(x)=\frac{\lambda-\alpha}{x}+o(\frac{1}{x}). \tag{14}$$

Thus, the tails of the WIE distribution exhibit a polynomial decay (order $1/x$).

3.3. Quantile function of the WIE distribution

The quantile function can help determine the location and dispersion of a distribution. The quantile function can be obtained by using the famous inverse transformation method for $F(x)$, that is,

$$x=F^{-1}(R).$$

Then, using Eq (2), we can obtain the following:

$$(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}=R. \tag{15}$$

Let $t=R\mathrm{e}^{\frac{\lambda}{x}}$; then, $\ln t-\ln R=\frac{\lambda}{x}$. Hence,

$$1+\frac{\alpha}{\lambda}\ln\frac{t}{R}=t,$$

$$-\frac{\lambda}{\alpha}te^{-\frac{\lambda}{\alpha}t}=-\frac{\lambda R}{\alpha e^{\frac{\lambda}{\alpha}}}.$$

Using the definition of the Lambert W function, we obtain

$$-\frac{\lambda}{\alpha}t=W(-\frac{\lambda R}{\alpha e^{\frac{\lambda}{\alpha}}}),$$

then

$$t = -\frac{\alpha}{\lambda} W\left(-\frac{\lambda R}{\alpha e^{\frac{\lambda}{\alpha}}}\right).$$

Here, W stands for the Lambert W function. According to the properties of the Lambert W function, $\ln W(x) = \ln(x) - W(x)$. Thus, we can obtain the following:

$$x = -\frac{\alpha\lambda}{\lambda + \alpha W\left(-\frac{e^{-\frac{\lambda}{\alpha}} R\lambda}{\alpha}\right)}. \tag{16}$$

3.4. Skewness and kurtosis of the WIE distribution

Skewness and kurtosis are commonly used in statistics to describe the characteristics of data. Skewness provides information about the symmetry of the distribution, which can help to understand the skew of the distribution. Kurtosis provides information about the tail area property of the distribution and can help identify the extreme values and the degree of spikes in the data distribution. The first, second, and third quartiles can be obtained by Eq (16), namely Q(1/4), Q(1/2), and Q(3/4). From these values, we can calculate Galton's skewness, which was detailed by Cooray [18]. At the same time, we can derive Moors [19] 's kurtosis. Let WSk represent skewness and WSr represent kurtosis in the WIE distribution, which are represented as follows:

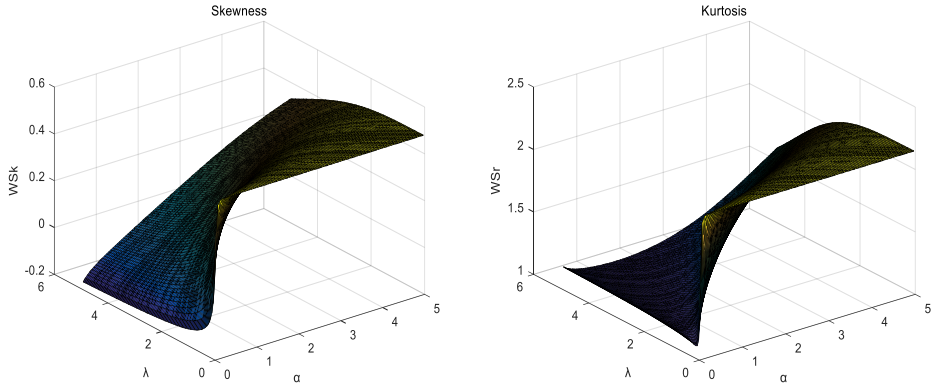
$$WSk = \frac{Q(3/4) + Q(1/4) - 2Q(1/2)}{Q(3/4) - Q(1/4)}, \tag{17}$$

$$WSr = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(3/4) - Q(1/4)}. \tag{18}$$

Table 1 shows the measured values of kurtosis and skewness of the WIE distribution with different parameter values. Obviously, when α is constant, the kurtosis and skewness of the WIE distribution increase with increasing λ values. When λ values remain constant, the kurtosis and skewness values of the WIE distribution are inversely proportional to the α values. Figure 4 displays the Galton's skewness and Moors' kurtosis for the WIE distribution in terms of the parameters α and λ. The surface morphology of skewness indicates that when α increases, the long tail on the right side of the WIE distribution is more pronounced. The surface morphology of kurtosis shows that when α increases, the WIE distribution shows sharp peaks and thick tails. On the contrary, when λ decreases, skewness right skewing and kurtosis spiking simultaneously occur.

Table 1. Values of the kurtosis and skewness of the WIE distribution with different parameter values.

λ	α	WSk	WSr
0.3	0.25	0.3784	1.7966
0.8		0.4715	2.1328
1.2		0.475	2.1398
1.5		0.4757	2.1408
0.3	0.5	0.1678	1.2002
0.8		0.4316	2.0111
1.2		0.4638	2.1146
1.5		0.4703	2.1303
0.3	0.75	0.0652	1.1063
0.8		0.3476	1.6711
1.2		0.4316	2.0111
1.5		0.4538	2.0862
0.3	1	0.0122	1.0888
0.8		0.2583	1.3761
1.2		0.3784	1.7966
1.5		0.4224	1.9762



(a) **(b)**

Figure 4. Plots of skewness and kurtosis for the WIE distribution.

3.5. Mode

As an important index to describe the frequency distribution, the mode can effectively reflect the central tendency of the distribution, especially the points with the highest frequency.

Let

$$\frac{d\ln(f(x))}{dx} = \frac{\lambda}{x^2} + \frac{\lambda-\alpha}{\alpha\lambda + x(\lambda-\alpha)} - \frac{3}{x} = 0.$$

Then, the mode of the WIE distribution can be solved as follows:

$$x = \frac{\pm\sqrt{8\alpha^2\lambda^2 + \lambda^4} + 4\alpha\lambda - \lambda^2}{4(\alpha - \lambda)}. \tag{19}$$

3.6. Order statistics of the WIE distribution

Order statistics are widely used and very important in statistics and probability theory, and they play an important role in describing, inferring, and analyzing the distribution and characteristics of the sample data. Let $X_1, X_2, ..., X_n$ be a sample from the WIE distribution, and $X_{(1)}, X_{(2)}, ..., X_{(n)}$ are the order statistics of $X_1, X_2, ..., X_n$. Here, $X_{(1)}, X_{(n)}$, and $X_{(i)}$ represent the minimum, maximum, and *i*-th order statistics, respectively.

The PDF of the *i*-th order statistics is stated as follows:

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

Then, using Eqs (2) and (3), the PDF of the *i*-th order statistics of the WIE distribution is given as follows:

$$f_{(i,n)}(x) = \frac{n!}{(i-1)!(n-i)!}[(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}]^{i-1}\{1-[(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}]\}^{n-i}\frac{e^{-\frac{\lambda}{x}}[\alpha\lambda + x(\lambda-\alpha)]}{x^3}, \quad x > 0. \tag{20}$$

Meanwhile, the PDF of the minimum order statistics is as follows:

$$f_{(1,n)}(x) = n\{1-[(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}]\}^{n-1}\frac{e^{-\frac{\lambda}{x}}[\alpha\lambda + x(\lambda-\alpha)]}{x^3}, x > 0. \tag{21}$$

Alternatively, the PDF of the maximum order statistics is as follows:

$$f_{(n,n)}(x) = n[(1+\frac{\alpha}{x})e^{-\frac{\lambda}{x}}]^{n-1}\frac{e^{-\frac{\lambda}{x}}[\alpha\lambda + x(\lambda-\alpha)]}{x^3}, x > 0. \tag{22}$$

4. Entropy of the WIE distribution

In information theory, entropy is used to quantify the degree of disorder or uncertainty of a system. It has applications in various fields. For example, in data compression, lower entropy values represent better compression. In communication systems, entropy measures the efficiency and reliability of information transmission. The importance of the concept of entropy has warranted numerous studies [20]. Shannon entropy is the oldest entropy measurement. Lin et al. [21] improved the existing distribution entropy by introducing time series information to improve the accuracy of the signal complexity analysis.

It is easy to prove that the Shannon entropy of the WIE distribution does not exist. Thus, in the following discussion, we will discuss the other entropies, such as Rényi entropy, generalized Tsallis

entropy, Havrda-Charvat entropy, and Arimoto entropy of the WIE distribution.

4.1. Rényi entropy of the WIE distribution

As an extension of information entropy, Rényi entropy provides a broader measure of information through parameterization. Skorski [22] improved the existing estimator with Rényi entropy to obtain a more efficient Rényi-entropy estimation. Singh et al. [23] modified and correctly used Rényi entropy to estimate the velocity distribution and promoted the development of fluid mechanics related to river flow. Tran and Kukal [24] used Rényi entropy's maximum entropy principle to derive a new heavy-tail distribution to reflect the return characteristics of financial assets. Mesfioui [25] advanced the development of life cycle predictability in engineering systems through the development of system signatures using Rényi entropy. Rényi entropy can be calculated by the following formula:

$$I_R(v) = \frac{1}{1-v}\log[I(v)], v > 0, v \neq 1; \tag{23}$$

here,

$$I(v) = \int_{-\infty}^{+\infty} f^v(x)dx, v > 0, v \neq 1. \tag{24}$$

By applying Eq (3), we can obtain the following:

$$I(v) = \int_0^{+\infty} \frac{e^{-\frac{v\lambda}{x}}[\alpha\lambda + x(\lambda-\alpha)]^v}{x^{3v}}dx, \tag{25}$$

which can be obtained by binomial expansion:

$$\begin{aligned}[\alpha\lambda + x(\lambda-\alpha)]^v &= [(\alpha+x)\lambda]^v[1-\frac{x\alpha}{(\alpha+x)\lambda}]^v \\ &= \begin{cases} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{2v-i-j}\lambda^{v-i}\binom{v}{i}\binom{i-v+j-1}{j}x^{i+j}, if \ \frac{x\alpha}{(\alpha+x)\lambda} \le 1, x < \alpha \\ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{3v-2i-j}\lambda^{v-i}\binom{v}{i}\binom{i-v+j-1}{j}x^{2i-v+j}, if \ \frac{x\alpha}{(\alpha+x)\lambda} \le 1, x > \alpha \\ \sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{v+1}\binom{v}{i}\binom{i}{j}x^{v+j}\alpha^{v-j-i}\lambda^i, if \ \frac{x\alpha}{(\alpha+x)\lambda} > 1 \end{cases}. \end{aligned} \tag{26}$$

By inserting Eq (26) into Eq (24), we can obtain the following:

$$I(v)=\int_0^{+\infty}\frac{e^{-\frac{v\lambda}{x}}[\alpha\lambda+x(\lambda-\alpha)]^v}{x^{3v}}dx$$

$$=\begin{cases}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{2v-i-j}\lambda^{j+1-2v}v^{i+j+1-3v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(3v-i-j-1), if\ \frac{x\alpha}{(\alpha+x)\lambda}\le 1, x<\alpha\\ \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{3v-2i-j}\lambda^{i+j+1-3v}v^{2+2i+j-4v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(4v-2i-j-1), if\ \frac{x\alpha}{(\alpha+x)\lambda}\le 1, x>\alpha\\ \sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{v+1}\binom{v}{i}\binom{i}{j}v^{2v+1-j}\alpha^{v-j-i}\lambda^{i-j+2v+1}\Gamma(j-2v-1), if\ \frac{x\alpha}{(\alpha+x)\lambda}>1\end{cases}. \quad (27)$$

Therefore, the Rényi entropy for the WIE distribution is as follows:

$$I_R(v)=\begin{cases}\frac{1}{1-v}\log\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{2v-i-j}\lambda^{j+1-2v}v^{i+j+1-3v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(3v-i-j-1)\right), if\ \frac{x\alpha}{(\alpha+x)\lambda}\le 1, x<\alpha\\ \frac{1}{1-v}\log\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{3v-2i-j}\lambda^{i+j+1-3v}v^{2+2i+j-4v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(4v-2i-j-1)\right), if\ \frac{x\alpha}{(\alpha+x)\lambda}\le 1, x>\alpha\\ \frac{1}{1-v}\log\left(\sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{v+1}\binom{v}{i}\binom{i}{j}v^{2v+1-j}\alpha^{v-j-i}\lambda^{i-j+2v+1}\Gamma(j-2v-1)\right), if\ \frac{x\alpha}{(\alpha+x)\lambda}>1\end{cases}. \quad (28)$$

4.2. Generalized Tsallis entropy of the WIE distribution

Generalized Tsallis entropy was proposed by Tsallis [26] as a generalization of Shannon entropy to solve this problem. Due to the flexible parameterization and other characteristics of Generalized Tsallis entropy, Nikoufar and Fazlolahi [27] gave information monotonicity on the basis of extending Generalized Tsallis entropy. Ghosh and Chaudhuri [28] applied generalized Tsallis entropy to the quantum domain, thus enabling the measurement of position and velocity in space. Then, the Generalized Tsallis entropy of the WIE distribution is given by the following:

$$I_T(v)=\frac{1}{v-1}(1-I(v)), 0<v\ne 1$$

$$=\begin{cases}\frac{1}{v-1}\left(1-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{2v-i-j}\lambda^{j+1-2v}v^{i+j+1-3v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(3v-i-j-1)\right), if\ \frac{x\alpha}{(\alpha+x)\lambda}\le 1, x<\alpha\\ \frac{1}{v-1}\left(1-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i\alpha^{3v-2i-j}\lambda^{i+j+1-3v}v^{2+2i+j-4v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(4v-2i-j-1)\right), if\ \frac{x\alpha}{(\alpha+x)\lambda}\le 1, x>\alpha\\ \frac{1}{v-1}\left(1-\sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{v+1}\binom{v}{i}\binom{i}{j}v^{2v+1-j}\alpha^{v-j-i}\lambda^{i-j+2v+1}\Gamma(j-2v-1)\right), if\ \frac{x\alpha}{(\alpha+x)\lambda}>1\end{cases}. \quad (29)$$

4.3. Havrda-Charvat entropy of the WIE distribution

Havrda-charvat entropy is a generalized form of Shannon entropy. Havrda-Charvat entropy has a unique role in dealing with non-uniform probability distributions and complex systems. For example,

Shi et al. [29] used the Havrda-Charvat entropy to distinguish the uncertainty of similar data, including stock markets. Wang and Shang [30] used Havrda-Charvat entropy to analyze the complexity of time series. Finding that it can effectively extract information of different time scales, Brochet et al. [31] applied Havrda-Charvat entropy to deep learning for pulmonary optical endoscopy classification. The expression of Havrda-Charvat entropy of the WIE distribution is as follows:

$$I_H = \frac{1}{2^{1-v}-1}[I(v)-1], v \neq 1, v>0 .$$

Using Eq (27), we can obtain the following:

$$I_H = \begin{cases} \frac{1}{2^{1-v}-1}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i \alpha^{2v-i-j}\lambda^{j+1-2v} v^{i+j+1-3v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(3v-i-j-1)-1\right), & \frac{x\alpha}{(\alpha+x)\lambda} \leq 1, x<\alpha \\ \frac{1}{2^{1-v}-1}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i \alpha^{3v-2i-j}\lambda^{i+j+1-3v} v^{2+2i+j-4v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(4v-2i-j-1)-1),i\right), & \frac{x\alpha}{(\alpha+x)\lambda} \leq 1, x>\alpha \\ \frac{1}{2^{1-v}-1}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{v+1}\binom{v}{i}\binom{i}{j} v^{2v+1-j}\alpha^{v-j-i}\lambda^{i-j+2v+1}\Gamma(j-2v-1)-1\right), & \frac{x\alpha}{(\alpha+x)\lambda}>1 \end{cases} . \quad (30)$$

4.4. Arimoto entropy of the WIE distribution

Arimoto entropy is more flexible and adaptable than traditional entropy measures in specific situations and has greater advantages to process and analyze probability distributions. Wu and Long [32] innovated a food image segmentation method based on Arimoto entropy of a linear histogram. Alotaibi et al. [33] studied the statical inference of the Arimoto entropy of Kavya-Manoharan's inverse length partial exponential distribution. Li et al. [34] used Arimoto entropy to develop a new method to judge the quality of fused images, which allowed for a more subjective fusion of images. Zhang and Fan [35] proposed an efficient threshold segmentation method using the excellent segmentation performance of two-dimensional Arimoto entropy. Using Eq (27), the Arimoto entropy of the WIE distribution can be obtained as follows:

$$I_A = \frac{v}{1-v}\{I(v)^{\frac{1}{v}}-1\}, v \neq 1, v>0$$

$$= \begin{cases} \frac{v}{1-v}\{[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i \alpha^{2v-i-j}\lambda^{j+1-2v} v^{i+j+1-3v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(3v-i-j-1)]^{\frac{1}{v}}-1\}, if & \frac{x\alpha}{(\alpha+x)\lambda} \leq 1, x<\alpha \\ \frac{v}{1-v}\{[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i \alpha^{3v-2i-j}\lambda^{i+j+1-3v} v^{2+2i+j-4v}\binom{v}{i}\binom{i-v+j-1}{j}\Gamma(4v-2i-j-1)]^{\frac{1}{v}}-1\}, if & \frac{x\alpha}{(\alpha+x)\lambda} \leq 1, x>\alpha \\ \frac{v}{1-v}\{[\sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{v+1}\binom{v}{i}\binom{i}{j} v^{2v+1-j}\alpha^{v-j-i}\lambda^{i-j+2v+1}\Gamma(j-2v-1)]^{\frac{1}{v}}-1\}, if & \frac{x\alpha}{(\alpha+x)\lambda}>1 \end{cases} . \quad (31)$$

4.5. Mathai-Haubold entropy of the WIE distribution

Mathai-haubold entropy has important applications in statistical physics and information theory.

Maya and Irshad [36] studied the kernel estimation of the Mathai-Haubold entropy and residual Mathai-Haubold entropy function under α mixed dependence. Almanjahie et al. [37] proposed and studied a new quantile-based Mathai-Haubold entropy and its properties. Using Eq (27), the Mathai-haubold entropy of the WIE distribution is as follows:

$$I_M = \frac{1}{v-1}[I(2-v)-1], v \neq 1, v < 2$$

$$= \begin{cases} \frac{1}{v-1}\{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i \alpha^{4-2v-i-j}\lambda^{j-3+2v}(2-v)^{i+j-5+3v}\binom{2-v}{i}\binom{i-v+j-1}{j}\Gamma(5-3v-i-j)-1\}, if\ \frac{x\alpha}{(\alpha+x)\lambda} \leq 1, x < \alpha \\ \frac{1}{v-1}\{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^i \alpha^{6-3v-2i-j}\lambda^{i+j-5+3v}(2-v)^{-6+2i+j+4v}\binom{2-v}{i}\binom{i+v+j-3}{j}\Gamma(-4v-2i-j+7)-1\}, if\ \frac{x\alpha}{(\alpha+x)\lambda} \leq 1, x > \alpha \\ \frac{1}{v-1}\{\sum_{i=0}^{\infty}\sum_{j=0}^{i}(-1)^{3-v}\binom{2-v}{i}\binom{i}{j}(2-v)^{-2v+5-j}\alpha^{2-v-j-i}\lambda^{i-j-2v+5}\Gamma(j+2v-5)-1\}, if\ \frac{x\alpha}{(\alpha+x)\lambda} > 1 \end{cases} . \quad (32)$$

4.6. Entropy measure under different parameters

We calculate and analyze the changes of each entropy measure of the WIE distribution under different parameters. We set the parameters values of the WIE distribution to be $\alpha = 0.25, 0.5, 0.75, 1$, $\lambda = 1, 1.2, 1.5, 1.75$, and the parameter of entropy to be $v = 0.75, 1.2$ to obtain the entropy measure under different parameters (see Tables 2 and 3).

Tables 2 and 3 show that when α and v are fixed, an increase in λ makes the tail of the WIE distribution more extended and the probability dispersion higher, thus leading to an increase in the uncertainty of the WIE distribution and a simultaneous rise in various entropy values. Additionally, when v and λ are fixed, the entropy measure values decrease with an increase of α, and the PDF of the WIE distribution is characterized by a right shift and contraction of the tails. An increase in v decreases the sensitivity of I_R, I_T, I_H, I_A and elevates the sensitivity of I_M to some extent. At the same time, these entropy measures are relatively small, thus reflecting that the information of the WIE distribution is concentrated and the uncertainty is relatively small.

Table 2. Various entropy measurements of the WIE distribution under different parameters when $v = 0.75$.

λ	α	I_R	I_T	I_H	I_A	I_M
1	0.25	2.5834	3.6305	4.7970	4.0976	1.2826
1.2		2.8172	4.0896	5.4037	4.6727	1.4352
1.5		3.0898	4.6603	6.1577	5.4027	1.6032
1.75		3.2715	5.0626	6.6893	5.9272	1.7093
1	0.5	2.2298	2.9848	3.9438	3.3083	1.0626
1.2		2.5395	3.5472	4.6869	3.9944	1.2674
1.5		2.8801	4.2180	5.5733	4.8355	1.4819
1.75		3.0975	4.6768	6.1795	5.4241	1.6120
1	0.75	1.7674	2.2223	2.9363	2.4072	0.8077
1.2		2.1985	2.9305	3.8720	3.2430	1.0758
1.5		2.6352	3.7300	4.9284	4.2212	1.3458
1.75		2.8992	4.2574	5.6253	4.8855	1.5040
1	1	1.0453	1.1946	1.5785	1.2506	0.5171
1.2		1.7561	2.2048	2.9133	2.3870	0.8614
1.5		2.3433	3.1858	4.2094	3.5515	1.1957
1.75		2.6710	3.7994	5.0202	4.3079	1.3859

Table 3. Various entropy measurements of the WIE distribution under different parameters when $v = 1.2$.

λ	α	I_R	I_T	I_H	I_A	I_M
1	0.25	1.5984	1.3681	2.1137	1.4032	3.0482
1.2		1.8298	1.5323	2.3675	1.5771	3.4328
1.5		2.1009	1.7154	2.6503	1.7725	3.9050
1.75		2.2820	1.8322	2.8307	1.8982	4.2342
1	0.5	1.2844	1.1327	1.7501	1.1563	2.5085
1.2		1.5747	1.3508	2.0870	1.3850	2.9831
1.5		1.9026	1.5825	2.4449	1.6305	3.5423
1.75		2.1151	1.7247	2.6646	1.7825	3.9204
1	0.75	0.9447	0.8608	1.3300	0.8741	1.8771
1.2		1.2997	1.1445	1.7682	1.1685	2.4742
1.5		1.6899	1.4340	2.2155	1.4728	3.1428
1.75		1.9367	1.6057	2.4809	1.6553	3.5793
1	1	0.5831	0.5504	0.8504	0.5557	1.0631
1.2		1.0090	0.9137	1.4117	0.9287	1.8839
1.5		1.4659	1.2705	1.9630	1.3005	2.6999
1.75		1.7493	1.4761	2.2805	1.5174	3.2081

5. Estimation of the WIE distribution

In this section, we introduce a Bayesian estimation and three classical estimators of the WIE distribution and compare these estimators through simulation.

5.1. Maximum likelihood estimation

The maximum likelihood (ML) estimation is an estimation method based on observations. Suppose that $(X_1, X_2, ..., X_n)$ is a sample from the WIE distribution, and $(x_1, x_2, ..., x_n)$ is the observation of $(X_1, X_2, ..., X_n)$. The ML function can be obtained by Eq (3) as follows:

$$L(\alpha, \lambda \mid x) = \prod_{i=1}^{n} \frac{[\alpha\lambda + x_i(\lambda - \alpha)]}{x_i^3} e^{-\frac{\lambda}{x_i}}. \tag{33}$$

The log-likelihood function is stated as follows:

$$l(\alpha, \lambda \mid x) = \sum_{i=1}^{n} \ln[\lambda\alpha + x_i(\lambda - \alpha)] - 3\sum_{i=1}^{n} \ln x_i - \lambda \sum_{i=1}^{n} \frac{1}{x_i}. \tag{34}$$

Find the first and second order derivatives of the parameters α, λ by Eq (34) as follows:

$$l_\alpha = \sum_{i=1}^{n} \frac{\lambda - x_i}{\lambda\alpha + x_i(\lambda - \alpha)}, \tag{35}$$

$$l_\lambda = \sum_{i=1}^{n} \left(\frac{\alpha + x_i}{\lambda\alpha + x_i(\lambda - \alpha)} - \frac{1}{x_i}\right), \tag{36}$$

$$l_{\alpha\alpha} = \frac{\partial^2 l(\alpha, \lambda | x)}{\partial \alpha^2} = -\sum_{i=1}^{n} \frac{(\lambda - x_i)^2}{(\alpha(\lambda - x_i) + \lambda x_i)^2}, \tag{37}$$

$$l_{\alpha\lambda} = \frac{\partial^2 l(\alpha, \lambda | x)}{\partial \alpha \partial \lambda} = \sum_{i=1}^{n} \frac{x_i^2}{(\alpha(\lambda - x_i) + \lambda x_i)^2} = l_{\lambda\alpha}, \tag{38}$$

$$l_{\lambda\lambda} = \frac{\partial^2 l(\alpha, \lambda | x)}{\partial \lambda^2} = -\sum_{i=1}^{n} \frac{(\alpha + x_i)^2}{(\alpha(\lambda - x_i) + \lambda x_i)^2}. \tag{39}$$

Since the estimated parameters cannot be obtained by directly calculating these equations when Eqs (35) and (36) are equal to 0, we use the Newton Raphson algorithm to obtain the approximate ML estimates of these parameters. The corresponding steps are as follows:

(i) Construct the gradient vector and Hessian matrix of the log-likelihood function as follows:

$$\nabla l(\alpha,\lambda \mid x)=\left(l_{\alpha},l_{\lambda}\right),\ H(\alpha,\lambda \mid x)=\begin{pmatrix} l_{\alpha\alpha} & l_{\alpha\lambda} \\ l_{\lambda\alpha} & l_{\lambda\lambda} \end{pmatrix};$$

(ii) define the initial parameter $(\alpha^{(0)},\lambda^{(0)})^{T}$;

(iii) update the parameter using the Newton-Raphson scheme as follows:

$$\begin{bmatrix} \alpha^{(k+1)} \\ \lambda^{(k+1)} \end{bmatrix}=\begin{bmatrix} \alpha^{(k)} \\ \lambda^{(k)} \end{bmatrix}-\begin{pmatrix} l_{\alpha\alpha} & l_{\alpha\lambda} \\ l_{\lambda\alpha} & l_{\lambda\lambda} \end{pmatrix}\begin{bmatrix} l_{\alpha} \\ l_{\lambda} \end{bmatrix};$$

(iv) repeat (iii) until the absolute difference between successive parameter estimates is less than a small, predefined threshold ε : $|\alpha^{(k+1)}-\alpha^{(k)}|<\varepsilon,|\lambda^{(k+1)}-\lambda^{(k)}|<\varepsilon$; and

(v) The estimated parameter values are denoted as $\hat{\lambda}=\lambda^{(k+1)},\hat{\theta}=\theta^{(k+1)}$.

5.2. Anderson-Darling estimation

The Anderson-Darling (AD) estimation is a commonly used parameter estimation method in statistics. For any distribution, the AD function is defined as follows:

$$AD(\alpha,\lambda)=-n-\frac{1}{n}\sum_{i=1}^{n}(2i-1)[\log(F(x_i;\alpha,\lambda))+\log(\overline{F}(x_{n+1-i};\alpha,\lambda))]. \tag{40}$$

Here, $\overline{F}(x;\alpha,\lambda)=1-F(x;\alpha,\lambda)$.

By plugging in Eq (2), Eq (40) into Eq (2), the following can be obtained:

$$AD(\alpha,\lambda)=-n-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log[(1+\frac{\alpha}{x_i})e^{-\frac{\lambda}{x_i}}]+\log[1-(1+\frac{\alpha}{x_{n+1-i}})e^{-\frac{\lambda}{x_{n+1-i}}}]\}. \tag{41}$$

The estimated parameters of the WIE distribution can be obtained by minimizing the above formula. However, it is very difficult to solve the minimum value by a general numerical method; thus, we use the same Newton Raphson algorithm as above to solve for the minimum value of Eq (41).

5.3. Cramer-von-Mises estimation

Cramer-von-Mises (CVM) is an alternative method to estimate the parameters of a distribution. The corresponding function for any distribution is given as follows:

$$C(\alpha,\lambda)=\frac{1}{12n}+\sum_{i=1}^{n}[F(x_i;\alpha,\lambda)-\frac{2i-1}{2n}]^{2}. \tag{42}$$

Using Eq (1), the above formula is converted to the following:

$$C(\alpha,\lambda)=\frac{1}{12n}+\sum_{i=1}^{n}[(1+\frac{\alpha}{x_i})e^{-\frac{\lambda}{x_i}}-\frac{2i-1}{2n}]^{2}. \tag{43}$$

Then, the estimated parameters can be obtained by minimizing the above formula. We use the Newton Raphson algorithm to get its estimated parameters.

5.4. Ordinary least square estimation

The ordinary least square (OLS) estimation is an important parameter estimation method. The OLS function is defined as follows:

$$OLS(\alpha,\lambda)=\sum_{i=1}^{n}[F(x_i;\alpha,\lambda)-\frac{i}{n+1}]^2 . \tag{44}$$

The corresponding OLS estimator of the WIE distribution is as follows:

$$OLS(\alpha,\lambda)=\sum_{i=1}^{n}[(1+\frac{\alpha}{x_i})e^{-\frac{\lambda}{x_i}}-\frac{i}{n+1}]^2 . \tag{45}$$

Similarly, we can obtain the parameters of the WIE distribution by minimizing the OLS function, and the estimated parameters are obtained by the Newton Raphson algorithm.

5.5. Existence and uniqueness

Due to the nonlinearity of each classical estimator, it is difficult to analyze the mathematical existence and uniqueness of the solution by a rigorous mathematical analysis. However, in this paper, the existence and uniqueness of the solution is necessary for simulation and real data fitting. Thus, it is verified with the help of contour plots [38]. For example, setting $(\alpha,\lambda)=(0.8,1.5)$ and $n=100$, we obtain one simulation sample and four classical estimators. Meanwhile, for convenience, we take negative values of Eqs (41), (43), and (45) to obtain the corresponding contour plots. Figure 5 shows the contour plots of the four classical estimators. It is easy to prove that, under the simulation sample, the estimated parameters of the four classical estimators are existential and unique with the corresponding values:

$$(\alpha_{ML},\lambda_{ML},\alpha_{AD},\lambda_{AD},\alpha_{CVM},\lambda_{CVM},\alpha_{OLS},\lambda_{OLS})=(0.798,1.623,1.213,1.854,0.654,1.297,1.076,1.547) .$$

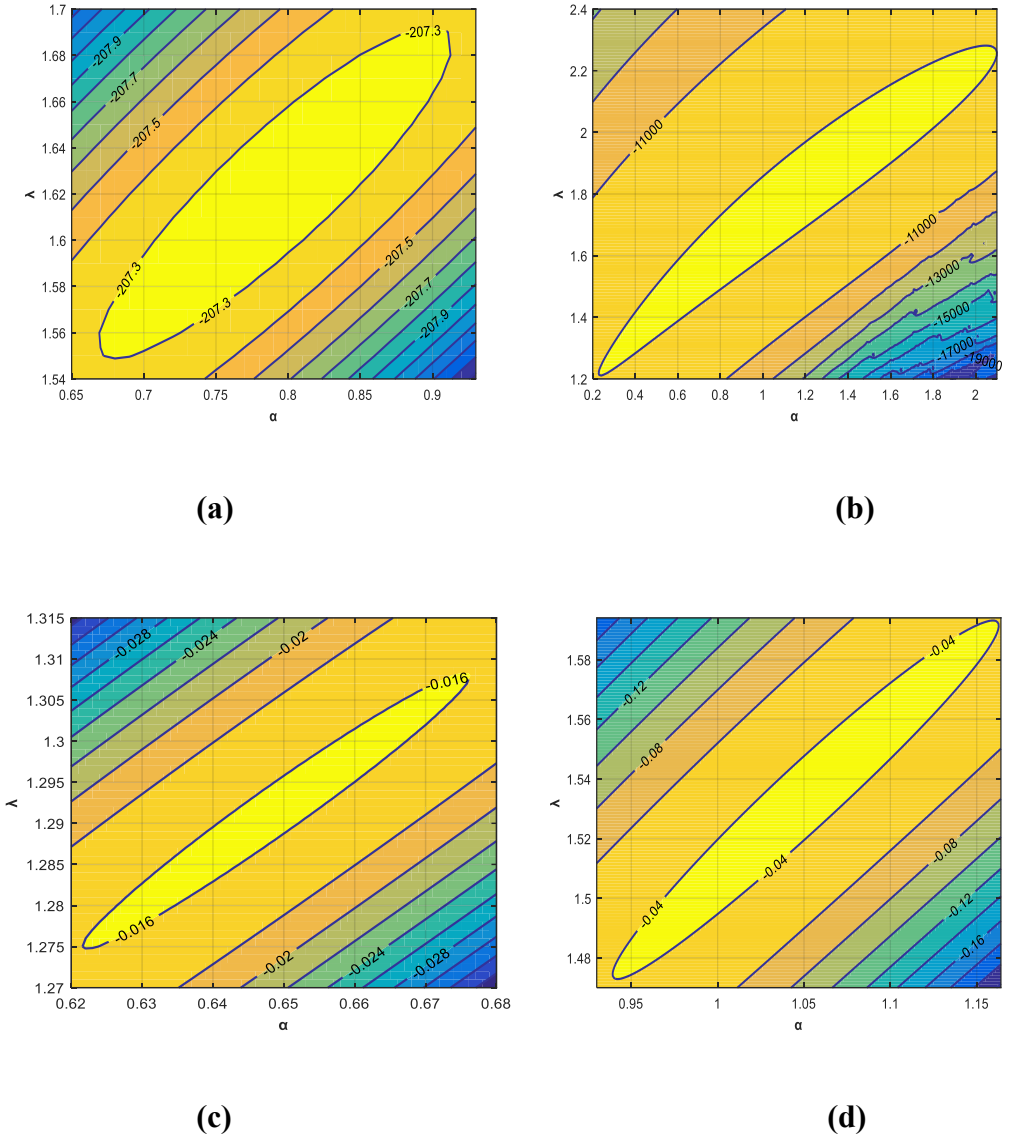


Figure 5. Contour plots for (a) ML estimator, (b) AD estimator, (c) CVM estimator, (d) OLS estimator under the simulation sample.

5.6. Bayesian estimation

In a Bayesian estimation, the prior is categorized into the informative prior and the noninformative prior. In this paper, we only consider the informative prior. Unlike the classical estimation method, which treats the parameters as fixed unknown constants, the core of a Bayesian estimation that uses the informative prior is to combine the probability distribution of the parameter before the observed data with the likelihood information of the observed data to derive the posterior distribution of the parameter, and ultimately make inferences based on the posterior distribution. An informative prior can help us address the shortcomings of exclusively relying on observed data to obtain more accurate parameter estimates. The selection of informative priors for real problems, based on extensive literature and fieldwork, ensures the reliability of prior knowledge and the suitability of problems. In the Bayesian estimation, we need to introduce a loss function to quantify the effect of decision

bias. In this section, we analyze the Bayesian estimation under the squared error loss function (SELF) to estimate the parameters of the WIE distribution. Gamma distributions are often used as prior distributions for a Bayesian estimation due to their non-negative support, intuitive representation of the information prior, and flexible shape to portray different prior information. This analysis of the Bayesian estimation under the SELF is based on the assumption that the unknown parameters independently follow the Gamma distribution, that is, $\alpha \sim Gamma(\sigma_1, \omega_1)$ and $\lambda \sim Gamma(\sigma_2, \omega_2)$. Then, the prior PDFs of α and λ are as follows:

$$\pi(\alpha|\sigma_1, \omega_1) = \frac{\omega_1^{\sigma_1}}{\Gamma(\sigma_1)} \alpha^{\sigma_1 - 1} e^{-\omega_1 \alpha}, \alpha > 0, \sigma_1 > 0, \omega_1 > 0, \tag{46}$$

$$\pi(\lambda|\sigma_2, \omega_2) = \frac{\omega_2^{\sigma_2}}{\Gamma(\sigma_2)} \lambda^{\sigma_2 - 1} e^{-\omega_2 \lambda}, \lambda > 0, \sigma_2 > 0, \omega_2 > 0. \tag{47}$$

Based on the assumption, the joint prior PDF of (α, λ) can be obtained as follows:

$$\pi(\alpha, \lambda) = \frac{\omega_1^{\sigma_1} \omega_2^{\sigma_2}}{\Gamma(\sigma_1)\Gamma(\sigma_2)} \alpha^{\sigma_1 - 1} \lambda^{\sigma_2 - 1} e^{-\omega_1 \alpha - \omega_2 \lambda}. \tag{48}$$

The joint posterior density function is calculated from the probability distribution of the parameter and the likelihood information of the observed data as follows:

$$\pi(\alpha, \lambda|x) = \frac{L(\alpha, \lambda|x)\pi(\alpha, \lambda)}{\int_0^\infty \int_0^\infty L(\alpha, \lambda|x)\pi(\alpha, \lambda) d\alpha d\lambda}. \tag{49}$$

Based on Eq (33) and Eq (48), the following can be obtained:

$$\pi(\alpha, \lambda|x) = \frac{\alpha^{n+\sigma_1 - 1} \lambda^{n+\sigma_2 - 1} \prod_{i=1}^{n} \frac{[\alpha\lambda + x_i(\lambda - \alpha)]}{x_i^3} e^{-\frac{\lambda}{x_i} - \omega_1 \alpha - \omega_2 \lambda}}{\int_0^\infty \int_0^\infty \prod_{i=1}^{n} \frac{[\alpha\lambda + x_i(\lambda - \alpha)]}{x_i^3} e^{-\frac{\lambda}{x_i} - \omega_1 \alpha - \omega_2 \lambda} \alpha^{\sigma_1 - 1} \lambda^{\sigma_2 - 1} d\alpha d\lambda}. \tag{50}$$

To evaluate the accuracy of the estimation, the SELF in the Bayes estimation is used to measure the discrepancy between the estimated value and the true value by calculating the square of their difference. Under the SELF, the Bayes estimator is equal to the mean of the posterior distribution. The SELF is defined as follows [39]:

$$L(\phi, \hat{\phi}) = (\phi - \hat{\phi})^2, \tag{51}$$

and as the corresponding Bayesian estimator of ϕ, $\hat{\phi}$ is defined as follows:

$$\hat{\phi}(\alpha,\lambda)=E[\phi|x]=\int_0^\infty\int_0^\infty \phi(\alpha,\lambda)\pi(\alpha,\lambda\,|\,x)d\alpha d\lambda\,. \tag{52}$$

Obviously, Eq (52) is difficult to derive. Therefore, we adopt two methods to solve this problem, namely the Lindley approximation and the Markov Chain Monte Carlo (MCMC) method.

5.6.1. The Lindley approximation

According to Lindley [40], we construct the integral ratio as follows:

$$I(x)=E[\phi(\alpha,\lambda)|x]=\frac{\int_0^\infty\int_0^\infty \phi(\alpha,\lambda)e^{h(\alpha,\lambda)+l(\alpha,\lambda|x)}d\alpha d\lambda}{\int_0^\infty\int_0^\infty e^{h(\alpha,\lambda)+l(\alpha,\lambda|x)}d\beta d\lambda}, \tag{53}$$

where $h(\alpha,\lambda)=\ln\pi(\alpha,\lambda)$, and $l(\alpha,\lambda\,|\,x)=\ln L(\alpha,\lambda\,|\,x)$.

Meanwhile, Lindley gave an approximate solution to $I(x)$. The Eq (53) can be expressed as follows:

$$\begin{aligned} I(x)\approx{}&\phi(\hat{\alpha},\hat{\lambda})+\frac{1}{2}[(\hat{\phi}_{\alpha\alpha}+2\hat{\phi}_\alpha\hat{h}_\alpha)\tau_{11}+(\hat{\phi}_{\lambda\alpha}+2\hat{\phi}_\lambda\hat{h}_\alpha)\tau_{21}+(\hat{\phi}_{\alpha\lambda}+2\hat{\phi}_\alpha\hat{h}_\lambda)\tau_{12}\\ &+(\hat{\phi}_{\lambda\lambda}+2\hat{\phi}_\lambda\hat{h}_\lambda)\tau_{22}]+\frac{1}{2}[(\hat{\phi}_\alpha\tau_{11}+\hat{\phi}_\lambda\tau_{12})(\hat{l}_{\alpha\alpha\alpha}\tau_{11}+\hat{l}_{\alpha\lambda\alpha}\tau_{12}+\hat{l}_{\lambda\alpha\alpha}\tau_{21}+\hat{l}_{\lambda\lambda\alpha}\tau_{22})\,.\\ &+(\hat{\phi}_\alpha\tau_{21}+\hat{\phi}_\lambda\tau_{22})(\hat{l}_{\lambda\alpha\alpha}\tau_{11}+\hat{l}_{\alpha\lambda\lambda}\tau_{12}+\hat{l}_{\lambda\alpha\lambda}\tau_{21}+\hat{l}_{\lambda\lambda\lambda}\tau_{22})] \end{aligned} \tag{54}$$

Here, $\hat{\alpha}$ and $\hat{\lambda}$ are the ML estimates of α and λ, respectively. Subscript α and λ to indicate the partial derivative of the corresponding function, for example, ϕ_α stands for the first partial derivative of $\phi(\alpha,\lambda)$ with respect to α, $\phi_{\lambda\lambda}$ stands for the second partial derivative of $\phi(\alpha,\lambda)$ with respect to λ, and τ_{ij} is the $(i,j)^{th}$ element of the inverse matrix of $\{-l_{ij}\}$. Then, there is the following:

$$\hat{l}_{\alpha\alpha\alpha}=\left.\frac{\partial^3 l(\alpha,\lambda|x)}{\partial\alpha\partial\alpha\partial\alpha}\right|_{\alpha=\hat{\alpha},\lambda=\hat{\lambda}}=\sum_{i=1}^n\frac{2(\hat{\lambda}-x_i)^3}{(\hat{\alpha}(\hat{\lambda}-x_i)+\hat{\lambda}x_i)^3}, \tag{55}$$

$$\hat{l}_{\alpha\alpha\lambda}=\left.\frac{\partial^3 l(\alpha,\lambda|x)}{\partial\alpha\partial\alpha\partial\lambda}\right|_{\alpha=\hat{\alpha},\lambda=\hat{\lambda}}=-\sum_{i=1}^n\frac{2x_i^2(\hat{\lambda}-x_i)}{(\hat{\alpha}(\hat{\lambda}-x_i)+\hat{\lambda}x_i)^3}=\hat{l}_{\lambda\alpha\alpha}, \tag{56}$$

$$\hat{l}_{\lambda\alpha\lambda}=\left.\frac{\partial^3 l(\alpha,\lambda|x)}{\partial\alpha\partial\alpha\partial\lambda}\right|_{\alpha=\hat{\alpha},\lambda=\hat{\lambda}}=-\sum_{i=1}^n\frac{2x_i^2(\hat{\alpha}+x_i)}{(\hat{\alpha}(\hat{\lambda}-x_i)+\hat{\lambda}x_i)^3}=\hat{l}_{\lambda\lambda\alpha}, \tag{57}$$

$$\hat{l}_{\lambda\lambda\lambda}=\left.\frac{\partial^3 l(\alpha,\lambda|x)}{\partial\lambda\partial\lambda\partial\lambda}\right|_{\alpha=\hat{\alpha},\lambda=\hat{\lambda}}=\sum_{i=1}^n\frac{2(\hat{\alpha}+x_i)^3}{(\hat{\alpha}(\hat{\lambda}-x_i)+\hat{\lambda}x_i)^3}, \tag{58}$$

$$\begin{bmatrix}\tau_{11} & \tau_{12}\\ \tau_{21} & \tau_{22}\end{bmatrix}=\frac{-1}{l_{11}l_{22}-l_{12}l_{21}}\begin{bmatrix}l_{22} & -l_{12}\\ -l_{21} & l_{11}\end{bmatrix}. \tag{59}$$

By defining

$$\varphi = \frac{1}{2}[\phi_{\alpha\alpha}\tau_{11} + \phi_{\alpha\lambda}\tau_{12} + \phi_{\lambda\alpha}\tau_{21} + \phi_{\lambda\lambda}\tau_{22}],$$

$$\xi_1 = \tau_{11}h_\alpha + \tau_{12}h_\lambda + \frac{1}{2}[\tau_{11}^2 l_{\alpha\alpha\alpha} + \tau_{21}\tau_{22}l_{\lambda\lambda\lambda} + (\tau_{11}\tau_{22} + 2\tau_{21}^2)l_{\alpha\lambda\lambda} + 3\tau_{12}\tau_{11}l_{\alpha\alpha\lambda}],$$

$$\xi_2 = \tau_{21}h_\alpha + \tau_{22}h_\lambda + \frac{1}{2}[\tau_{22}^2 l_{\lambda\lambda\lambda} + \tau_{12}\tau_{11}l_{\alpha\alpha\alpha} + (\tau_{11}\tau_{22} + 2\tau_{12}^2)l_{\alpha\alpha\lambda} + 3\tau_{21}\tau_{22}l_{\alpha\lambda\lambda}],$$

the Bayesian estimate of $\phi(\alpha,\lambda)$ can be expressed as

$$\phi = \phi(\hat{\alpha},\hat{\lambda}) + \varphi + \xi_1\phi_\alpha + \xi_2\phi_\lambda .$$

Because of $h(\alpha,\lambda) = \ln\pi(\alpha,\lambda)$, the following is easy to obtain:

$$\hat{h}_\alpha = \frac{\sigma_1 - 1}{\hat{\alpha}} - \omega_1, \hat{h}_\lambda = \frac{\sigma_2 - 1}{\hat{\lambda}} - \omega_2. \tag{60}$$

Based on the SELF, the Bayesian estimates for different parameters are illustrated below:

If $\phi(\alpha,\lambda) = \alpha$, then

$$\hat{\alpha}_{BS} = \hat{\alpha} + \xi_1. \tag{61}$$

If $\phi(\alpha,\lambda) = \lambda$, then

$$\hat{\lambda}_{BS} = \hat{\lambda} + \xi_2. \tag{62}$$

5.6.2. MCMC method

The MCMC method is an algorithm that utilizes a conditional posterior distribution to generate MCMC samples and obtain Bayesian estimates of unknown parameters. From Eq (50), the conditional posterior distributions of the WIE distribution parameters α and λ, are

$$\pi_1(\alpha \mid \lambda, x) = \alpha^{n+\sigma_1-1}e^{-\omega_1 n\alpha}\prod_{i=1}^{n}\frac{[\alpha\lambda + x_i(\lambda-\alpha)]}{x_i^3}, \tag{63}$$

$$\pi_2(\lambda \mid \alpha, x) = \lambda^{n+\sigma_2-1}e^{-\sum_{i=1}^{n}\frac{\lambda}{x_i} - \omega_2 n\lambda}\prod_{i=1}^{n}\frac{[\alpha\lambda + x_i(\lambda-\alpha)]}{x_i^3}, \tag{64}$$

respectively.

Due to the complex nonlinearity of Eqs (63) and (64), we adopt the Metropolis-Hastings (M-H)

algorithm for sampling MCMC samples. The specific steps are as follows [41,41]:

(i) Set the initial values $k=1$ and $(\alpha,\lambda)=(\alpha^{(0)},\lambda^{(0)})$;

(ii) establish candidate values α^* and λ^* from $N(\alpha^{(k-1)},\omega_\alpha^2)$ and $N(\lambda^{(k-1)},\omega_\lambda^2)$, respectively, and the candidate values should meet the following requirements: $\lambda^*\geq\alpha^*>0$, otherwise re-select the values. Moreover, $\lambda^{(k-1)}$ and $\theta^{(k-1)}$ represent the previous state, and ω_α^2 and ω_λ^2 represent the variance of the previous state;

(iii) define the acceptance criterion $\varpi_{(\alpha_{k-1},\alpha^*)}=\min(1,\frac{\pi_1(\alpha^*\mid\lambda^{(k-1)},x)}{\pi_1(\alpha^{(k-1)}\mid\lambda^{(k-1)},x)})$ and $\varpi_{(\lambda_{k-1},\lambda^*)}=\min(1,\frac{\pi_2(\lambda^*\mid\alpha^{(k)},x)}{\pi_2(\lambda^{(k-1)}\mid\alpha^{(k)},x)})$;

(iv) generate random variables u_1 and u_2 from the uniform distribution $U(0,1)$ and exercise decision-making judgment:

a) If $u_1\leq\varpi_{(\alpha_{k-1},\alpha^*)}$, then $\alpha^{(k)}=\alpha^*$, otherwise $\alpha^{(k)}=\alpha^{(k-1)}$,

b) if $u_2\leq\varpi_{(\lambda_{k-1},\lambda^*)}$, then $\lambda^{(k)}=\lambda^*$, otherwise $\lambda^{(k)}=\lambda^{(k-1)}$;

(v) let $k=k+1$;

(vi) repeat (ii)–(iv) N times to get $\{\alpha^{(k)}\}$ and $\{\lambda^{(k)}\}$ $(k=1,2,\ldots,N)$;

(vii) discard the first N^* samples of $\{\alpha^{(k)}\}$ and $\{\lambda^{(k)}\}$ to mitigate the initialization bias, thus retaining the remaining samples for inference; and

(viii) compute the Bayesian estimate based on SELF of α, λ : $\hat{\alpha}_S=\frac{1}{N-N^*}\sum_{i=N^*+1}^{N}\alpha^{(i)}$, $\hat{\lambda}_S=\frac{1}{N-N^*}\sum_{i=N^*+1}^{N}\lambda^{(i)}$.

5.7. Simulation

In this section, we use the Monte Carlo simulation method and combine the five parameter estimation methods mentioned above to conduct a comprehensive simulation study. The simulation process of this paper adopts Matlab R2016a, and the relevant code is presented in the appendix. We do this by generating 1000 random samples of size $n=30,50,100,120$ that follow the WIE distribution with parameter values $(\alpha,\lambda)=(0.5,0.5)$ and $(\alpha,\lambda)=(0.25,1.5)$. Of course, the actual value of α is not necessarily less than 1. Other values of α can be easily incorporated. In order to improve the effectiveness of the Bayesian estimator, we first generate a sample set with the size of 1000 as historical experience data and choose the prior parameter as $(\sigma_1,\omega_2,\sigma_1,\omega_2)=(s,1,s,1)$, where s is the sample mean. The choice of prior parameters for the Bayesian estimation is based on the following considerations: We keep the prior mean equal to the sample mean, and then take the one with smaller prior variance, which makes the prior parameter contain more informative [43]. The 12,000 MCMC samples are generated by the M-H algorithm and the first 2000 MCMC samples are discarded as perturbation samples. The remaining 10,000 MCMC samples are used as formal samples for the Bayesian estimation calculations based on the MCMC method. The average estimate (AE), the absolute deviation (AB), the mean relative error (MRE), and the mean square error (MSE) are used to evaluate the performance of estimators and the mathematical expressions are shown in Table 4.

Table 4. Numerical evaluation criteria in Monte Carlo simulations.

Name	Formula of errors
AE	$\bar{\hat{\alpha}} = \frac{1}{1000}\sum_{i=1}^{1000}\hat{\alpha}_i$
AB	$AB_{\bar{\hat{\alpha}}} = \frac{1}{1000}\sum_{i=1}^{1000}\lvert\hat{\alpha}_i - \alpha\rvert$
MRE	$MRE_{\bar{\hat{\alpha}}} = \frac{1}{1000}\sum_{i=1}^{1000}\left\lvert\frac{\hat{\alpha}_i - \alpha}{\alpha}\right\rvert$
MSE	$MSE_{\bar{\hat{\alpha}}} = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\alpha}_i - \alpha)^2$

Tables 5 and 6 illustrate the calculation results of AE, AB, MRE, and MSE under different sample sizes and two couple of values of (α, λ). Here, E denotes "$\times 10^{\wedge}$" in scientific notation.

The following findings can be found in Tables 5 and 6:

(1) All five estimation methods provide accurate estimates with low AB, MSE, and MRE values.

(2) Relatively, the CVM and OLS estimates are not sensitive enough to the parameter estimation of the WIE distribution.

(3) In the Bayesian estimation, it is observed through simulation experiments that the Lindley approximation and the MCMC method give approximate results for parameter estimation. However, from the computational cost point of view, the computational overhead of the MCMC method is significantly higher.

(4) In most cases, the sample size is inversely proportional to the estimated error. With the increase of the sample size, the smaller the estimated error is, the closer the estimated value is to the actual value.

(5) According to the parameter estimates of the WIE distribution, the ML and the Bayesian estimation had lower AB, MSE, and MRE values than other estimation methods. However, it is necessary to mention that the Bayesian estimation will be more computationally expensive than the ML estimation when the prior information about the parameters is uncertain.

Table 5. Classical and Bayesian estimates of α and λ when $(\alpha,\lambda)=(0.5,0.5)$.

n		ML		AD		CVM		OLS		Lindley		MCMC	
		α	λ	α	λ	α	λ	α	λ	α	λ	α	λ
30	AE	0.5217	0.5839	0.1176	0.1416	1.60E-08	0.1767	2.14E-08	0.1759	0.6102	0.6825	0.2585	0.4455
	AB	0.0755	0.1578	0.4379	0.3994	0.5	0.3233	0.5	0.3241	0.1438	0.4115	0.2416	0.0977
	MSE	0.0098	0.0623	0.2583	0.1998	0.25	0.1047	0.25	0.1055	0.0539	0.8285	0.2508	0.1201
	MRE	0.1509	0.3155	0.8758	0.7988	1	0.6466	1	0.6482	0.2876	0.8229	0.4831	0.1954
50	AE	0.5274	0.5556	0.1499	0.1729	1.72E-08	0.1768	1.74E-08	0.1762	0.5729	0.5944	0.2616	0.4216
	AB	0.0646	0.1662	0.4188	0.3826	0.5	0.3232	0.5	0.3239	0.092	0.2343	0.2386	0.12
	MSE	0.0077	0.057	0.243	0.195	0.25	0.1047	0.25	0.1051	0.0336	0.3602	0.2476	0.1462
	MRE	0.1292	0.3324	0.8375	0.7652	1	0.6463	1	0.6477	0.184	0.4686	0.4772	0.24
100	AE	0.5239	0.5794	0.1759	0.2093	1.71E-08	0.1769	1.68E-08	0.177	0.5308	0.5339	0.2641	0.4716
	AB	0.0483	0.1247	0.4003	0.3495	0.5	0.3231	0.5	0.323	0.0488	0.1316	0.2359	0.0678
	MSE	0.0047	0.0473	0.2441	0.1767	0.25	0.1049	0.25	0.1046	0.0057	0.1091	0.2487	0.084
	MRE	0.0965	0.2493	0.8006	0.699	1	0.6462	1	0.646	0.0976	0.2631	0.4718	0.1356
120	AE	0.5268	0.5827	0.1838	0.2204	2.39E-08	0.179	2.44E-08	0.1784	0.5315	0.5538	0.279	0.4872
	AB	0.0454	0.1151	0.3799	0.3317	0.5	0.3211	0.5	0.3216	0.0459	0.1248	0.2225	0.0507
	MSE	0.004	0.036	0.2342	0.1772	0.25	0.104	0.25	0.1042	0.0045	0.0893	0.2378	0.0633
	MRE	0.0907	0.2302	0.7597	0.6634	1	0.6421	1	0.6432	0.0919	0.2495	0.445	0.1014

Table 6. Classical and Bayesian estimates of α and λ when $(\alpha,\lambda)=(0.25,1.5)$.

n		**ML**		**AD**		**CVM**		**OLS**		**Lindley**		**MCMC**	
		α	λ	α	λ	α	λ	α	λ	α	λ	α	λ
30	AE	0.1957	1.463	0.8565	1.259	0.6757	1.5551	0.8155	1.6169	0.1955	1.4183	0.4205	1.3282
	AB	0.3338	0.2235	0.8441	0.6345	0.835	0.9759	0.9552	1.1012	0.3051	0.2244	0.1712	0.1793
	MSE	0.0710	0.1325	0.3164	0.7550	0.7518	3.1239	0.8957	3.6870	0.0503	0.1790	0.0462	0.3285
	MRE	1.335	0.149	3.3762	0.423	3.34	0.6506	3.8208	0.7341	1.2203	0.1496	0.6848	0.1195
50	AE	0.2035	1.4352	0.8095	1.2288	0.3941	1.2708	0.5919	1.3824	0.1907	1.4268	0.405	1.3619
	AB	0.3015	0.1971	0.8117	0.6485	0.5725	0.7904	0.7520	0.9534	0.2808	0.1992	0.1579	0.1399
	MSE	0.2152	0.0662	1.2059	0.5199	1.7005	1.3058	2.5902	1.8269	0.1379	0.065	0.1739	0.1614
	MRE	1.206	0.1314	3.2469	0.4323	2.29	0.5269	3.0081	0.6356	1.1233	0.1328	0.6317	0.0933
100	AE	0.1905	1.4254	0.7219	1.1587	0.3149	1.1478	0.3181	1.1075	0.2033	1.433	0.3871	1.3887
	AB	0.2793	0.183	0.7241	0.64	0.5005	0.7589	0.4968	0.7917	0.2653	0.1737	0.1539	0.1129
	MSE	0.0438	0.0867	0.2557	0.7976	0.3410	1.6145	0.3354	1.6758	0.0338	0.0783	0.0426	0.1938
	MRE	1.117	0.122	2.8965	0.4265	2.0021	0.5059	1.9872	0.5278	1.061	0.1158	0.6158	0.0753
120	AE	0.1914	1.4219	0.6404	1.068	0.2167	1.0531	0.2749	1.0701	0.206	1.4302	0.3761	1.3976
	AB	0.2590	0.1749	0.6557	0.6743	0.4122	0.7131	0.4626	0.7742	0.2647	0.1739	0.1499	0.1085
	MSE	0.1205	0.0497	0.8855	0.6358	0.9142	0.861	1.1631	1.0222	0.131	0.0483	0.1667	0.1237
	MRE	1.0361	0.1166	2.6229	0.4495	1.6487	0.4754	1.8505	0.5161	1.0586	0.1159	0.5996	0.0723

6. Real application case

In this section, we use real data to test the model utility of the WIE distribution and its derivative

distribution generated by $\alpha = \lambda$. We fit the two distributions to the real dataset, use ML estimation for parameter estimation, and compare these results with the original efficient model.

6.1. Superiority of the WIE distribution

Dataset 1. This dataset comes from a two-arm experiment adopted from Efron [44] and Mudholkar et al. [45]. The survival times (in days) of the 31 patients are as follows:

37, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1557 and 1776.

Compare the WIE distribution to some common distributions, namely Modi inverse exponential distribution (MIE), the inverse generalized Weibull distribution (IGWD), generalized inverse generalized Weibull distribution (GIGWD), and the new two-parameter weighted-exponential distribution (WE). The corresponding PDFs are illustrated in Table 7.

Table 7. The corresponding PDFs for competitors of the WIE distribution (dataset 1).

Distribution	PDF	Parameter Range	Reference
MIE	$f(x) = \dfrac{\theta\gamma(\gamma+1)e^{\frac{\theta}{x}}}{(x\gamma e^{\frac{\theta}{x}} + x)^2}$	$\theta, \gamma > 0$	Bengalath and Punathumparambath [46]
IGWD	$f(x) = \alpha\beta\lambda^{\beta} e^{-\left(\frac{\lambda}{x}\right)^{\beta}} x^{-(\beta+1)}(1-e^{-(\frac{\lambda}{x})^{\beta}})^{\alpha-1}$	$\alpha, \beta, \lambda > 0$	Jain et al. [47]
GIGWD	$f(x) = \alpha\beta\gamma\lambda^{\beta} e^{-\gamma(\frac{\lambda}{x})^{\beta}} x^{-(\beta+1)}(1-e^{-\gamma(\frac{\lambda}{x})^{\beta}})^{\alpha-1}$	$\alpha, \beta, \lambda, \gamma > 0$	Jain et al. [47]
WE	$f(x) = \sigma^{-1}(\sigma+1)\gamma e^{-\gamma x}(1-e^{-\sigma\gamma x})$	$\sigma, \gamma > 0$	Elshahhat et al. [48]

The standard errors (St.ES) are used to reflect the reliability and stability of the estimation results. In order to judge how well these distributions fit the dataset, three selection criteria are used: Akaike information criteria (AIC), Bayesian information (BIC), and Kolmogorov-Smirnov (KS) statistics with its P-value. The expressions for the selection criteria are as follows:

(1) $AIC = -2\hat{l} + 2k$;

(2) $BIC = -2\hat{l} + k\log n$; and

(3) $KS = \max_{1\le i\le n} |F_{\exp}(x_i) - F(x_i)|$, (4) $P \approx 1 - F_K(\sqrt{n}KS)$.

Here, k represents the number of parameters contained in the model, $F_{\exp}(x)$ is the empirical density function of the dataset, and $F_K(x)$ is the PDF of the Kolmogorov distribution.

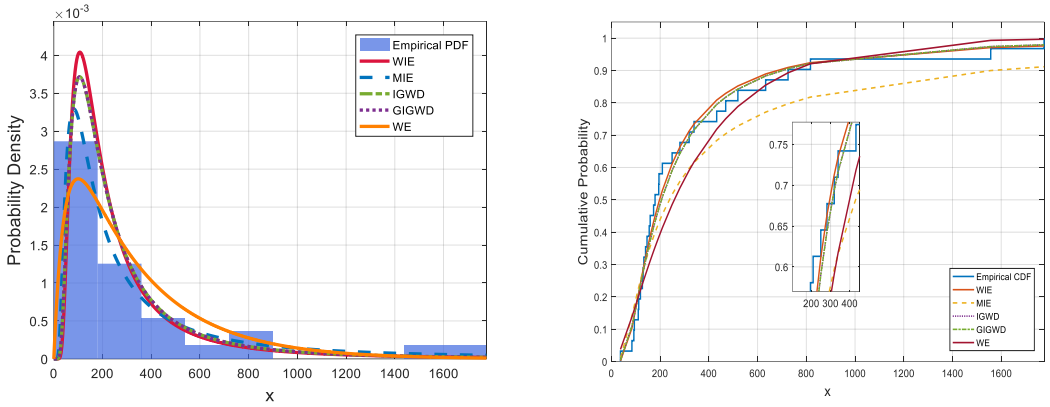
Table 8 shows the results of fitting these distributions to dataset 1. Table 8 shows that the WIE

distribution has lower AIC, BIC, and KS values with a higher P-value; thus, we conclude that the newly proposed model is more suitable for dataset 1 than other models.

Table 8. The estimated parameters and selection criteria for each distribution of dataset 1 obtained by ML estimation.

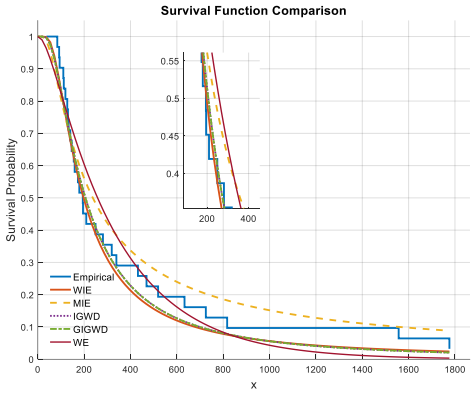
Model	Parameter	Estimate (St.ES)	l	AIC	BIC	KS(P-value)
WIE	α	297.3148(111.4426)	-206.5092	417.0184	419.8864	0.086625(0.9586)
	λ	318.103(54.7037)				
MIE	γ	9.8112(16.9085)	-208.5024	421.0048	423.8727	0.13936(0.5379)
	θ	167.1524(25.1004)				
IGWD	α	2.0354(1.7838)	-206.2951	418.5903	422.8922	1.1059(4.3500e-35)
	β	0.94678(0.4123)				
	λ	258.324(203.1935)				
GIGWD	α	2.0354(1.7838)	-206.2951	420.5903	426.3262	1.1059(4.3500e-35)
	β	0.94678(0.4123)				
	λ	25.8399(1.48e+7)				
	γ	8.8443(4.79e+4)				
WE	σ	5.8134(4.1426)	-209.8054	423.6109	426.4788	0.19682(0.1581)
	γ	0.0033(0.0006)				

At the same time, the fitting of the WIE distribution to this dataset can be further judged by the experience function, fitting density function, fitting distribution function, survival function, and probability–probability (PP) plot drawn in Figure 6. Dataset 1 shows that right skewness is more suitable for the WIE distribution. Figure 6(e) presents a contour plot of the log-likelihood function of the WIE distribution, which demonstrates the existence and uniqueness of the ML estimation α and λ based on dataset 1.

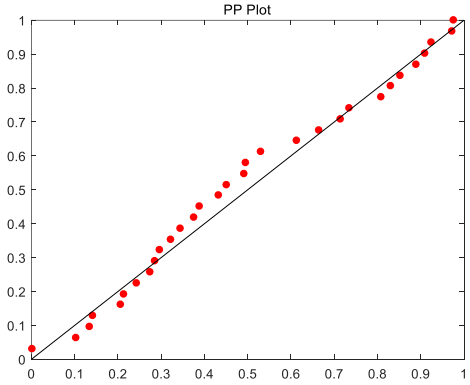


(a)

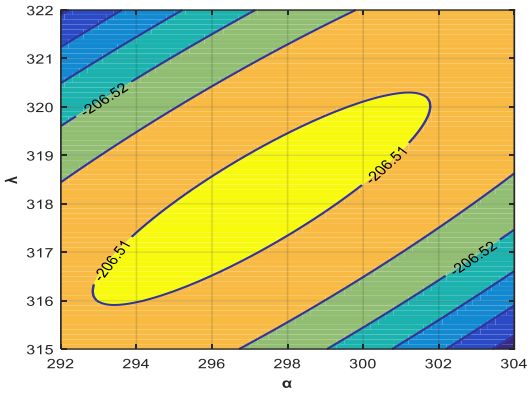
(b)



(c)



(d)



(e)

Figure 6. (a) Fitted PDF, (b) Fitted CDF, (c) Kaplan-Meier survival function, (d) PP plot of the WIE distribution, (e) Contour of the log-likelihood function of the WIE distribution for dataset 1.

6.2. Superiority of the SWIE distribution

This section explores the fact that the SWIE distribution is also well characterized. The good properties of the SWIE distribution are verified as a derivative distribution of the WIE distribution for the parameter $\alpha = \lambda$, which also confirms the flexibility of the WIE distribution.

Dataset 2. The dataset has been previously applied by Eldessouky et al. [49] for information on the time it takes for an infection to develop in individuals receiving renal dialysis. The data are shown below:

12.5, 13.5, 3.5, 3.5, 7.5, 12.5, 3.5, 9.5, 10.5, 11.5, 7.5, 14.5, 2.5, 2.5, 7.5, 8.5, 14.5, 21.5, 4.5, 5.5, 6.5, 6.5, 7.5, 25.5, 27.5, 21.5, 22.5 and 22.5.

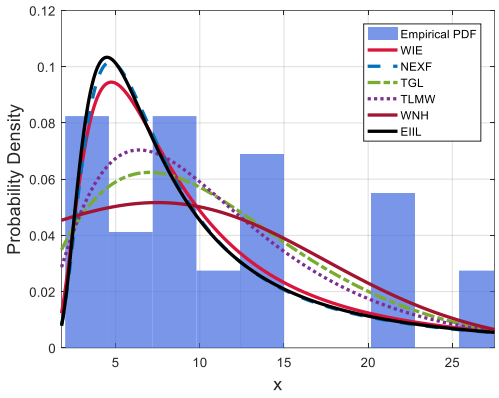
The newly proposed models, which demonstrate commendable performance, include the new exponential-X Fréchet distribution (NEXF), the new transmuted generalized Lomax distribution (TGL), the Topp–Leone modified Weibull model (TLMW), the weighted Nadarajah–Haghighi distribution (WNH), and the extended exponentiated inverse Lindley distribution (EIIL). The PDF of these models are shown in Table 9. These models are fitted to dataset 2 in a simultaneous manner with the WIE distribution, and the resulting fits are shown in Table 10. To demonstrate the fit of each distribution, Figure 7 visualizes the density overlay, cumulative distribution overlay, and survival overlay plots of each distribution for dataset 2, the PP plot of the SWIE distribution, and the plot of the log-likelihood function of the SWIE distribution based on dataset 2. Figure 7 shows that the SWIE distribution fits dataset 2 better and that the solution to the log-likelihood function under dataset 2 exists and is unique. This suggests that the WIE distribution has some advantages in dealing with right-skewed data.

Table 9. The corresponding PDFs for competitors of the SWIE distribution (dataset 2).

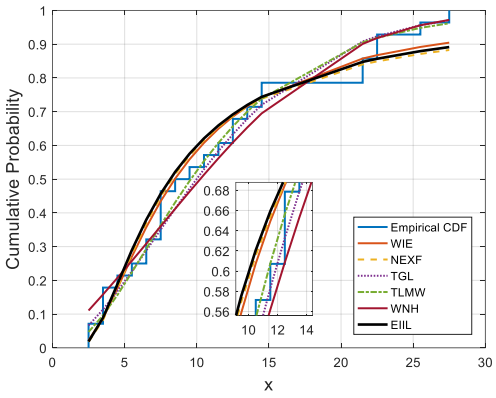
Distribution	PDF	Parameter Range	Reference
NEXF	$g(x) = \dfrac{\frac{\lambda}{x}\left(\frac{\alpha}{x}\right)^{\lambda} e^{-\left(\frac{\alpha}{x}\right)^{\lambda}}\left[1+\beta\left(1-e^{-\left(\frac{\alpha}{x}\right)^{\lambda}}\right)\right]}{e^{\beta e^{-\left(\frac{\alpha}{x}\right)^{\lambda}}}}$	$\alpha, \lambda, \beta > 0$	Alzeley et al. [50]
TGL	$f(x) = \delta\vartheta\gamma\left(1+\gamma x\right)^{\delta\pi-1} e^{\frac{\vartheta}{\pi}\left[1-(1+\gamma x)^{\delta\pi}\right]}\left[1-\beta+2\beta e^{\frac{\vartheta}{\pi}\left[1-(1+\gamma x)^{\delta\pi}\right]}\right]$	$\pi, \beta, \delta, \gamma, \vartheta > 0$	Azm et al. [51]
TLMW	$f(x) = 2\lambda\left(\alpha+\theta\beta x^{\theta-1}\right) e^{-2\left(\alpha x+\beta x^{\theta}\right)}\left[1-e^{-2\left(\alpha x+\beta x^{\theta}\right)}\right]^{\lambda-1}$	$\alpha, \beta, \lambda, \theta > 0$	Alyami et al. [52]
WNH	$f(x) = \dfrac{2\delta\mu(1+\mu x)^{\delta-1} e^{1-(1+\mu x)^{\delta}}}{\left(1+e^{1-(1+\mu x)^{\delta}}\right)^2}$	$\delta, \mu > 0$	Elshahhat, and Mohammed [53]
EIIL	$f(x;\alpha,\beta,\gamma,\theta) = \dfrac{\alpha\theta\beta^2}{\beta+\gamma}\left(\dfrac{\gamma+x^{\alpha}}{x^{2\alpha+1}}\right) e^{-\frac{\beta\theta}{x^{\alpha}}}\left[\left(1+\dfrac{\gamma\beta}{(\gamma+\beta)x^{\alpha}}\right)\right]^{\theta-1}$	$\alpha, \beta, \theta, \gamma > 0$	Eltehiwy [54]

Table 10. The estimated parameters and selection criteria for each distribution of dataset 2 obtained by ML estimation.

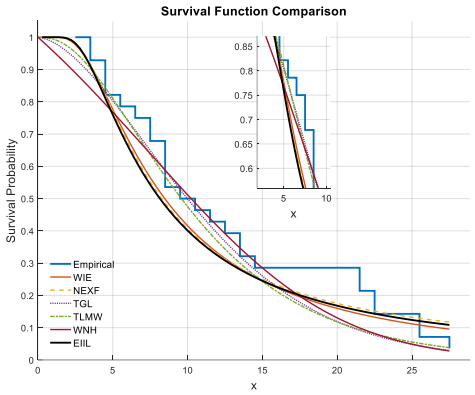
Model	Parameter	Estimate (St.ES)	l	AIC	BIC	KS(P-value)
SWIE	α	14.2247(1.9009)	-92.4404	186.881	188.213	0.0955(0.82597)
NEXF	α	8.8073(2.9669)	-93.6568	193.3137	197.3103	0.1172(0.7097)
	β	0.8203(0.9893)				
	λ	1.2142(0.2143)				
TGL	θ	0.00053(0.0016)	-91.3982	192.7964	199.4574	0.1209(0.8079)
	δ	0.5379(0.8528)				
	γ	11.0152(5.0147)				
	β	0.2829(0.6292)				
	π	3.2119(4.9239)				
TLMW	α	0.0761(0.0209)	-91.1122	190.2245	195.5533	0.1006(0.7711)
	β	0.001(0.0212)				
	λ	2.6787(0.8091)				
	θ	1(3.987e-07)				
WNH	δ	2.7013(2.7388)	-92.6109	189.2219	191.8863	0.10394(0.8119)
	μ	0.0309(0.0397)				
EIIL	α	1.4777(0.2072)	-93.2266	194.4532	199.7820	0.1086(0.6854)
	β	4.9327(5.1537)				
	γ	0.0001(4.2564)				
	θ	3.1207(4.0570)				



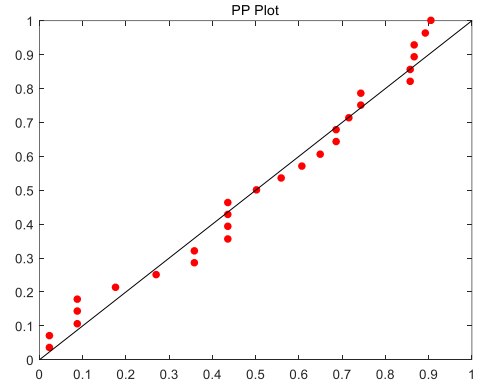
(a)



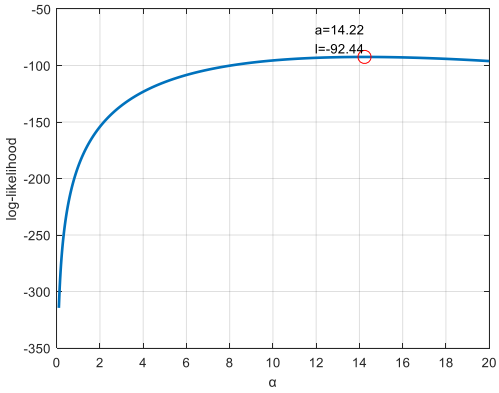
(b)



(c)



(d)



(e)

Figure 7. (a) Fitted PDF, (b) Fitted CDF, (c) Kaplan-Meier survival function, (d) PP plot (e) The log-likelihood function of the WIE distribution for dataset 1.

7. Conclusions

This paper presented a new weighted inverse exponential distribution, which has a simple mathematical expression and can be more flexible to adjust the tail shape and the central tendency and dispersion. A series of mathematical features of the WIE distribution were investigated, and by studying its PDF plots, tail area property, kurtosis, and skewness, it was concluded that the WIE distribution is a heavy-tailed, unimodal, and right-skewed distribution. Five entropy measures based on the distribution were systematically analyzed, and the conclusion was that the information distribution of the WIE distribution is concentrated, and the uncertainty is relatively small. Four classical estimators of the WIE distribution were inferred and the existence and uniqueness of the parameter estimates of the WIE distribution in real datasets and simulations were demonstrated using contour plots. A Bayesian estimator under the SELF was accomplished using the Lindley approximation and MCMC method. Through a comparative analysis, the ML and Bayesian estimation is more suitable for the parameter estimation of the proposed distribution, and the ML estimation has a smaller computational cost than the Bayesian estimation due to the uncertainty in the prior information of the parameters. Based on the real data of the survival time of patients and the time required for patients to become infected, various distributions with good properties were fitted to them. Through the AIC, BIC, and KS statistics with P-value selection criteria, it was concluded that the WIE distribution and its derivative distribution have relatively good fitting effects and better performances in analyzing the dataset.

In the future, we will explore the potential of the WIE distribution in other fields. In addition, we plan to study the fit of the WIE distribution under the censored sample and make a systematic comparison with the existing standard distribution. These works will lay a solid foundation for the promotion and popularization of the WIE distribution in practical applications.

Author contributions

Haiping Ren: Methods, writing-original draft, writing-review & editing; Jiajie Shi: Writing-original draft, method, software, writing-review & editing; Lianwu Yang: Method, software; Laijun Luo: Method, software, writing-review. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by National Natural Science Foundation of China, grant number 71661012 and the Foundation of Jxust (No.XJG-2020-27).

Conflict of interest

The authors declare no conflict of interest.

References

1. N. Alsadat, C. Taniş, L. P. Sapkota, A. Kumar, W. Marzouk, A. M. Gemeay, Inverse unit exponential probability distribution: classical and Bayesian inference with applications, *AIP Adv.*, **14** (2024), 055108. https://doi.org/10.1063/5.0210828
2. M. R. Mahmoud, A. E. Ismail, M. A. M. Ahmad, Exponential-inverse exponential [Weibull]: A new distribution, *Math. Stat.*, **11** (2023), 308–314. https://doi.org/10.13189/ms.2023.110209
3. M. Pizon, On the distribution of the product of inverse pareto and exponential random variables, *Recoletos Multidiscip. Res. J.*, **11** (2023), 27–31. https://doi.org/10.32871/rmrj2311.02.03
4. E. Chiodo, M. Fantauzzi, G. Mazzanti, The compound inverse Rayleigh as an extreme wind speed distribution and its Bayes estimation, *Energies*, **15** (2022), 861. https://doi.org/10.3390/en15030861
5. M. A. U. Haq, On Poisson moment exponential distribution with applications, *Ann. Data Sci.*, **11** (2024), 137–158. https://doi.org/10.1007/s40745-022-00400-0
6. J. Barahona, Y. Gómez, E. G. Déniz, O. Venegas, H. Gómez, Scale mixture of exponential distribution with an application, *Mathematics-Basel*, **12** (2024), 156. https://doi.org/10.3390/math12010156
7. A. Z. Keller, A. R. R. Kamath, U. D. Perera, Reliability analysis of CNC machine tools, *Reliab. Eng.*, **3** (1982), 449–473. https://doi.org/10.1016/0143-8174(82)90036-1
8. C. T. Lin, B. S. Duran, T. O. Lewis, Inverted gamma as a life distribution, *Microelectron. Reliab.*, **29** (1989), 619–626. https://doi.org/10.1016/0026-2714(89)90352-1
9. Q. Chen, W. Gui, Statistical inference of the generalized inverted exponential distribution under joint progressively type-II censoring, *Entropy*, **24** (2022), 576. https://doi.org/10.3390/e24050576
10. S. Ibrahim, B. Olalekan, H. Lawal, On the extended generalized inverse exponential distribution with its applications, *J. Probab. Stat.*, **7** (2020), 14–27. https://doi.org/10.9734/ajpas/2020/v7i330184
11. N. Alsadat, C. Taniş, L. P. Sapkota, A. Kumar, W. Marzouk, A. M. Gemeay, Inverse unit exponential probability distribution: Classical and Bayesian inference with applications, *AIP Adv.*, **14** (2024), 055108. https://doi.org/10.1063/5.0210828
12. M. N. Khan, A. Saboor, G. M. Cordeiro, M. Nazir, R. R. Pescim, A weighted Nadarajah and Haghighi distribution, *UPB Sci. Bull. Ser. A Appl. Math. Phys.*, **80** (2018), 133–140.
13. G. P. Patil, C. R. Rao, Weighted distributions and size-biased sampling with applications to wild life populations and human families, *Biometrics*, **34** (1978), 179–189. https://doi.org/10.2307/2530008
14. S. Dey, Inverted exponential distribution as a life distribution model from a Bayesian viewpoint, *Data Sci. J.*, **6** (2007), 107–113. https://doi.org/10.2481/dsj.6.107
15. G. Lugosi, S. Mendelson, Mean estimation and regression under heavy-tailed distributions: A survey, *Found. Comput. Math.*, **19** (2019), 1145–1190. https://doi.org/10.1007/s10208-019-09427-x
16. J. Nair, A. Wierman, B. Zwart, *The fundamentals of heavy-tails: Properties, emergence, and identification*, USA, New York: Association for Computing Machinery, **41** (2013), 387–388. https://doi.org/10.1145/2465529.2466587
17. R. Leipus, J. Šiaulys, S. Danilenko, J. Karaseviciene, Randomly stopped sums, minima and maxima for heavy-Tailed and light-tailed distributions, *Axioms*, **13** (2024), 355. https://doi.org/10.3390/axioms13060355

18. K. Cooray, Generalized Gumbel distribution, *J. Appl. Stat.*, **37**(2009), 171–179. https://doi.org/10.1080/02664760802698995
19. J. Moors, A quantile alternative for kurtosis, *J. Roy. Stat. Soc. D.*, **37** (1988), 25–32. https://doi.org/10.2307/2348376
20. H. Ren, Q. Gong, X. Hu, Estimation of entropy for generalized Rayleigh distribution under progressively type-II censored samples, *Axioms*, **12** (2023), 776. https://doi.org/10.3390/axioms12080776
21. G. Lin, A. Lin, Y. Mi, Modified hierarchical time-delayed distribution entropy: An effective method for signal complexity measurement, *Nonlinear Dynam.*, **113** (2025), 1191–1207. https://doi.org/10.1007/s11071-024-10297-4
22. M. Skorski, Towards more efficient Renyi entropy estimation, *Entropy*, **25** (2023), 185. https://doi.org/10.3390/e25020185
23. G. Singh, R. Khosa, M. K. Jain, T. Moramarco, V. P. Singh, Influence of the channel bed slope on Shannon, Tsallis, and Renyi entropy parameters, *J. Hydroinform.*, **25** (2023), 2195–2209. https://doi.org/10.2166/hydro.2023.008
24. Q. Tran, J. Kukal, Renyi entropy based design of heavy tailed distribution for return of financial assets, *Physica A*, **637** (2024), 129531. https://doi.org/10.1016/j.physa.2024.129531
25. M. Mesfioui, M. Kayid, M. Shrahili, Renyi entropy of the residual lifetime of a reliability system at the system level, *Axioms*, **12** (2023), 320. https://doi.org/10.3390/axioms12040320
26. C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, *J. Stat. Phys.*, **52** (1988), 479–487. https://doi.org/10.1007/BF01016429
27. I. Nikoufar, M. Fazlolahi, Some bounds for the generalized tsallis relative operator entropy, *Mediterr. J. Math.*, **20** (2023), 115. https://doi.org/10.1007/s00009-023-02312-6
28. A. Ghosh, P. Chaudhuri, Generalized position and momentum Tsallis entropies, *Int. J. Theor. Phys.*, **39** (2000), 2423–2438. https://doi.org/10.1023/A:1026432919521
29. Y. Shi, Y. Wu, P. Shang, Research on weighted Havrda-Charvat's entropy in financial time series, *Physica A*, **572** (2021), 125914. https://doi.org/10.1016/j.physa.2021.125914
30. Z. Wang, P. Shang, Analysis of the dispersion Havrda–Charvat entropy plane in financial time series, *Int. J. Bifurcat. Chaos*, **32** (2022), 2250234. https://doi.org/10.1142/S0218127422502340
31. T. Brochet, J. L. Lahorgue, S. Bougleux, M. Salaün, S. Ruan, Deep learning using Havrda-Charvat entropy for classification of pulmonary optical endomicroscopy, *IRBM*, **42** (2021), 400–406. https://doi.org/10.1016/j.irbm.2021.06.006
32. Y. Q. Wu, Y. L. Long, Line intercept histogram-based Arimoto entropy or Arimoto gray entropy for food image segmentation, *Mod. Food. Sci. Tech.*, **32** (2016), 164–169. https://doi.org/10.13982/j.mfst.1673-9078.2016.1.026
33. N. Alotaibi, A. F. Hashem, I. Elbatal, S. A. Alyami, A. S. A. Moisheer, M. Elgarhy, Inference for a Kavya-Manoharan inverse length biased exponential distribution under progressive-stress model based on progressive Type-II censoring, *Entropy*, **24** (2022), 1033. https://doi.org/10.3390/e24081033
34. B. Li, R. Li, Z. Liu, C. Li, Z. Wang, An objective non-eeference metric based on Arimoto entropy for assessing the quality of fused images, *Entropy*, **21** (2019), 879. https://doi.org/10.3390/e21090879
35. H. Zhang, J. L. Fan, Two-dimensional Arimoto entropy linear-type threshold segmentation method, *Acta Photonica Sinic.*, **42** (2013), 234–240. https://doi.org/10.3788/gzxb20134202.0234
36. R. Maya, M. Irshad, Kernel estimation of Mathai-Haubold entropy and residual Mathai-Haubold entropy functions under α-Mixing dependence condition, *Am. J. Math. Manag. Sci.*, **41** (2022), 148–159. https://doi.org/10.1080/01966324.2021.1935366

37. I. M. Almanjahie, J. G. Dar, A. I. A. Omari, A. Mir, Quantile version of Mathai-Haubold entropy of order statistics, *Cmes-Comp. Model. Eng.*, **128** (2021), 907–925. https://doi.org/10.32604/cmes.2021.014896
38. R. M. Alotaibi, H. R. Rezk, A. Elshahhat, Analysis of generalized type-II progressively hybrid Lindley-exponential data and its modeling in physics, engineering, and management, *Aip. Adv.*, **14** (2024), 045213. https://doi.org/10.1063/5.0197082
39. M. Nagy, Expected Bayesian estimation based on generalized progressive hybrid censored data for Burr-XII distribution with applications, *Aip. Adv.*, **14** (2024), 015357. https://doi.org/10.1063/5.0184910
40. D. V. Lindley, Approximate Bayesian method, *Trab. Estad. Invest. Oper.*, **31** (1980), 223–237. https://doi.org/10.1007/BF02888353
41. A. R. E. Saeed, E. M. Almetwally, On algorithms and approximations for progressively type-I censoring schemes, *Stat. Anal. Data Min.*, **17** (2024), e11717. https://doi.org/10.1002/sam.11717
42. E. S. A. E. Sherpieny, E. M. Almetwally, H. Z. Muhammed, Bayesian and non-Bayesian estimation for the parameter of Bivariate generalized Rayleigh distribution based on Clayton Copula under progressive type-II censoring with random removal, *Sankhya A.*, **85** (2023), 1205–1242. https://doi.org/10.1007/s13171-021-00254-3
43. D. Kundu, Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring, *Technometrics*, **50** (2008), 144–154. https://doi.org/10.1198/004017008000000217
44. B. Efron, Logistic regression, survival analysis, and the Kaplan-Meier curve, *J. Am. Stat. Assoc.*, **83** (1988), 414–425. https://doi.org/10.1080/01621459.1988.10478612
45. G. S. Mudholkar, D. K. Srivastava, G. D. Kollia, A Generalization of the Weibull distribution with application to the analysis of survival data, *J. Am. Stat. Assoc.*, **91** (1996), 1575–1583. https://doi.org/10.1080/ 01621459.1996.10476725
46. J. Bengalath, B. Punathumparambath, A novel extension of inverse exponential distributions: A heavy-tailed model with upside down bathtub shaped hazard rate, *Reliab. Theory Appl.*, **18** (2023), 112–127. https://doi.org/10.24412/1932-2321-2023-476-112-127
47. K. Jain, N. Singla, S. Sharma, The Generalized inverse generalized weibull distribution and its properties, *J. Probab.*, 2024. https://doi.org/10.1155/2014/736101
48. A. Elshahhat, E. M. Almetwally, S. Dey, H. S. Mohammed, Analysis of WE parameters of life using adaptive-progressively type-II hybrid censored mechanical equipment data, *Axioms*, **12** (2023), 690. https://doi.org/10.3390/axioms12070690
49. E. A. Eldessouky, O. H. Hassan, B. Aloraini, I. Elbatal. Modeling to medical and economic data using: The transmuted power unit inverse Lindley distribution, *Alex. Eng. J.*, **113** (2025), 633–647. https://doi.org/10.1016/j.aej.2024.11.008
50. O. Alzeley, E. M. Almetwally, A. M. Gemeay, H. M. Alshanbari, E. H. Hafez, M. H. A. Moussa, Statistical inference under censored data for the new exponential-X Fréchet distribution: Simulation and application to Leukemia data, *Comput. Intel. Neurosc.*, **2021** (2021), 2167670. https://doi.org/10.1155/2021/2167670
51. W. S. A. E. Azm, E. M. Almetwally, S. N. A. Aziz, A. A. H. E. Bagoury, R. Alharbi, O. E. A. Kasem, A new transmuted generalized Lomax distribution: Properties and applications to COVID-19 data, *Comput. Intel. Neurosc.*, **2021** (2021), 5918511. https://doi.org/10.1155/2021/5918511
52. S. A. Alyami, I. Elbatal, N. Alotaibi, E. M. Almetwally, H. M. Okasha, M. Elgarhy, Topp–Leone modified Weibull model: Theory and applications to medical and engineering Data, *Appl. Sci.*, **12** (2022), 10431. https://doi.org/10.3390/app122010431

53. A. Elshahhat, H. S. Mohammed, Survival analysis and applications of weighted NH parameters using progressively censored data, *Symmetry*, **15** (2023), 735. https://doi.org/10.3390/sym 15030735
54. M. Eltehiwy, Extended exponentiated inverse Lindley distribution: Model, properties and applications, *J. Indian Soc. Prob. St.*, **20** (2019), 281–300. https://doi.org/10.1007/s41096-019-00068-5

Appendix

A. ML estimation of the WIE distribution

```
clear
tic
n = 100;
num_simulations = 1000;
estimated_parameters1 = zeros(num_simulations, 1);
estimated_parameters2 = zeros(num_simulations, 1);
for sim = 1:num_simulations
   rng('shuffle')
x = rand(1, n);
a1=0.25;b1=1.5;
x_values=-a1.*b1./(b1+a1.*lambertw(-1,-exp(-b1./a1).*x.*b1./a1));
estimated_parameters1(sim, :) = a;
estimated_parameters2(sim, :) = b;
end
mean_a =mean(estimated_parameters1);
mse_a = mean((estimated_parameters1 - a1).^2);
AB_a=mean(abs(estimated_parameters1-a1));
MRE_a=mean(abs((estimated_parameters1 - a1)./a1));
mean_b =mean(estimated_parameters2);
mse_b = mean((estimated_parameters2 - b1).^2);
AB_b= mean(abs(estimated_parameters2-a2));
MRE_b=mean(abs((estimated_parameters2 - b1)./b1));
toc
```

B. Lindley approximation for Bayesian estimation

```
function [a_BE,b_BE]=Lindley(n,x,a1,b1,a2,b2,a3,b3)
L11=@(a,b)-sum((b-x).^2./(a.*(b-x)+b.*x).^2);
L22=@(a,b) -sum((a+x).^2./(a.*(b-x)+b.*x).^2);
L12=@(a,b) sum(x.^2./(a.*(b-x)+b.*x).^2);
L21=L12;
L111=@(a,b) sum(2.*(b-x).^3./(a.*(b-x)+x.*b).^3);
L121=@(a,b) -sum(2.*x.^2.*(b-x)./(a.*(b-x)+b.*x).^3);
L211=L121;
L221=@(a,b)-sum(2.*x.^2.*(a+x)./(a.*(b-x)+x.*b).^3);
L122=L221;
L212=L221;
```

```
L222=@(a,b) sum(2.*(a+x).^3./(a.*(b-x)+x.*b).^3);
P1=@(a,b)((a1-1)./a)-b1;
P2=@(a,b)((a2-1)./b)-b2;
L11=L11(a3,b3);L12=L12(a3,b3);L21=L21(a3,b3);L22=L22(a3,b3);
L111=L111(a3,b3);L122=L122(a3,b3);L212=L122;L221=L122;L222=L222(a3,b3);L112=L121(a3,b
3);
P1=P1(a3,b3);P2=P2(a3,b3);
O=-1./(L11*L22-L12*L21)*([L22,-L12;-
L21,L11]);O11=O(1,1);O12=O(1,2);O21=O(2,1);O22=O(2,2);
U1=a3;
U1_1=1;
U1_2=0;U1_11=0;U1_22=0;U1_12=0;U1_21=0;
a_BE=U1+P1*O11+P2*O12+0.5*(O11^2*L111+O21*O22*L222+(O11*O22+2*O21^2)*L122+3*
O12*O11*L112);
U2=b3;
U2_1=1;
U2_2=0;U2_11=0;U2_22=0;U2_12=0;U2_21=0;
b_BE=U2+P1*O21+P2*O22+0.5*(O22^2*L222+O12*O11*L111+(O11*O22+2*O12^2)*L112+3*
O21*O22*L122);
end
```

C. MCMC method for Bayesian estimation

```
clear
tic
a0=0.25; b0=1.5; m=30;n=30;num=1000;
Ra=[zeros(1,m-1),n-m];
Rb=[floor((n-m)/2),zeros(1,m-2),(n-m-floor((n-m)/2))];
Rc=[n-m,zeros(1,m-1)];
a_BMH=zeros(num,1);
b_BMH=zeros(num,1);
S_BMH=zeros(num,1);
h_BMH=zeros(num,1);
aHPD=zeros(num,1);
bHPD=zeros(num,1);
SHPD=zeros(num,1);
hHPD=zeros(num,1);
i = 1;
while i <= num
G=rand(1,m);
for j=1:m
H(j)=G(j).^(1/(j+sum(Rc(m-j+1:m))));
end
for j=1:m
```

```
Z(j)=1-prod(H(m-j+1:m));
end
x=-a0.*b0./(b0+a0.*lambertw(-1,-exp(-b0./a0).*Z.*b0./a0));
[a_BMH(i,1),b_BMH(i,1),S_BMH(i,1),h_BMH(i,1),aHPD(i,1),bHPD(i,1),cp1,cp2,cp3,cp4]=BMH(t
,Rc,n,x,0.25,1.5);
c1=c1+cp1;c2=c2+cp2;c3=c3+cp3;c4=c4+cp4;
    i=i+1;
end
aBMH=mean(a_BMH)
mse_a=sqrt(mean((a_BMH-a0).^2))
MRE_a=mean(abs((a_BMH-a0)./a0))
bBMH=mean(b_BMH)
mse_b=sqrt(mean((b_BMH-b0).^2))
MRE_b=mean(abs((b_BMH-b0)./b0))
toc
```

D. Real Data Fitting

```
data=[12.5, 13.5, 3.5, 3.5, 7.5, 12.5, 3.5, 9.5, 10.5, 11.5, 7.5, 14.5, 2.5, 2.5, 7.5, 8.5, 14.5, 21.5, 4.5, 5.5,
6.5, 6.5, 7.5, 25.5, 27.5, 21.5, 22.5, 22.5];
[a,AIC, BIC, KS_statistic, l, p,SE]= ES(data,1);
disp(['Estimated parameters: a = ', num2str(a)]);
disp(['AIC: ', num2str(AIC)]);
disp(['BIC: ', num2str(BIC)]);
disp(['l: ', num2str(l)]);
disp(['KS statistic: ', num2str(KS_statistic)]);
disp(['p-value: ', num2str(p')]);
```