STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, November 2025, pp 2208–2227.
Published online in International Academic Press (www.IAPress.org)

A Three-Parameter Alpha Power Erlang Distribution for Modeling Lifetime Data

Maryam Aliabadi, Amir Hossein Ghatari*, Esmaile Khorram

Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran

Abstract The alpha power family of distributions introduces a new type of distribution by adding an extra parameter to its baseline distribution, making it highly suitable for lifetime data analysis. In this paper, we propose the three-parameter alpha power Erlang (APEr) distribution, where the baseline Erlang distribution is composed of the sum of multiple exponential distributions and is a special case of the gamma distribution. After defining the probability density and cumulative distribution functions of the APEr distribution, we present its theoretical properties, including quantiles, moments, and maximum likelihood estimation (MLE) of its parameters. Furthermore, using a simulation study, we verify the consistency of the parameter estimators and demonstrate the improvement in inference quality with increasing sample size. Finally, we compare the performance of the proposed distribution in fitting on two real-world datasets against the alpha power exponential, exponential, and Erlang distributions, demonstrating its superiority in these applications.

Keywords Alpha power distribution, Consistency, Goodness of fit, Lifetime data.

AMS 2010 subject classifications 62E10, 62N05

DOI: 10.19139/soic-2310-5070-2647

1. Introduction

Statistical distributions modeling lifetime data are widely applied in various fields such as reliability engineering, survival analysis, and queuing theory. These distributions are crucial for understanding and predicting the time-to-event data in systems or processes. By incorporating flexibility into their mathematical forms, they can adapt to various data patterns. Adding an extra parameter to an existing family of distribution functions is a common practice in statistical distribution theory. This approach often enhances the flexibility of the cumulative and probability density functions, making them more suitable for data analysis. Such modifications allow statisticians to better fit the model to empirical data, improving accuracy and interpretability in practical applications.
Although classical lifetime distributions such as the Gamma, Weibull, and Generalized Exponential are commonly used in reliability and survival analysis, they sometimes do not provide enough flexibility to model different shapes of hazard functions, especially non-monotonic patterns like the bathtub curve. The alpha power transformation (APT) is a general method that improves the flexibility of standard distributions by adding an extra shape parameter. Based on this idea, the proposed APEr distribution extends the Erlang model into a three-parameter version. As discussed in Section 3, this model allows for controlling skewness and kurtosis, and can represent various shapes of hazard rates. These features make the APEr distribution useful for modeling complex lifetime data.
In the last decade of the 20th century and the first decade of the 21st century, enhancing standard statistical distributions has become a fundamental aspect of statistical theory. One common approach involves introducing additional parameters to the existing distributions for creating new ones [8]. These modifications aim to improve

*Correspondence to: Amir Hossein Ghatari (Email: a.h.ghatari@aut.ac.ir). Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2025 International Academic Press

the analytical flexibility and adaptability of classical distributions, enabling them to effectively model complex data structures. For instance, [21] and [19] proposed methodologies for embedding new parameters into existing distributions to increase their versatility. Similarly, [11] introduced the concept of beta-generated distributions, where the basic distribution is beta, and the baseline distribution corresponds to the cumulative distribution function (CDF) of any continuous random variable. Building on this idea, [12] replaced the beta distribution with the Kumaraswamy distribution, offering an alternative approach to achieve greater tractability. Moreover, [3] advanced the development of the T-X family of continuous distributions by replacing the beta distribution's probability density function (PDF) with that of any continuous random variable and using a functional form of the CDF satisfying specific conditions. [14] provided an extensive review of various methods for generating univariate continuous distributions, offering insights into their theoretical and practical implications. For more recent years, [2] proposed a generalized version of the Kumaraswamy distribution and demonstrated its superiority in modeling various types of real-life data. [26] introduced the Balakrishnan-alpha-beta-skew-Laplace distribution, which incorporates shape and skewness parameters to better capture asymmetry and tail behavior. Moreover, [16] developed the Exponentiated-Gompertz-Marshall-Olkin-G family of distributions, offering greater modeling power through an advanced compounding technique.

[18] introduced the alpha power transformation (APT) as a flexible framework for generating univariate distributions, focusing on the alpha power-transformed exponential distribution as a special case. Building on this foundation, [20] provided a comprehensive analysis of the alpha power transformation family, exploring its theoretical properties and a range of potential applications. Subsequent advancements have led to various extensions and specialized applications of the APT framework. [22] proposed the alpha power extended inverted Weibull distribution, investigating its properties, inference methods, and practical uses. [24] studied the alpha power exponential distribution under progressive Type-II censored samples, presenting theoretical insights and practical implementations. [27] extended the framework by introducing the alpha power transformed inverse Lomax distribution, comparing estimation techniques, and highlighting its utility in real-world data analysis.

Recently, [5] introduced the Marshall-Olkin alpha power inverse Weibull distribution with Bayesian and non-Bayesian estimation methods. [23] proposed the alpha power inverse Weibull distribution, showcasing its effectiveness in modeling lifetime data, particularly gastric cancer cases. [25] further expanded the alpha power framework by integrating beta parameters, and applying the resulting distributions to exponential models. The Generalized Alpha Power Inverted Weibull (GAPIW) distribution was introduced as a novel probability model by [9], extending the alpha power family by compounding it with the inverted Weibull distribution. Their study explored its various statistical properties, including quantile function, moments, and stress-strength reliability, with a practical application to air pollution data. The Erlang distribution is the sum of several exponential distributions. The exponential distribution, being one of the most intriguing topics in lifetime distributions, motivates us to introduce the alpha power version of the Erlang distribution and investigate its properties.

In this paper, we introduce the alpha power version of the Erlang distribution (APEr), which is essentially a composition of multiple exponential distributions. Using the general structure of alpha power distributions, we derive the probability density function (PDF), cumulative distribution function (CDF), survival function, and hazard rate function for the APEr, a three-parameter distribution. Next, based on fundamental statistical definitions, we provide the methodology for calculating quantiles (particularly quartiles), the mean, variance, moments, order statistics, and other mathematical properties of the new distribution. Subsequently, we theoretically examine the maximum likelihood estimation (MLE) of the distribution parameters and discuss the lack of closed-form solutions for the estimates. To evaluate the performance of the distribution, we conduct a simulation study to assess the consistency of the MLEs and analyze estimation errors under various conditions of simulated data. The results indicate that the MLEs for the distribution parameters are consistent and converge to their true values as the sample size increases. Finally, we compare the performance of the proposed APEr distribution with other related distributions, such as the alpha power exponential distribution, Erlang distribution, and exponential distribution, using two real-world datasets.

In Section 2, we recall the definition of alpha power transformation and Erlang distribution. In Section 3, we introduce alpha power Erlang (APEr) distribution, then we provide mathematical aspects like quantiles, moments and etc. In Section 4, we study the performance of APEr using simulated and real datasets.

2. Alpha Power Distribution

Mahdavi and Kundu [18] introduced the alpha power transformation (APT), adding a parameter to a family of distributions. This parameter increases the flexibility of the CDF and PDF of an APT family, which can be expressed as

$$F_{\text{APT}}(x) = \begin{cases} \frac{\alpha^{F(x)}-1}{\alpha-1} & \text{if } \alpha > 0, \alpha \neq 1 \\ F(x) & \text{if } \alpha = 1 \end{cases} \tag{1}$$

where $F(x)$ represents the CDF of any continuous distribution, furthermore, it should be noted that $x > 0$ in this type of distribution makes it suitable for lifetime data.

The corresponding probability density function (PDF) can be written as:

$$f_{\text{APT}}(x) = \begin{cases} \frac{\log \alpha}{\alpha-1}\alpha^{F(x)}f(x) & \text{if } \alpha > 0, \alpha \neq 1 \\ f(x) & \text{if } \alpha = 1 \end{cases} \tag{2}$$

where $f(x)$ represents the PDF of any continuous distribution.

For $\alpha > 0$, $\alpha \neq 1$, Eq (2) can be expanded using the power series:

$$\alpha^z = \sum_{k=0}^{\infty} \frac{(\log \alpha)^k}{k!} z^k.$$

Thus, the APT family density can be easily verified that:

$$f_{\text{APT}}(x) = \frac{1}{(\alpha-1)} \sum_{k=0}^{\infty} \frac{(\log \alpha)^{k+1}}{k!} f(x)F(x)^k.$$

Mead et al.[20] considered $h_{k+1}(x) = (k+1)f(x)F(x)^k$ to represent the exponentiated-F (Exp-F) distribution with parameter $k+1$ (for $k \geq 0$). This distribution arises by exponentiating the baseline CDF, leading to flexible shapes useful in modeling skewed data. Thus, APT densities are linear combinations of Exp-F distributions.

$$f_{\text{APT}}(x) = \sum_{k=0}^{\infty} b_k h_{k+1}(x), \tag{3}$$

where $b_k = (\log \alpha)^{k+1}/[(\alpha-1)(k+1)!]$.

Hence, the structural features of the APT family can be obtained from Exp-F properties. The integration of (3) results in:

$$F_{\text{APT}}(x) = \sum_{k=0}^{\infty} b_k H_{k+1}(x), \tag{4}$$

where $H_{k+1}(x)$ is the CDF of the Exp-F family with parameter $k+1$.

2.1. Erlang Distribution

The Erlang distribution is a special case of the gamma distribution, when the shape parameter ω is a positive integer. It is named after A. K. Erlang, who introduced it in queuing systems [10]. The Erlang distribution has two parameters: the shape ω (a positive integer) and the rate λ (a positive real number). Its probability density function (PDF) is given by:

$$f(x;\omega,\lambda) = \frac{\lambda^{\omega}x^{\omega-1}e^{-\lambda x}}{\Gamma(\omega)} \quad \text{for } x > 0, \tag{5}$$

where $\Gamma(\cdot,\cdot)$ is Gamma function and $\Gamma(\omega) = (\omega-1)!$. Also the cumulative distribution function (CDF) is:

$$F(x;\omega,\lambda)=\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}=1-e^{-\lambda x}\sum_{n=0}^{\omega-1}\frac{(\lambda x)^n}{n!}. \tag{6}$$

Also, the lower incomplete Gamma function $\gamma(\cdot,\cdot)$ is defined as follows:

$$\gamma(\omega,\lambda x)=\int_0^{\lambda x}e^{-t}t^{\omega-1}dt.$$

Note that the exponential distribution is Erlang with $\omega=1$. In essence, the Erlang distribution models the sum of ω independent, exponentially distributed random variables, making it useful for modeling multiple stages of processes or waiting times.Its expected value and variance are:

$$E(X)=\frac{\omega}{\lambda}$$

and

$$\text{Var}(X)=\frac{\omega}{\lambda^2}$$

Applications of the Erlang distribution are common in telecommunications, reliability engineering, and queuing theory, where it models service times and lifetimes of systems. Its relationship with the gamma distribution allows broader flexibility when ω is not an integer, but the Erlang form remains fundamental in practical scenarios involving integer event [17, 1, 4].

3. Alpha Power Erlang (APEr) Distribution

Before presenting the formal definition of the APEr distribution, it is helpful to briefly explain the intuition behind the model. The Erlang distribution is widely used for modeling waiting times in queuing systems and reliability analysis. However, it lacks flexibility in capturing various hazard rate shapes. To address this, we introduce an additional shape parameter using the alpha power transformation, which enables the model to adapt to a wider range of data behaviors. This extension results in the Alpha Power Erlang (APEr) distribution.

In this section, we use the APT method on the Erlang distribution by substituting (5) and (6) into (1) and (2). The resulting distribution, known as the alpha power Erlang (APEr), has three parameters that $\lambda(>0)$ is the rate parameter, and ω and α are shape parameters. Thus random variable X following the APEr distribution will be denoted by $X\sim\text{APEr}(\alpha,\omega,\lambda)$ and X has the PDF and CDF as follows:

$$F_{APEr}(x)=\begin{cases}\frac{\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}-1}{\alpha-1} & \text{if } \alpha>0,\alpha\neq 1\\ \frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)} & \text{if } \alpha=1\end{cases} \tag{7}$$

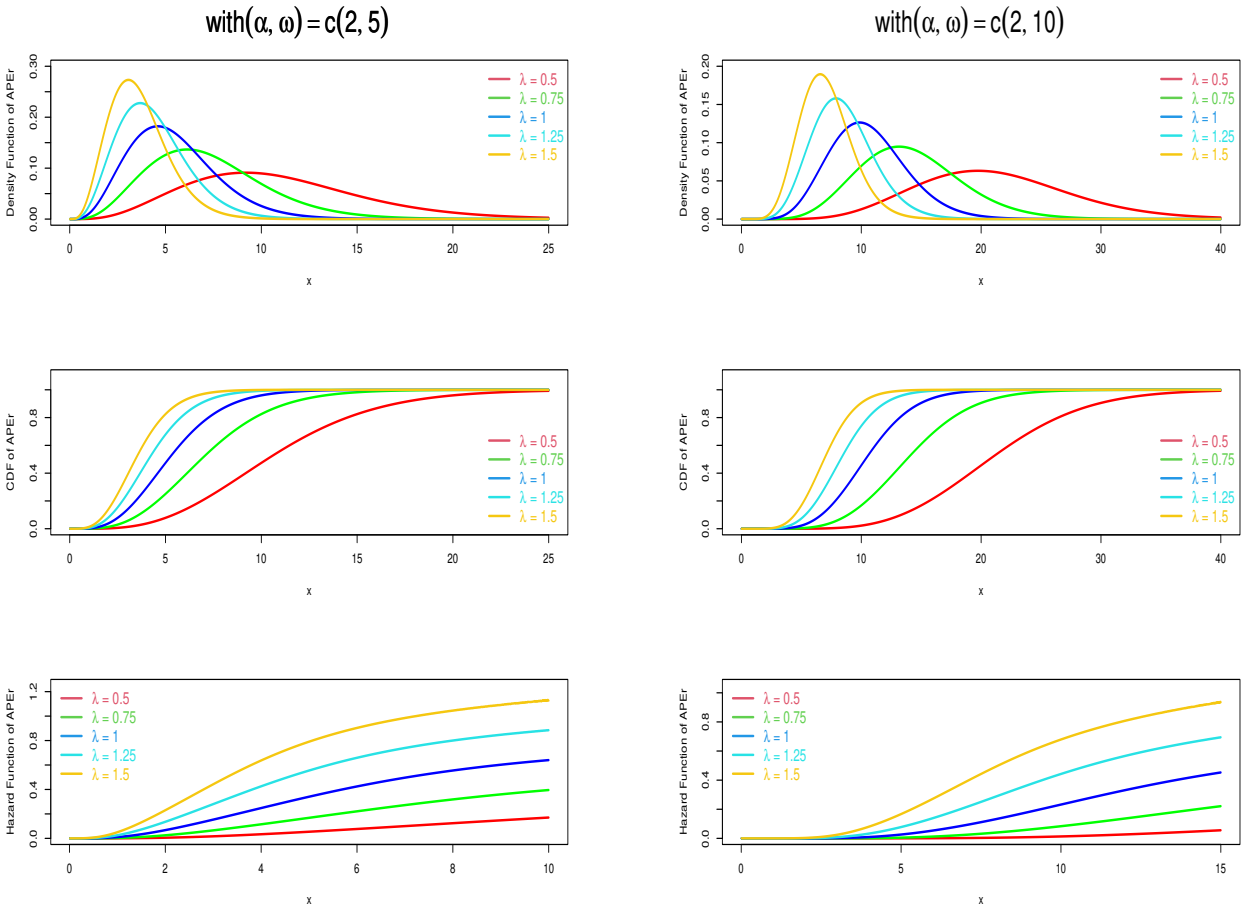
$$f_{APEr}(x)=\begin{cases}\frac{(\log\alpha)\lambda^\omega}{(\alpha-1)\Gamma(\omega)}x^{\omega-1}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}} & \text{if } \alpha>0,\alpha\neq 1\\ \frac{\lambda^\omega x^{\omega-1}e^{-\lambda x}}{\Gamma(\omega)} & \text{if } \alpha=1\end{cases} \tag{8}$$

This PDF introduces a new parameter α which modulates the tail behavior of the distribution. When $\alpha=1$, the distribution reduces to the standard Erlang. Increasing or decreasing α allows for greater or lesser skewness, which improves fit for datasets with varying shapes. The following is how the survival and hazard rate functions are determined, respectively:

$$S_{APEr}(x)=\begin{cases}\frac{\alpha-\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}}{\alpha-1} & \text{if } \alpha>0,\alpha\neq 1\\ 1-\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)} & \text{if } \alpha=1\end{cases} \tag{9}$$

$$h_{APEr}(x) = \begin{cases} \frac{(\log\alpha)\lambda^{\omega}}{\Gamma(\omega)} x^{\omega-1} e^{-\lambda x} \left(\frac{\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}}{\alpha - \alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}} \right) & \text{if } \alpha > 0, \alpha \neq 1 \\ \frac{\lambda^{\omega} x^{\omega-1} e^{-\lambda x}}{\Gamma(\omega) - \gamma(\omega,\lambda x)} & \text{if } \alpha = 1 \end{cases}$$

The hazard rate function of the APEr distribution exhibits diverse shapes including increasing, decreasing, and bathtub forms depending on the choice of parameters. This makes the APEr model suitable for a wide variety of lifetime data applications.



(a) PDF, CDF and HRF for $\omega = 5$ and $\alpha = 2$.

(b) PDF, CDF and HRF for $\omega = 10$ and $\alpha = 2$.

Figure 1. Side-by-side plots for various values of λ.

Figure 1 includes the PDF,CDF and HRF plots of the APEr distribution for $\omega = 5, 10$, $\alpha = 2$, and various values of λ. According to the plots of the figure, skewness and kurtosis both decrease as ω increases, while the other two parameters are constant. It is evident that the density range shrinks as ω decreases until it approaches zero, indicating an increase in kurtosis. Furthermore, λ has a direct impact on how skewed the distribution is; a higher λ equals a mo re skewed distribution. The distribution is extremely close to symmetry when the values of λ equal 0.5 and omega equal 5. When $\lambda = 1.5$ and $\omega = 5$, the condition is the most skewed and peaked. The density function's height reduces as ω grows, indicating that the dispersion rises and the concentration of frequencies involve the data mode.

Remark 1. It is important to note that the APEr distribution tends the Erlang distribution when the shape parameter $\alpha = 1$. consequently, all subsequent calculations will focus solely on the $\alpha > 0, \alpha \neq 1$ for the sake of brevity.

3.1. Quantile Function

Theorem 3.1. *Let $X \sim APEr(\alpha, \omega, \lambda)$, then p^{th} quantile is :*

$$x_p = \frac{\gamma^{-1}\left(\omega, \Gamma(\omega)\log_\alpha\left((\alpha-1)p+1\right)\right)}{\lambda}$$

where $0 < p < 1$ and $\gamma^{-1}(\cdot,\cdot)$ is the inverse of the incomplete Gamma function.

Proof
One definition of the quantile function is the inverse of the distribution function. Examine the identity

$$F(X) = U \Rightarrow X = F^{-1}(U)$$

then

$$U = \frac{\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}} - 1}{\alpha - 1}$$

and

$$\Gamma(\omega)\log_\alpha\left[(\alpha-1)U+1\right] = \gamma(\omega,\lambda x)$$

after simple mathematical calculations, Eq (13) is obtained. □

Using Theorem 3.1, the median of APEr can be calculated as follows:

$$x_{0.50} = \frac{\gamma^{-1}\left(\omega, \Gamma(\omega)\log_\alpha\left(\frac{1}{2}(\alpha+1)\right)\right)}{\lambda}$$

Additionally, the first quantile (25^{th} percentile) and the third quantile (75^{th} percentile) of the APEr are provided as:

$$Q_1 = x_{0.25} = \frac{\gamma^{-1}\left(\omega, \Gamma(\omega)\log_\alpha\left(\frac{1}{4}(\alpha+3)\right)\right)}{\lambda}$$

$$Q_3 = x_{0.75} = \frac{\gamma^{-1}\left(\omega, \Gamma(\omega)\log_\alpha\left(\frac{1}{4}(3\alpha+1)\right)\right)}{\lambda}.$$

3.2. Mode

The equation (10) can be solved to determine the distribution mode:

$$\frac{d}{dx}f_{APEr}(x) = \left\{\frac{(\log\alpha)\lambda^\omega}{(\alpha-1)\Gamma(\omega)}\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}x^{\omega-2}e^{-\lambda x}\right\}g(x) = 0 \tag{10}$$

The APEr distribution is unimodal and has a distinct mode at $x = x_0$ .When $\omega = 1$, $x_0 = \frac{\log(\log\alpha)}{\lambda}$ is the mode and when $\omega > 1$, the mode at x_0 is solution of $g(x) = 0$, where

$$g(x) = \frac{(\log\alpha)\lambda^\omega}{\Gamma(\omega)}x^\omega e^{-\lambda x} + (\omega-1) - \lambda x.$$

3.3. Moments

The moment generating function (MGF) $M_X(t) = \mathbb{E}(e^{tX})$ of APEr distribution can be expressed from (3) as

$$M_x(t) = \sum_{k=0}^{\infty}(k+1)b_k\tau(t,k), \tag{11}$$

where $\tau(t,k) = \int_0^1 \exp\left[tQ_F(u)\right] u^k du$ and

$$Q_F(u) = F^{-1}(v) = \frac{\gamma^{-1}(\omega, u\Gamma(\omega))}{\lambda}.$$

The moment generating function (MGF) of the APEr distribution is represented as an infinite weighted sum, where each term involves an integral of the form $\tau(t,k) = \int_0^1 e^{tQF(u)} u^k du$. Notably, the quantile function $Q_F(u)$ has an explicit form based on the inverse incomplete gamma function, allowing for numerical evaluation with high precision. This semi-analytic structure provides a tractable yet flexible approach to compute moments of various orders without resorting to purely simulation-based methods. By taking the derivative of (11), we can write the r^{th} order moment of APEr distribution as

$$\mu'_r = E\left(x^r\right) = \sum_{K=0}^{\infty} (k+1) b_k \delta_{r,k} \tag{12}$$

where

$$\delta_{r,k} = \int_0^1 \left[Q_F(u)\right]^r u^k = \int_0^1 \left(\frac{\gamma^{-1}(\omega, u\Gamma(\omega))}{\lambda}\right)^r u^k du$$

Therefore, it is possible to compute the expected value, second raw moment, and variance of the APEr distribution as

$$E(X) = \sum_{k=0}^{\infty} (k+1) b_k \delta_{1,k} \tag{13}$$

$$E\left(X^2\right) = \sum_{k=0}^{\infty} (k+1) b_k \delta_{2,k}$$

$$Var(x) = \sum_{k=0}^{\infty} \frac{(\log\alpha)^{k+1}}{(\alpha-1)k!} \delta_{2,k} - \left(\sum_{i,j=0}^{\infty} \frac{(\log\alpha)^{i+j+2}}{(\alpha-1)^2 i! j!} \delta_{1,i}\delta_{1,j}\right).$$

In this formulation, the shape parameter α directly affects the weighting coefficients b_k, which in turn influence the skewness and kurtosis of the distribution. As a result, the APEr distribution allows for rich tail behavior and moment structures that can adapt to different empirical data shapes. The measure of skewness (S_k) and kurtosis (K) for the APEr distribution are given, respectively, as follows

$$\begin{aligned} S_K &= \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1)^{\frac{3}{2}}}, \\ K &= \frac{\mu_4 - 4\mu_3\mu_1 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1)^2} \end{aligned} \tag{14}$$

where μ_1,μ_2,μ_3, and μ_4 are the 1^{th}, 2^{th}, 3^{th}, and 4^{th} moments obtained from (12), respectively.

Since closed-form expressions for the moments rely on infinite series and are not available necessarily, numerical methods are needed to approximate them. In this situation, one can approximate them using numerical methods. For example, the infinite series can be truncated after a sufficient number of terms, and the integrals can be evaluated using numerical quadrature techniques such as Gaussian quadrature. These methods usually provide accurate results with manageable computational effort. To implement such calculations, software tools like R or MATLAB can be used, as they offer built-in functions for numerical integration and summation.

3.4. Mean Residual Life and Mean Waiting Time

The mean residual life (MRL) and mean waiting time (MWT) functions provide practical interpretations of the distribution in reliability and survival contexts. The MRL function represents the expected remaining lifetime given survival up to time t, while the MWT describes the average time until failure from the perspective of time t.

The mean residual life (MRL) function is defined as the expected additional lifetime that a component has survived until time t, assuming that $X$is a continuous random variable with a survival function shown in (9). Consider that $\mu(t)$ is obtained from

$$\mu(t) = \frac{1}{p(X > t)} \int_0^\infty p(X > x)dx \qquad t \geqslant 0$$

$$\mu(t) = \frac{1}{S(t)} \left(E(t) - \int_0^t xf(x)dx \right) - t \qquad t \geqslant 0 \tag{15}$$

where

$$\int_0^t xf(x)dx = \frac{(\log \alpha)\lambda^\omega}{(\alpha - 1)\Gamma(\omega)} \xi_{\alpha,\omega,\lambda}(t) \tag{16}$$

and

$$\xi_{\alpha,\omega,\lambda}(t) = \int_0^t x^\omega e^{-\lambda x} \alpha^{\frac{\gamma(\alpha,\lambda x)}{\Gamma(\omega)}} dx$$

Thus, the MRL of APEr is obtained by substituting (13) and (15) into (14), as follows:

$$\mu(t) = \frac{\log \alpha}{\Gamma(\omega) \left(\alpha - \alpha^{\frac{\gamma(\omega,\lambda t)}{\Gamma(\omega)}} \right)} \left[\sum_{k=0}^{\infty} (k+1)(\log \alpha)\delta_{1,k} - \lambda^\omega \xi_{\alpha,\omega,\lambda}(t) \right] - t$$

Similarly, the mean waiting time (MWT) of X, $\bar{\mu}(t)$, can be defined as follows:
If it possesses the CDF (7):

$$\bar{\mu}(t) = t - \frac{1}{F(t)} \int_0^t xf(x)dx \tag{17}$$

where $\int_0^t xf(x)dx$ represents the first incomplete moment as defined by Eq(24). Thus, the MWT of the APEr can be obtained by substituting the Eqs (4), (7), and (16) in (17),

$$\bar{\mu}(t) = t - \frac{\lambda^\omega \xi_{\alpha,\omega,\lambda}{}^{(t)}}{\sum_{k=0}^{\infty} (\log \alpha)^k H_{k+1}(t)}.$$

While these expressions involve nested integrals and summations, they are particularly useful in engineering applications where understanding the aging or failure behavior of a system is important. The MRL function, for instance, is increasing when the system becomes more likely to survive longer as time passes, which is controlled by α and ω jointly.

3.5. Stochastic Ordering

Stochastic ordering is crucial for assessing and comparing the behavior of continuous random variables by establishing structured relationships between them.The pdf, cdf, and hrf of two positive random variables $X_i (i = 1, 2)$ are represented by $f_{X_i}(\cdot)$, $F_{X_i}(\cdot)$, and $h_{X_i}(\cdot)$, respectively. Then, the following definitions are given:

- If $\frac{f_{X_1}(x)}{f_{X_2}(x)}$ decreases in x then the likelihood ratio order $X_1 \leq_{(lro)} X_2$;
- If $h_{X_1}(x) \leq h_{X_2}(x)$ for all x then the hazard rate order $X_1 \leq_{(hro)} X_2$;
- If $F_{X_1}(x) \leq F_{X_2}(x)$ for all x then the stochastic order $X_1 \leq_{(sto)} X_2$.

A distribution that follows likelihood ratio order (lro) will also follow hazard rate (hro) and stochastic order (sto). Furthermore, if a distribution family exhibits likelihood ratio ordering, a uniformly most powerful test exists [6].

Theorem 3.2. *Let* $X_1 \sim APEr(\alpha_1, \omega, \lambda)$ *and* $X_2 \sim APEr(\alpha_2, \omega, \lambda)$. *If* $\alpha_1 < \alpha_2$, *then* $[X_1 \leq_{lro} X_2]$, $[X_1 \leq_{hro} X_2]$, *and* $[X_1 \leq_{sto} X_2]$.

Proof
For any x the likelihood ratio is given by

$$\frac{f_{X_1(x)}}{f_{X_2(x)}} = \left(\frac{\log \alpha_1}{\log \alpha_2}\right)\left(\frac{\alpha_{2-1}}{\alpha_{1-1}}\right) \quad \left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}$$

$$\frac{d}{dx}\left(\log\left(\frac{f_{x_1}(x)}{f_{x_2}(x)}\right)\right) = \frac{\left(\log\left(\frac{\alpha_1}{\alpha_2}\right)\right)\lambda^{\omega}}{\Gamma(\omega)} x^{\omega-1}e^{-\lambda x} < 0 \qquad \text{if } \alpha_1 < \alpha_2$$

Hence,

$$X_1 \leq_{(lro)} X_2$$

for all x , other ordering follow that

$$X_1 \leq_{(hro)} X_2 \Rightarrow X_1 \leq_{(sto)} X_2.$$

□

This result shows that increasing the α parameter leads to a stochastic shift in the distribution, making the random variable stochastically larger. This monotonic behavior in α is especially helpful when ranking systems or components by reliability, as it ensures consistent ordering under different model configurations.

3.6. Order statistics

The following theorem proposes the structure of obtaining order statistics in APEr distributions.

Theorem 3.3. *Let* $X_1, X_2, ..., X_n$ *be the observed values with sample size of* n *from APEr and* $X_{r:n}$ *denote* r^{th} *order statistics, then the density of* $X_{r:n}$ *stated as follows:*

$$f_{r:n}(x) = \sum_{k=0}^{\infty} c_k h_{k+1}(x)$$

where

$$c_k = \frac{(\log \alpha)^{k+1}}{(k+1)!} \sum_{j=0}^{s+r-1} \sum_{s=0}^{n-r} \frac{(-1)^{s+j}(j+1)^k}{B(r, n-r+1)(1-\alpha)^{s+r}} \binom{n-r}{s} \binom{s+r-1}{j}.$$

and $B(\cdot,\cdot)$ *refers to the Beta function.*

Proof
The density of $X_{r:n}$ can be defined as

$$f_{r:n}(x) = \frac{n!}{(r-1)(n-r)!} f_x(x) \left[F_x(x)\right]^{r-1} \left[1 - F_x(x)\right]^{n-r}$$

the following is an expansion of the binomial series:

$$(x-z)^n = \sum_{s=0}^{n} (-1)^s \binom{n}{s} x^{n-s} z^s \tag{18}$$

using the binomial series expansion as presented in (18), $f_{r:n}(x)$ can be expressed as follows

$$f_{r:n}(x) = \frac{f_{APEr}(x)}{B(r, n-r+1)} \sum_{s=0}^{n-r} (-1)^s \binom{n-r}{s} F_{APEr}(x)^{s+r-1}$$

for $f_{APEr} \cdot F_{APEr}(x)$, (18) can be derived as

$$f_{APEr}\, F_{APEr}(x)^{s+r-1} = \frac{\log \alpha}{(1-\alpha)^{s+r}} f(x) \sum_{j=0}^{s+r-1} (-1)^j \binom{s+r-1}{j} \left(\alpha^{F(x)}\right)^{j+1}$$

Considering the PDF and CDF of the APEr distribution in Eqs (7) and (8) and also applying the above equation, we can conclude that:

$$f_{APEr}(x) F_{APEr}(x)^{s+r-1} = \sum_{j=0}^{s+r-1} \sum_{k=0}^{\infty} \frac{(-1)^j (j+1)^k (\log \alpha)^{k+1}}{(1-\alpha)^{s+r} k!} \binom{s+r-1}{j} \left(\frac{\lambda^\omega x^{\omega-1} e^{-\lambda x}}{\Gamma(\omega)}\right) \left(\frac{\gamma(\omega, \lambda x)}{\Gamma(\omega)}\right)^k.$$

□

3.7. Rényi and Shannon Entropies

Entropy measures the variability or uncertainty associated with a random variable, X. Two widely used entropy measures are Rényi and Shannon entropies.

The Rényi entropies for arandom variable X, which has a probability density function $f(x)$, is defined as follows:

$$RE_X(\nu) = \frac{1}{1-\nu} \log \left(\int_0^x f(x)^\nu dx \right); \nu > 0, \nu \neq 1$$

using the PDF specified in Eq (10), $RE_X(\nu)$ has no closed-form solution and cannot be caculated with mathematical formulas; it can be solved using numerical methods.

Theorem 3.4. *It can be demonstrated that the Shannon entropy of random variable X, assuming it has an APEr distribution, is as follows:*

$$SE_X = \log \left(\frac{\lambda^\omega \log \alpha}{(\alpha-1)\Gamma(\omega)} \right) + \left(\frac{\lambda^\omega \log \alpha}{(\alpha-1)\Gamma(\omega-1)} \right) \left(\zeta_{\alpha,\omega,\lambda} - \frac{\lambda}{\omega} \eta_{\alpha,\omega,\lambda} \right) + \Gamma(\omega) \left(1 - \frac{\log \alpha}{(\alpha-1)} \right) \tag{19}$$

where

$$\zeta_{\alpha,\omega,\lambda} = \int_0^\infty (\log x)\, x^{\omega-1} e^{-\lambda x} \alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}} dx,$$

and

$$\eta_{\alpha,\omega,\lambda} = \int_0^\infty x^\omega e^{-\lambda x} \alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}} dx.$$

Proof

The Shannon entropy SE_X of random variable X with PDF $f(x)$ is obtained as

$$SE_X = E[-\log f(x)] = -\int_0^\infty \log\left(f(x)\right) f(x)dx. \tag{20}$$

Inserting (8) into (20) can be derived as

$$\begin{aligned}
SE_x = & \\
&\log\left(\frac{\log\alpha}{\alpha-1}\right)\left(\frac{\lambda^\omega\log\alpha}{(\alpha-1)\Gamma(\omega)}\right)\int_0^\infty x^{\omega-1}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda(x)}{\Gamma(\omega)}}dx + \frac{\lambda^\omega(\log\alpha)^2}{(\alpha-1)\Gamma(\omega)}\int_0^\infty \gamma(\omega,\lambda x)x^{\omega-1}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda(x)}{\Gamma(\omega)}}dx \\
&+ \frac{\lambda^\omega\log\lambda^\omega\log\alpha}{(\alpha-1)\Gamma(\omega)}\int_0^\infty x^{\omega-1}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda(x)}{\Gamma(\omega)}}dx + \frac{\lambda^\omega\log\alpha}{(\alpha-1)\Gamma(\omega-1)}\int_0^\infty (\log x)\, x^{\omega-1}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda(x)}{\Gamma(\omega)}}dx \\
&- \frac{\lambda^{\omega+1}\log\alpha}{(\alpha-1)\Gamma(\omega)}\int_0^\infty x^{\omega}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda(x)}{\Gamma(\omega)}}dx - \frac{\lambda^\omega\log(\Gamma(\omega))\log\alpha}{(\alpha-1)\Gamma(\omega)}\int_0^\infty x^{\omega-1}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda(x)}{\Gamma(\omega)}}dx
\end{aligned}$$

after some simplification, we can easily reach to Equation (19). □

Remark 2. Note that

$$\eta_{\alpha,\omega,\lambda} = \lim_{t\to\infty}\xi_{\alpha,\omega,\lambda}(t) = \int_0^\infty x^{\omega}e^{-\lambda x}\alpha^{\frac{\gamma(\omega,\lambda x)}{\Gamma(\omega)}}dx.$$

3.8. Maximum likelihood estimation

In this section, we determine the maximum likelihood estimates (MLEs) for the parameters of APEr distribution. Assume $X_1, X_2, ..., X_m$ be random sample from the APEr distribution with parameters in $\Theta = (\alpha, \omega, \lambda)$. Then likelihood function for Θ reduces to

$$L = \prod_{i=1}^m f\left(x_i;\alpha,\omega,\lambda\right)\left(\frac{\alpha\lambda^\omega(\log\alpha)}{(\alpha-1)\Gamma(\omega)}\right)^m \alpha^{-\sum_{n=0}^{\omega-1}\sum_{i=1}^m \frac{1}{n!}e^{-\lambda x_i}(\lambda x_i)^n}\left(\prod_{i=1}^m x_i\right)^{\omega-1} e^{-\lambda\sum_{i=1}^m x_i} \tag{21}$$

by logarithmic transformation, Equation (21) changes to

$$l = l(\Theta) = m\log\left(\frac{\alpha\lambda^\omega(\log\alpha)}{(\alpha-1)\Gamma(\omega)}\right) - \log\alpha\sum_{n=0}^{\omega-1}\sum_{i=1}^{\infty}\frac{1}{n!}e^{-\lambda x_i}\left(\lambda x_i\right)^n + (\omega-1)\sum_{i=1}^m\log x_i - \lambda\sum_{i=1}^m x_i. \tag{22}$$

When we take partial derivatives of the log-likelihood (22) in relation to the parameters and set the results to zero, we get

$$\frac{\partial l}{\partial\alpha} = \frac{m}{\alpha} + \frac{m}{\alpha\log\alpha} - \frac{m}{\alpha-1} - \frac{1}{\alpha}\left\{\sum_{n=0}^{\omega-1}\sum_{i=1}^m\frac{1}{n!}e^{-\lambda x_i}\left(\lambda x_i\right)^n\right\} = 0, \tag{23a}$$

$$\frac{\partial l}{\partial\omega} = m\log\lambda - m\psi(\omega) - \frac{\log\alpha}{\Gamma(\omega)}\left\{\sum_{i=1}^m e^{-\lambda x_i}\left(\lambda x_i\right)^{\omega-1}\left[\log\left(\lambda x_i\right) - \psi(\omega)\right]\right\} + \sum_{i=1}^m\log\left(x_i\right) = 0, \tag{23b}$$

$$\frac{\partial L}{\partial\lambda} = \frac{m\omega}{\lambda} - \log\alpha\left\{\sum_{i=1}^m\sum_{n=1}^{\omega-1}\frac{\left(\lambda x_i\right)^{n-1}x_i e^{-\lambda x_i}}{(n-1)!}\left(1-\frac{\lambda x_i}{n}\right) - \sum_{i=1}^m x_i e^{-\lambda x_i}\right\} - \sum_{i=1}^m x_i = 0. \tag{23c}$$

where $\psi(\omega) = \frac{\Gamma'(\omega)}{\Gamma(\omega)}$.

By solving the system of (23a)-(23c), we can then determine the maximum likelihood estimates (MLEs) for each parameter.

As shown in the aforementioned equations, the MLEs of the parameters do not admit closed-form solutions. This is a common feature among generalized or transformed distributions, especially those involving incomplete gamma functions or quantile-based expressions. In practice, we use numerical optimization methods, such as

the Newton-Raphson or BFGS algorithm, to compute the MLEs. Despite the lack of closed-form solutions, the estimators remain consistent and asymptotically normal, as confirmed through simulation studies. Hence, a numerical procedure is required to obtain MLEs. Based on the log-likelihood function and its partial derivatives with respect to the parameters, we apply an iterative root-finding algorithm to obtain the MLEs. The following algorithm outlines the computational steps used for parameter estimation.

Algorithm 1: Numerical Estimation of MLEs for APEr Distribution

Input: Observed data $x_1, x_2, \ldots, x_n$
Output: MLE estimates of parameters $\hat{\alpha}, \hat{\lambda}, \hat{k}$
Step 1. Define the log-likelihood function $\ell(\alpha, \lambda, k)$ based on Equation (21).
Step 2. Compute the partial derivatives $\frac{\partial \ell}{\partial \alpha}$, $\frac{\partial \ell}{\partial \lambda}$, and $\frac{\partial \ell}{\partial k}$ as in Equations (23a–c).
Step 3. Choose reasonable initial guesses $(\alpha^{(0)}, \lambda^{(0)}, k^{(0)})$.
Step 4. Use a numerical root-finding method (e.g., Newton-Raphson or BFGS) to solve the system of equations:

$$\frac{\partial \ell}{\partial \alpha} = 0, \quad \frac{\partial \ell}{\partial \lambda} = 0, \quad \frac{\partial \ell}{\partial k} = 0$$

Step 5. Update the parameter vector iteratively until convergence is achieved (e.g., when the change in parameters is below a predefined threshold ϵ).
Step 6. Return the final estimates $(\hat{\alpha}, \hat{\lambda}, \hat{k})$ as the MLEs.

We apply this algorithm in the numerical study to estimate the parameters of APEr.

4. Numerical Experiments

In this section, we study the performance of the APEr from the viewpoints of the consistency and unbiasedness of ML estimates of the parameters using simulated data, and its goodness of fit in comparison to alpha power exponential (APE), exponential and Erlang distributions using two real datasets.

4.1. Simulation Study

We divided simulation sutdy into two main parts. First we consider a simulated experiment to check if the ML estimators of α, λ and ω are consistent or not. Then, we organized a comparison study based on different values of α, λ, ω, and sample size n to obtain a general view about the distribution and the effects of changing n.

4.1.1. Study of Consistency To evaluate the consistency of the MLEs for the APEr distribution, we consider the definition of consistency: as n increases, the estimator tends to the true parameter value, and as $n \to \infty$, this difference converges to zero. In other words, for a consistent estimator, we expect the mean squared error (MSE) to decrease and approach zero as n increases.
To verify this property for the APEr distribution, we simulate datasets using the parameters $\alpha = 2$, $\lambda = 1$, $\omega = 5$, and $n = 10{,}000$. We generate 25 replicates of these datasets and compute the sequence of MSE values for increasing n. If the MSE values exhibit a decreasing trend and approach zero with increasing sample size, it confirms the consistency of the MLEs for the parameters of the APEr distribution.

Figure 2 shows the MSE curves of ML estimators for the parameters of APEr (α, λ and ω) as the sample size n increases. As observed, the MSE values for the MLEs of all three parameters of the APEr distribution decrease toward zero as n increases, with this convergence occurring around $n = 5000$. Additionally, the MSE values for the estimators of λ and ω are significantly smaller compared to α. For α, a larger sample size is essential to achieve convergence of the MSE toward zero.

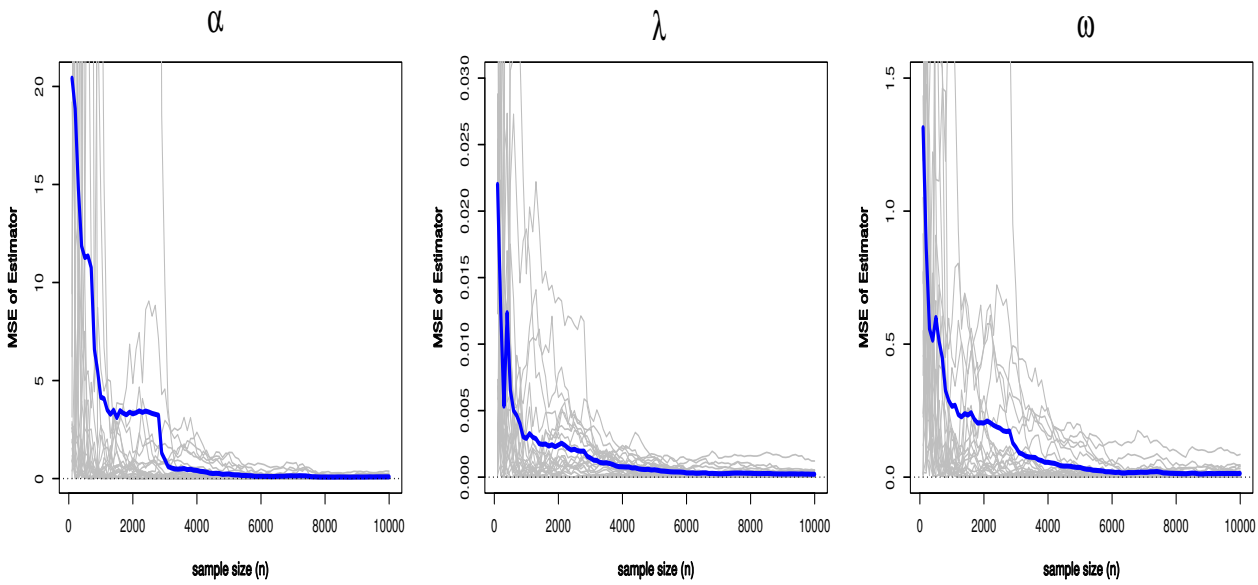


Figure 2. MSE curves of the ML estimators for the parameters of APEr through increasing the sample size n.

In other words, the MSE sequence for λ and ω consistently tends to zero over a narrow range as the sample size grows, even maintaining low values for smaller n. However, for α, the convergence to zero occurs more gradually, requiring a larger sample size for noticeable improvement. Overall, we can conclude that ML estimators for the parameters of APEr are consistent based on the resutls of our simulation study.

4.1.2. Comparison Study In the comparison study, we examine the precision of maximum likelihood estimators for APEr parameters. The mean of square error (MSE), the mean of the obtained MLE and the bias of ML estimates for each parameter as comparison tools. We follow the following steps:

1. Consider 4 sample size $n = 100, 200, 400, 1000$ for $\alpha = 2, 4, \lambda = 0.5, 0.75, 1, 1.5$, and $\omega = 5, 10$
2. Repeat the above step $N = 1000$ times
3. Calculate the MSE, mean of MLE and bias of each parameter over the 1000 iterations.

The different values of n and the parameters make 64 different possible situation to generate data and will have 1000 iterations from each situation. The results of generating random sample and calculating the MSE, the mean of the obtained MLE and the bias values are prepared in Table 1.

As can obtain from the observations in the table, increasing n across all 64 possible situations leads to a reduction in both MSE and bias, while the MLE values for the parameters approach their true values. For ω and λ, similar to the consistency analysis, it is observed that even at smaller values of n, the bias and MSE are significantly lower compared to α. Regarding α, although its performance improves as n increases, it still exhibits higher bias and MSE relative to the other two parameters. This observation aligns with the results of the consistency simulation study, where the MSE of the estimator for α converges to zero more slowly compared to the estimators for ω and λ.
It is also worth remarking that the simulation results indirectly demonstrate the robustness of the proposed MLE procedure. As sample size increases, the bias consistently decreases and the MSE stabilizes across different parameter settings, indicating that the estimation method remains reliable even under varying model specifications.

Table 1. Mean of MSE, MLE and bias of each parameter of APEr over the 1000 iterations.

α	ω	λ	n	$Mean(\hat{\alpha})$	$Mean(\hat{\lambda})$	$Mean(\hat{\omega})$	$MSE(\hat{\alpha})$	$MSE(\hat{\lambda})$	$MSE(\hat{\omega})$	$Bias(\hat{\alpha})$	$Bias(\hat{\lambda})$	$Bias(\hat{\omega})$
2	5	0.5	100	1.997	0.513	5.151	0.226	0.005	0.602	0.464	0.057	0.603
			200	1.998	0.506	5.071	0.215	0.002	0.293	0.448	0.039	0.429
			400	2.021	0.503	5.032	0.208	0.001	0.150	0.436	0.028	0.307
			1000	2.003	0.501	5.024	0.185	0.001	0.068	0.403	0.018	0.203
		0.75	100	1.997	0.770	5.151	0.226	0.012	0.602	0.464	0.086	0.603
			200	1.998	0.758	5.071	0.215	0.006	0.293	0.448	0.058	0.429
			400	2.021	0.754	5.032	0.208	0.003	0.150	0.436	0.042	0.307
			1000	2.003	0.751	5.024	0.185	0.001	0.068	0.404	0.027	0.203
		1	100	1.997	1.026	5.151	0.226	0.022	0.602	0.464	0.115	0.603
			200	1.998	1.011	5.071	0.215	0.010	0.293	0.448	0.078	0.429
			400	2.021	1.005	5.032	0.208	0.005	0.150	0.436	0.056	0.307
			1000	2.003	1.001	5.024	0.185	0.002	0.068	0.404	0.035	0.203
		1.5	100	1.997	1.539	5.151	0.226	0.049	0.602	0.464	0.172	0.603
			200	1.997	1.517	5.071	0.215	0.022	0.294	0.448	0.117	0.429
			400	2.021	1.508	5.032	0.208	0.011	0.150	0.436	0.084	0.307
			1000	2.003	1.502	5.024	0.185	0.005	0.068	0.404	0.053	0.203
	10	0.5	100	1.997	0.513	10.296	0.226	0.005	2.386	0.464	0.057	1.199
			200	1.998	0.506	10.137	0.216	0.002	1.142	0.448	0.039	0.846
			400	2.020	0.503	10.061	0.209	0.001	0.580	0.438	0.028	0.602
			1000	2.003	0.501	10.042	0.187	0.001	0.253	0.406	0.018	0.391
		0.75	100	1.997	0.770	10.296	0.226	0.012	2.386	0.464	0.086	1.199
			200	1.998	0.758	10.137	0.216	0.006	1.142	0.448	0.059	0.846
			400	2.020	0.754	10.061	0.209	0.003	0.580	0.438	0.042	0.602
			1000	2.003	0.751	10.042	0.187	0.001	0.253	0.406	0.026	0.391
		1	100	1.997	1.026	10.296	0.226	0.022	2.386	0.464	0.115	1.199
			200	1.998	1.011	10.137	0.216	0.010	1.142	0.448	0.078	0.846
			400	2.020	1.005	10.061	0.209	0.005	0.580	0.438	0.056	0.602
			1000	2.003	1.002	10.042	0.187	0.002	0.253	0.406	0.035	0.391
		1.5	100	1.997	1.539	10.296	0.226	0.049	2.386	0.464	0.172	1.199
			200	1.998	1.517	10.137	0.216	0.022	1.142	0.448	0.117	0.846
			400	2.020	1.508	10.061	0.209	0.011	0.580	0.438	0.084	0.602
			1000	0.013	0.007	0.049	0.001	0.000	0.001	0.002	0.000	0.002
4	5	0.5	100	3.711	0.516	5.252	1.437	0.006	0.793	1.144	0.059	0.689
			200	3.784	0.508	5.146	1.354	0.003	0.396	1.099	0.040	0.494
			400	3.831	0.504	5.098	1.257	0.001	0.217	1.053	0.029	0.370
			1000	3.848	0.502	5.078	1.113	0.001	0.118	0.969	0.018	0.274
		0.75	100	3.711	0.774	5.252	1.437	0.013	0.793	1.144	0.088	0.689
			200	3.784	0.762	5.146	1.354	0.006	0.396	1.099	0.060	0.494
			400	3.831	0.757	5.098	1.257	0.003	0.217	1.053	0.043	0.370
			1000	3.848	0.754	5.078	1.113	0.001	0.118	0.969	0.028	0.274
		1	100	3.711	1.032	5.252	1.437	0.023	0.793	1.144	0.118	0.689
			200	3.784	1.016	5.146	1.354	0.011	0.396	1.099	0.081	0.494
			400	3.831	1.009	5.098	1.257	0.005	0.217	1.053	0.058	0.370
			1000	3.848	1.005	5.078	1.113	0.002	0.118	0.969	0.037	0.274
		1.5	100	3.711	1.548	5.252	1.437	0.052	0.793	1.144	0.177	0.689
			200	3.784	1.524	5.146	1.354	0.024	0.396	1.099	0.121	0.494
			400	3.831	1.513	5.098	1.257	0.012	0.217	1.053	0.086	0.370
			1000	3.848	1.507	5.078	1.113	0.005	0.118	0.969	0.055	0.274
	10	0.5	100	3.711	0.517	10.473	1.439	0.006	3.023	1.145	0.059	1.343
			200	3.782	0.508	10.270	1.359	0.003	1.483	1.102	0.041	0.954
			400	3.829	0.505	10.179	1.268	0.001	0.797	1.059	0.029	0.707
			1000	3.844	0.503	10.138	1.125	0.001	0.416	0.977	0.019	0.512
		0.75	100	3.711	0.775	10.473	1.439	0.013	3.023	1.145	0.089	1.343
			200	3.782	0.763	10.270	1.359	0.006	1.483	1.102	0.061	0.954
			400	3.829	0.757	10.179	1.268	0.003	0.797	1.059	0.044	0.707
			1000	3.844	0.754	10.138	1.125	0.001	0.416	0.977	0.028	0.512
		1	100	3.711	1.034	10.473	1.439	0.024	3.023	1.145	0.119	1.343
			200	3.782	1.017	10.270	1.359	0.011	1.483	1.102	0.082	0.954
			400	3.829	1.010	10.179	1.268	0.006	0.797	1.059	0.058	0.707
			1000	3.844	1.006	10.138	1.125	0.002	0.416	0.977	0.038	0.512
		1.5	100	3.711	1.551	10.473	1.439	0.053	3.023	1.145	0.178	1.343
			200	3.782	1.525	10.270	1.359	0.025	1.483	1.102	0.123	0.954
			400	3.829	1.515	10.179	1.268	0.012	0.797	1.059	0.088	0.707
			1000	3.875	1.506	10.107	1.088	0.005	0.396	0.956	0.058	0.509

4.2. Real Data Analysis

In this subsection, we apply the APEr distribution on two real datasets to demonstrate its importance and flexibility. First, we analyze a dataset containing remission times for 128 cancer patients. These remission times are part of a dataset from a bladder cancer study. Second dataset is Covid-19 mortality rates data in the United Kingdom used by [13].

4.2.1. Remission Times Dataset Detailed information about this dataset can be found on page 230 and Table 9.3, of [15]. In continue, we refer to this dataset as the remission times dataset for summarization. Table 2 shows descriptive statistics of the remission times dataset.

Table 2. Descriptive values of remission times dataset

Min	1st.Qu	Median	Mean	3rd.Qu	Max	Skewness	Kurtosis
0.080	3.348	6.395	9.367	11.838	79.050	3.287	18.483

Table 3 shows the MLEs of parameters of the models (distributions) fitted to remission times dataset. In continue, we first test the goodness of fit for the fitted models on data and check which ones are rejected based the statistical tests. Indeed, we examine the goodness of fit for the fitted models over the remission times dataset. we

Table 3. MLE values of the model parameters for fitted distributions on remission times dataset data.

Distribution	α	λ	ω
APEr	0.0561	0.1301	2
APE	0.0001	0.0130	-
Er	-	0.2135	2
E	-	0.1068	-

consider and compute the Kolmogorov–Smirnov ($K--S$), and Anderson–Darling (AD^*) statistics for testing the goodness of fit. The smaller values of $K-S$ and AD^* indicate the better fitting of the model. [7] provided detailed information about Anderson–Darling test.

Table 4 presents the values of AD^*, and $K--S$ statistics. Moreover, it provides $p-values$ of Anderson–Darling and $K--S$ tests. According to the Table 4, the Erlang distribution is not accepted as a good fit over the data. The results accept the fitting of APEr, APE and E. Also, based on the $K-S$ statistic $APEr$ has the best performance and according to AD^* statistic Exponential distribution (E) is the best fit.

Table 4. Goodness-of-fit tests for remission times dataset.

Distribution	AD^*	$p-value$	$K-S$	$p-value$
APEr	1.912	0.103	0.079	0.407
APE	1.594	0.156	0.103	0.129
Er	5.359	0.002	0.141	0.012
E	1.174	0.278	0.085	0.318

Now, we analyze the performance of accepted models by the goodness-of-fit tests to reach the best conclusion about the best fitted model on remission times dataset. For more comparison and reaching better decision about the best model fitted over the remission times dataset, we apply likelihood ratio (LR) test. We should calculate the maximum amounts of log-likelihoods of null and alternative hypothesis to obtain the LR statistics. In LR test, if Θ be the parameter space of problem, we divide the parameter space as $\Theta = (\Theta_0, \Theta_1)$. Θ_0 is the parameter space

of model under null hypothesis (H_0) and Θ_1 is the parameter space of model under alternative hypothesis (H_1). For example, if we want to asses the $APEr$ model against E, Θ_0 is the parameter space under the assumption of $APEr$ distribution and Θ_1 is the parameter space when we consider E for data. $\hat{\Theta}_0$ and $\hat{\Theta}_1$ be MLE under H_0 and H_1, the LR statistic for this problem is calculated as:

$$LR = -2\log\left[\frac{L\left(\hat{\Theta}_0\right)}{L\left(\hat{\Theta}_1\right)}\right] = 5.7436.$$

Table 5 shows the information about LR test for APEr against APE and E. Note that Erlang distribution is not considered here (because of rejection based on goodness-of-fit tests). For the aforementioned example about testing $APEr$ versus E, $p - value = 0.024$ and we have another evidence to obtain $APEr$ as the best model fitted over the remission times dataset.

Table 5. Likelihood ratio test results for remission times dataset.

Model	LR statistic	p-value
APEr vs APE	5.991	0.045
APEr vs E	5.743	0.024

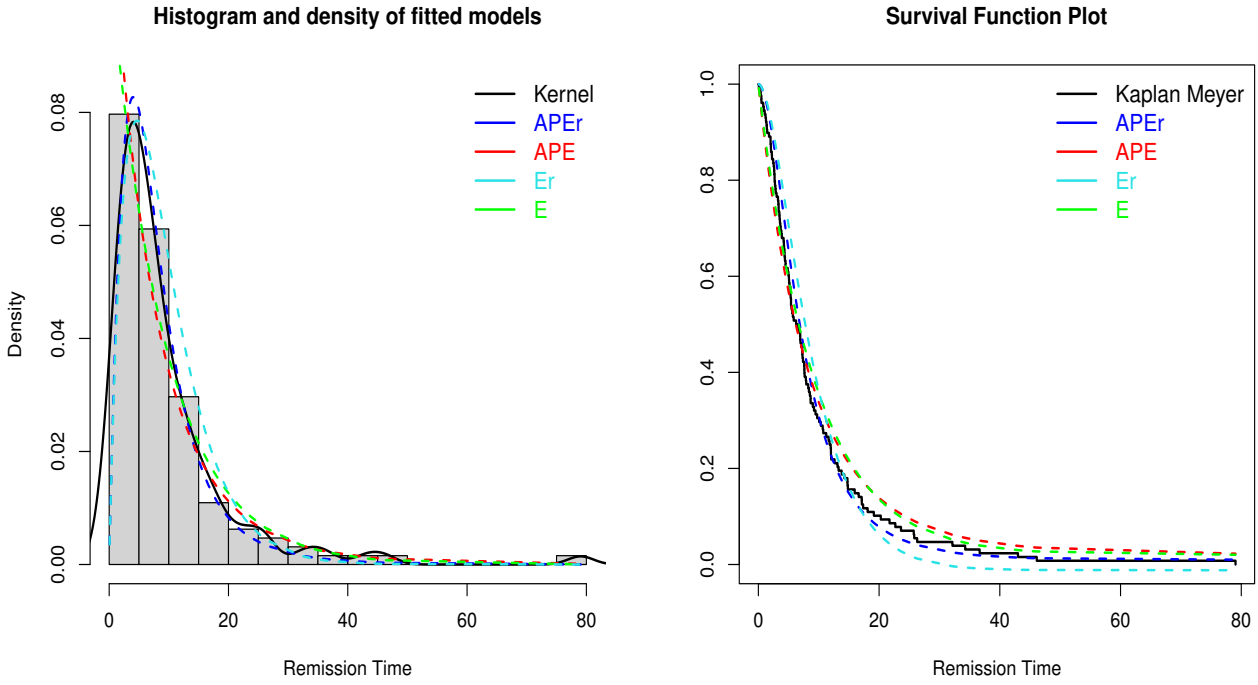


Figure 3. The right plot depicts the Kaplan-Meier survival function estimate for the remission times dataset alongside the survival functions of the fitted distributions (APEr, APE, E, and Er). The left plot illustrates the histogram and Kernel density estimate of the remission times dataset along with the density functions of the fitted distributions.

Figure 3 compares the performance of the fitted distributions using survival and density function. The right plot shows the Kaplan-Meier estimate of the survival function for the actual data used, along with the survival functions

of the fitted distributionsAPEr (blue), APE (red), E (green), Er (light blue). As observed from the plot, APEr provides a closer fit to the Kaplan-Meier estimate of the survival function compared to other fitted distributions.

The left plot presents the histogram and the probability density estimate of the data using the Kernel method (a non-parametric method) alongside the density functions of the fitted distributions with their estimated parameters. The peak height of the Kernel estimate is close to that of APEr. For Er, it can also be noted that it bears more resemblance to APEr compared to other distributions. In contrast, the modes estimated by E and APE for the data are significantly distant from the observed mode, indicating the poor fit of these distributions to the data.

4.2.2. UK Covid-19 Dataset Table 6 shows descriptive statistics of Covid-19 mortality rates data in the United Kingdom (UK Covid-19 Dataset).

Table 6. Descriptive values of UK Covid-19 dataset

Min	1st.Qu	Median	Mean	3rd.Qu	Max	Skewness	Kurtosis
0.242	0.638	1.117	1.322	1.834	3.385	0.911	2.939

Table 7 shows the MLEs of parameters of the models (distributions) fitted to UK Covid-19 dataset. Now, we test the goodness of fit for the fitted models on data and check which ones are rejected based the statistical tests. Similar

Table 7. MLE values of the model parameters for fitted distributions on UK Covid-19 dataset data.

Distribution	α	λ	ω
APEr	0.4239	1.8545	3
APE	69.9921	1.5735	-
Er	-	3	1.9848
E	-	0.7565	-

to first dataset, we consider and compute the Kolmogorov–Smirnov ($K--S$), and Anderson–Darling (AD^*) statistics for testing the goodness of fit. Table 8 presents the values of AD^*, and $K--S$ statistics. Moreover, it provides $p-values$ of Anderson–Darling and $K--S$ tests. According to the Table 8, the Exponential distribution is not accepted as a good fit over the data. The results accept the fitting of APEr, APE and Er. Based on the AD^* and $K-S$ statistics, $APEr$ has the best performance. It is worth noting that although the estimated shape parameter $\omega = 2$ in the APEr model corresponds to a two-stage Erlang structure, the Erlang distribution was rejected by Goodness-of-fit tests. Thus, despite the numerical resemblance, the additional shape flexibility provided by the α parameter contributes to a better overall fit. This highlights the importance of balancing parsimony and modeling power when analyzing real-world lifetime data. Now, for reaching better decision about the best model fitted over

Table 8. Goodness-of-fit tests for UK Covid-19 dataset.

Distribution	AD^*	$p-value$	$K-S$	$p-value$
APEr	0.425	0.824	0.065	0.897
APE	0.708	0.552	0.082	0.682
Er	0.502	0.745	0.075	0.789
E	5.021	0.003	0.209	0.003

the UK Covid-19 dataset, we apply likelihood ratio (LR) test. Table 9 shows the information about LR test for APEr against APE and Er. Note that Exponential distribution is not considered here (because of rejection based on goodness-of-fit tests). For the aforementioned example about testing $APEr$ versus Er, $p-value = 0.034$ and we have another evidence to obtain $APEr$ as the best model fitted over the UK Covid-19 dataset.

Table 9. Likelihood ratio test results for the UK Covid-19 dataset.

Model	LR statistic	p-value
APEr vs APE	5.473	0.041
APEr vs Er	4.977	0.034

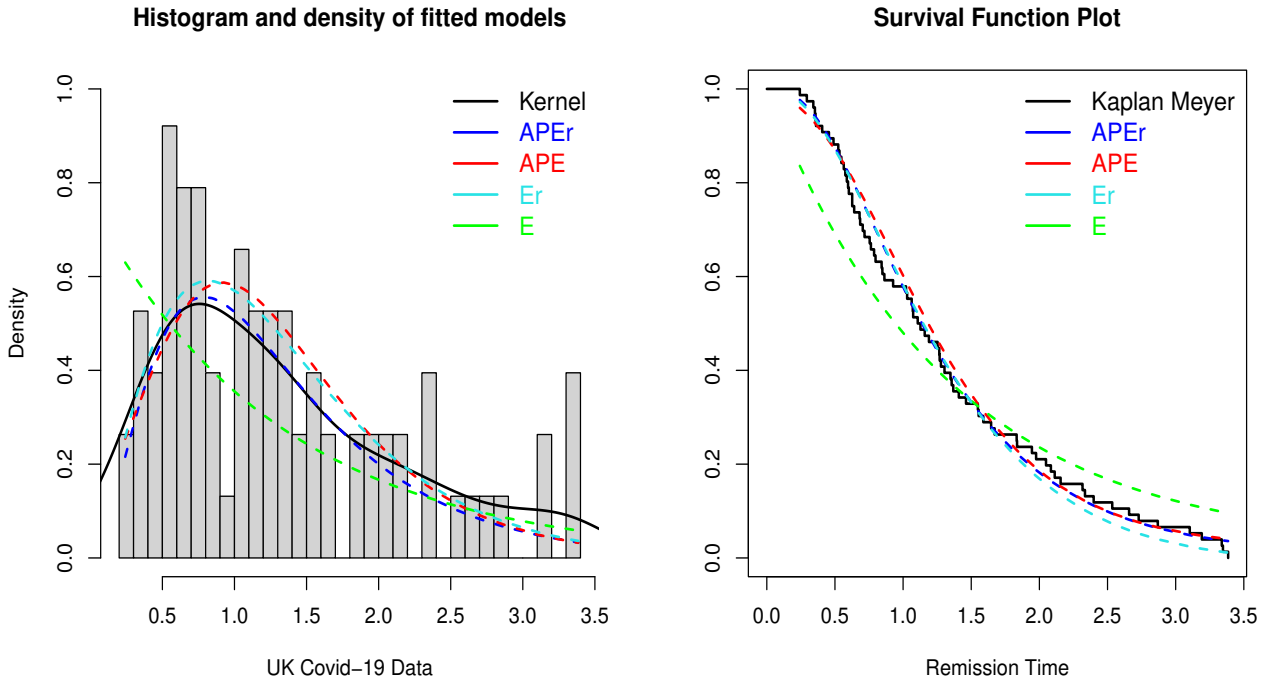


Figure 4. The right plot depicts the Kaplan-Meier survival function estimate for the UK Covid-19 dataset alongside the survival functions of the fitted distributions (APEr, APE, E, and Er). The left plot illustrates the histogram and Kernel density estimate of the UK Covid-19 dataset along with the density functions of the fitted distributions.

Figure 4 compares the performance of the fitted distributions using survival and density function. The right plot shows the Kaplan-Meier estimate of the survival function for the actual data used, along with the survival functions of the fitted distributionsAPEr (blue), APE (red), E (green), Er (light blue). As observed from the plot, APEr provides a closer fit to the Kaplan-Meier estimate of the survival function compared to other fitted distributions.

The left plot presents the histogram and the probability density estimate of the data using the Kernel method (a non-parametric method) alongside the density functions of the fitted distributions with their estimated parameters. The peak height of the Kernel estimate is close to that of APEr. For Er, it can also be noted that it bears more resemblance to APEr compared to other distributions. In contrast, the modes estimated by E and APE for the data are significantly distant from the observed mode, indicating the poor fit of these distributions to the data.

As a final remark before, it is worth emphasizing that the APEr distribution is a theoretical generalization of the Alpha Power Exponential (APE) distribution, one of the most foundational models in the alpha power family. Given that APE has already been extensively compared with standard lifetime distributions such as Weibull, Exponential, and Pareto in previous studies, our inclusion of APE in the comparative analysis naturally provides an indirect benchmark against these distributions. Furthermore, while additional comparisons with other alpha power-based models are conceivable, their structural motivations and hazard characteristics often differ significantly from the APEr framework. For these reasons, we focused our comparisons on models that share a more direct lineage with APEr.

Conclusion

Considering lifetime distributions and introducing new models for analyzing such data has been the focus of extensive research. The alpha power family of distributions, which generates a new class of distributions by adding a parameter to the baseline distribution, has proven to be a valuable tool for modeling lifetime data. On the other hand, the exponential distribution is one of the most well-known models for lifetime data. In problems involving the sum of multiple exponential distributions, the Erlang distribution, as a special case of the gamma distribution, is commonly used. In this paper, we introduced the alpha power Erlang (APEr) distribution as a special case of the alpha power gamma distribution. While the exponential distribution has been extensively studied in previous research, the direct exploration of the alpha power gamma distribution has remained largely untouched.

Along with introducing a new distribution designed for lifetime data, the APEr distribution has important mathematical features. It offers simple closed-form formulas for its quantiles and moments. Using theorems provided in this paper, we also found closed-form solutions for order statistics and entropy measures. Even though the maximum likelihood estimators (MLEs) do not have closed-form solutions, simulation results showed that the MLEs are consistent, meaning they become more accurate as the sample size increases, allowing for reliable analysis with larger datasets. The performance of the APEr distribution was compared with other related distributions in terms of goodness-of-fit on two real-world datasets. The results demonstrated that the APEr distribution provides a better fit for the studied data.

For future studies, the censored version of APEr can be considered by applying the setting of truncated distributions on the current APEr. Also, the direct exploration of the alpha power gamma distribution, without the constraint of the scale parameter being an integer (as in the Erlang case), could offer broader applicability. Additionally, applying probability rules to develop an alpha power distribution for mixtures of Erlang or exponential distributions with heavy-tailed distributions may provide a more robust approach for handling outliers. Moreover, such extensions may contribute to improving the generalizability of the proposed model across various lifetime data contexts. For instance, mixture-based models allow for better adaptation to heterogeneous populations and can capture multi-modal or skewed failure behaviors more effectively. While these extensions pose non-trivial theoretical challenges, such as deriving closed-form expressions or ensuring identifiability, they offer valuable directions for broadening the practical impact and scope of the alpha power family of distributions.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding

No funding was received.

REFERENCES

1. AL-Darraji, Z.S., Badr, D.: Comparison of the methods of estimating the parameters and reliability function for the erlang distribution with two parameters with a practical application. In: AIP Conference Proceedings, vol. 2834. AIP Publishing (2023)
2. Alshkaki, R.: A generalized modification of the kumaraswamy distribution for modeling and analyzing real-life data. Statistics, Optimization & Information Computing **8**(2), 521–548 (2020)
3. Alzaatreh, A., Lee, C., Famoye, F.: A new method for generating families of continuous distributions. Metron **71**(1), 63–79 (2013)
4. Bakhareva, N., Tarasov, V.: Application of two forms of the erlang distribution law in queueing theory. In: 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), pp. 639–642. IEEE (2021)
5. Basheer, A.M., Almetwally, E.M., Okasha, H.M.: Marshall-olkin alpha power inverse weibull distribution: non bayesian and bayesian estimations. Journal of Statistics Applications & Probability **10**(2), 327–345 (2021)
6. Boland, P.J., Hu, T., Shaked, M., Shanthikumar, J.G.: Stochastic ordering of order statistics ii. Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications pp. 607–623 (2002)

7. Chen, G., Balakrishnan, N.: A general purpose approximate goodness-of-fit test. Journal of Quality Technology **27**(2), 154–161 (1995)
8. Dey, S., Sharma, V.K., Mesfioui, M.: A new extension of weibull distribution with application to lifetime data. Annals of Data Science **4**, 31–61 (2017)
9. Dhungana, G.P., Chaudhary, A.K., Tharu, R.P., Kumar, V.: Generalized alpha power inverted weibull distribution: Application of air pollution in kathmandu, nepal. Annals of Data Science pp. 1–25 (2024)
10. Erlang, A.K.: The theory of probabilities and telephone conversations. Nyt. Tidsskr. Mat. Ser. B **20**, 33–39 (1909)
11. Eugene, N., Lee, C., Famoye, F.: Beta-normal distribution and its applications. Communications in Statistics-Theory and methods **31**(4), 497–512 (2002)
12. Jones, M.: Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. Statistical methodology **6**(1), 70–81 (2009)
13. Kilai, M., Waititu, G.A., Kibira, W.A., Alshanbari, H.M., El-Morshedy, M.: A new generalization of gull alpha power family of distributions with application to modeling covid-19 mortality rates. Results in Physics **36**, 105339 (2022)
14. Lee, C., Famoye, F., Alzaatreh, A.Y.: Methods for generating families of univariate continuous distributions in the recent decades. Wiley Interdisciplinary Reviews: Computational Statistics **5**(3), 219–238 (2013)
15. Lee, E.T., Wang, J.: Statistical methods for survival data analysis, vol. 476. John Wiley & Sons (2003)
16. Lekhane, O., Oluyede, B., Gabaitiri, L., Mabikwa, O.V.: The exponentiated-gompertz-marshall-olkin-g family of distributions: Properties and applications. Statistics, Optimization & Information Computing **13**(5), 1752–1788 (2025)
17. Machihara, F.: Mobile telecommunication systems and generalized erlang loss formula. IEICE transactions on communications **88**(1), 183–189 (2005)
18. Mahdavi, A., Kundu, D.: A new method for generating distributions with an application to exponential distribution. Communications in Statistics-Theory and Methods **46**(13), 6543–6557 (2017)
19. Marshall, A.W., Olkin, I.: A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. Biometrika **92**(2), 505–505 (2005)
20. Mead, M.E., Cordeiro, G.M., Afify, A.Z., Al Mofleh, H.: The alpha power transformation family: properties and applications. Pakistan Journal of Statistics and Operation Research pp. 525–545 (2019)
21. Mudholkar, G.S., Srivastava, D.K.: Exponentiated weibull family for analyzing bathtub failure-rate data. IEEE transactions on reliability **42**(2), 299–302 (1993)
22. Ogunde, A., Fayose, S., Ajayi, B., Omosigho, D.: Properties, inference and applications of alpha power extended inverted weibull distribution. International Journal of Statistics and Probability **9**(6), 90–107 (2020)
23. Rasjid, J.M., Nurrohmah, S., Fithriani, I.: Alpha power inverse weibull distribution: A new lifetime distribution with application to gastric cancer data. In: ITM Web of Conferences, vol. 61, p. 01009. EDP Sciences (2024)
24. Salah, M.M.: On progressive type-ii censored samples from alpha power exponential distribution. Journal of Mathematics **2020**(1), 2584184 (2020)
25. Semary, H., Hussain, Z., Hamdi, W.A., Aldahlan, M.A., Elbatal, I., Nagarjuna, V.B.: Alpha–beta-power family of distributions with applications to exponential distribution. Alexandria Engineering Journal **100**, 15–31 (2024)
26. Shah, S., Hazarika, P.J., Chakraborty, S., Alizadeh, M.: The balakrishnan-alpha-beta-skew-laplace distribution: Properties and applications. Statistics, Optimization & Information Computing **11**(3), 755–772 (2023)
27. ZeinEldin, R.A., Ahsan ul Haq, M., Hashmi, S., Elsehety, M.: Alpha power transformed inverse lomax distribution with different methods of estimation and applications. Complexity **2020**(1), 1860813 (2020)