RESEARCH ARTICLE

Survival analysis of cancer patients using a new extended Weibull distribution

Hadeel S. Klakattawi ⓘ *

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

* hklakattawi@kau.edu.sa

Abstract

One of the most important applications of statistical analysis is in health research and applications. Cancer studies are mostly required special statistical considerations in order to find the appropriate model for fitting the survival data. Existing classical distributions rarely fit such data well and an increasing interest has been shown recently in developing more flexible distributions by introducing some additional parameters to the basic model. In this paper, a new five-parameters distribution referred as alpha power Kumaraswamy Weibull distribution is introduced and studied. Particularly, this distribution extends the Weibull distribution based on a novel technique that combines two well known generalisation methods, namely, alpha power and T-X transformations. Different characteristics of the proposed distribution, including moments, quantiles, Rényi entropy and order statistics are obtained. The method of maximum likelihood is applied in order to estimate the model parameters based on complete and censored data. The performance of these estimators are examined via conducting some simulation studies. The potential importance and applicability of the proposed distribution is illustrated empirically by means of six datasets that describe the survival of some cancer patients. The results of the analysis indicated to the promising performance of the alpha power Kumaraswamy Weibull distribution in practice comparing to some other competing distributions.

1 Introduction

Many statistical distributions have been extensively utilized for analysing time to event data also referred to as survival or reliability data, in different areas of applicability, including the medical field. Medical scientists are mostly interested in studying the survival of patients with cancer in the applied research. These research are most often require special statistical attentions and adjustments in the context of finding and choosing the appropriate model that accurately determine and estimate the survival data and yielded in reliable results and valid inferences. It is possible to consider the Weibull distribution [1], to be one of the most popular distributions for modeling such data that explain the mortality and failure. However, the classical two-parameter Weibull distribution is less suitable for fitting when data show non-monotonic failure rates due to its limitation in modeling only monotonically increasing and decreasing hazard functions. Therefore, there is a crucial need in many cases to enhance the

traditional Weibull for modeling biomedical data. It follows that many attempts have been made to extend the baseline Weibull model by adding one or more additional parameters to achieve more flexibility in generating different shapes of data. To illustrate, [2] suggested the exponentiated Weibull distribution by applying the exponentiated method [3] in which a shape parameter is added to a baseline distribution. Beta-Weibull ditribution [4] is introduced based on the beta-generated method by [5]. Marshall–Olkin extended Weibull distribution [6] has been suggested to modify the Weibull distribution using the technique by [7]. This distribution has been applied to fit a dataset representing the remission times of bladder cancer patients. Furthermore, the Maxwell–Weibull distribution is introduced by [8] to model lifetime data. On the basis of the zero truncated Poisson model, [9] proposed a new compound distribution called the quasi Poisson Burr X exponentiated Weibull distribution, which accommodated many important failure rates. Moreover, in a recent study, [10] have derived a bimodal form of the Weibull distribution.

Researchers have shown a keen interest in developing new methods for expanding lifetime distributions. [11] developed a new method that add two extra shape parameters $a, b > 0$ to an arbitrary baseline distribution, called Kumaraswamy generalized (Kum-G) with a cumulative distribution function (cdf) defined as

$$F^{Kum-G}(x; a, b, \boldsymbol{\theta}) = 1 - [1 - (G(x; \boldsymbol{\theta}))^a]^b,$$

where X is a continuous random variable whose baseline (cdf) is $G(x;\boldsymbol{\theta})$ with a vector of parameter(s) $\boldsymbol{\theta}$. A number of studies have been applied this method to develop new distribution such as, the Kumaraswamy Gumbel by [12], the Kumaraswamy Birnbaum-Saunders by [13], the Kumaraswamy Burr XII distribution by [14], the Kumaraswamy generalized Rayleigh distribution by [15], the Kumaraswamy Laplace distribution by [16], the Kumaraswamy half-logistic distribution by [17], the Kumaraswamy exponentiated Weibull by [18], the Kumaraswamy Marshall-Olkin exponential distribution by [19] and the Kumaraswamy Pareto IV distribution by [20], among others.

[21] introduced the Kumaraswamy Weibull (KumW) distribution as a generalization of the Weibull distribution and demonstrated its flexibility to fit failure data. The proposed distribution can be obtained by assuming $G(x) = 1 - e^{-(\lambda x)^c}$ of the Weibull distribution with scale parameter $\lambda > 0$ and shape parameter $c > 0$. Thus, the cdf and probability density function (pdf) of the KumW is obtained respectively as

$$F^{KumW}(x; a, b, c, \lambda) = 1 - [1 - (1 - e^{-(\lambda x)^c})^a]^b, \tag{1}$$

and

$$f^{KumW}(x; a, b, c, \lambda) = abc\lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{a-1} [1 - (1 - e^{-(\lambda x)^c})^a]^{b-1}. \tag{2}$$

The KumW distribution has been considered by some authors, for example, [22, 23] discussed different types of statistical inference for constant stress accelerated life tests based on censored sampling data from the KumW distribution. [24] discussed some Bayesian analyses for the KumW distribution. [25] considered a regression model for bivariate random variables based on the bivariate KumW distribution. Although the KumW has been perfectly described many datasets, it has been modified by some authors. For instance, [26] in which the KumW is generalised by considering the new modified Kumaraswamy-G in [27]. Additionally, [28] who consider the exponentiated class in [3] to generalise the KumW. More recently, [29] generalised the KumW by considering the transmuted class in [30].

On the other hand, [31] suggested a new approach, called alpha power transformation (APT), for generating distributions with additional parameter α in order to add more

flexibility. Then, the APT for an arbitrary baseline cdf G and pdf g for a random variable X with a vector of parameter(s) $\boldsymbol{\theta}$ can be obtained as follows

$$F^{APT}(x; \alpha, \boldsymbol{\theta}) = \begin{cases} \dfrac{\alpha^{G(x;\boldsymbol{\theta})} - 1}{\alpha - 1} & \text{if } \alpha > 0, \alpha \neq 1 \\ G(x; \boldsymbol{\theta}) & \text{if } \alpha = 1, \end{cases} \tag{3}$$

with the corresponding pdf as

$$f^{APT}(x; \alpha, \boldsymbol{\theta}) = \begin{cases} \dfrac{log\alpha}{\alpha - 1} g(x;\boldsymbol{\theta}) \alpha^{G(x;\boldsymbol{\theta})} & \text{if } \alpha > 0, \alpha \neq 1 \\ g(x; \boldsymbol{\theta}) & \text{if } \alpha = 1. \end{cases} \tag{4}$$

[31] applied their suggested way to a one-parameter exponential distribution to develop alpha power exponential distribution with two-parameters. Several authors have been applied the method of APT to extend some exiting distributions in the literature. Examples include the alpha power Weibull distribution by [32, 33], the alpha-power inverse Weibull distribution by [34, 35], the alpha power inverted exponential by [36], the alpha power transformed extended exponential distribution by [37], the alpha power transformed power Lindley by [38], the alpha power transformed Lindley by, [39], the alpha power transformed inverse Lindley by [40], the alpha power transformed inverse Lomax by [41], alpha power Maxwell distribution by [42], the alpha power exponentiated inverse Rayleigh by [43] and the alpha power Weibull–exponential by [44], among others.

Motivated by the idea that developing some new distributions will eliminate some issues that inherent in the existing distributions, the main objective of this paper is to introduce a novel generalization for the Weibull distribution. This distribution is constructed by combining the works of [21, 31] introducing a new five-parameter distribution refereed to as the alpha power Kumaraswamy Weibull (APKumW) distribution. As compared to other probability distributions presented in the literature, the proposed model will increase the flexibility and adaptability for describing different shapes of hazard-rate functions, such as decreasing, increasing, bath-tub and upside down bath-tub shaped, which might extensively experienced in real life data. Particularly, as indicated by [45] for the effectiveness of employing the APT distributions for cancer research, this paper focuses in exploring the adaptability of the proposed distribution to describe the survival time by analyzing some cancer datasets. Additionally, another objective is to estimate the unknown model parameters using maximum likelihood method for both complete and censored cancer datasets.

The rest of the paper is organized as follows. In Section 2, the APKumW distribution is defined and its special cases are presented along with an useful expansion for its pdf. In Section 3, some of the properties of the proposed distribution are discussed. The maximum likelihood estimators (MLEs) of the distribution parameters are obtained in Section 4 based on uncensored and censored data. Consequently, some different simulation studies are carried out to assess the performance of the MLEs in Section 5. Finally, different applications of the APKumW distribution to complete and censored datasets are presented in Section 6. All computations throughout this paper were performed using the statistical programming language R.

2 Alpha power Kumaraswamy Weibull distribution

The APKumW distribution is suggested in this paper based on substituting by Eqs (1) and (2) respectively in Eqs (3) and (4). That is, the random variable X is said to have the APKumW

distribution with five parameters $\boldsymbol{\theta} = \{a, b, c, \lambda, \alpha\}$, if the cdf of X is

$$F(x; \boldsymbol{\theta}) = \begin{cases} \dfrac{\alpha^{1-[1-(1-e^{-(\lambda x)^c})^a]^b} - 1}{\alpha - 1} & \text{if } \alpha > 0, \alpha \neq 1 \\[3mm] 1 - [1 - (1 - e^{-(\lambda x)^c})^a]^b & \text{if } \alpha = 1, \end{cases} \tag{5}$$

and its corresponding pdf is

$$f(x; \boldsymbol{\theta}) = \begin{cases} \dfrac{\log(\alpha)}{\alpha - 1} \; abc\lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{a-1} \times \\[2mm] \qquad [1 - (1 - e^{-(\lambda x)^c})^a]^{b-1} \alpha^{1-[1-(1-e^{-(\lambda x)^c})^a]^b} & \text{if } \alpha > 0, \alpha \neq 1 \\[3mm] abc\lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{a-1} [1 - (1 - e^{-(\lambda x)^c})^a]^{b-1} & \text{if } \alpha = 1. \end{cases} \tag{6}$$

Additionally, the survival and hazard rate functions of the APKumW distribution are respectively given by

$$SF(x; \boldsymbol{\theta}) = \begin{cases} \dfrac{\alpha}{\alpha - 1}\left(1 - \alpha^{-[1-(1-e^{-(\lambda x)^c})^a]^b}\right) & \text{if } \alpha > 0, \alpha \neq 1 \\[3mm] [1 - (1 - e^{-(\lambda x)^c})^a]^b & \text{if } \alpha = 1, \end{cases} \tag{7}$$

and

$$h(x; \boldsymbol{\theta}) = \begin{cases} \log(\alpha) \, abc\lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{a-1} \times \\[2mm] \qquad [1 - (1 - e^{-(\lambda x)^c})^a]^{b-1} \dfrac{\alpha^{-[1-(1-e^{-(\lambda x)^c})^a]^b}}{1 - \alpha^{-[1-(1-e^{-(\lambda x)^c})^a]^b}} & \text{if } \alpha > 0, \alpha \neq 1 \\[3mm] abc\lambda^c x^{c-1} e^{-(\lambda x)^c} \dfrac{(1 - e^{-(\lambda x)^c})^{a-1}}{1 - (1 - e^{-(\lambda x)^c})^a} & \text{if } \alpha = 1. \end{cases} \tag{8}$$

Incorporating skewness to the base distribution is done by adding the parameter α. The APKumW model is therefore a suitable model to describe positively skewed patterns in biomedical and public health data. Fig 1 displays some of the shapes that the pdf and hazard functions of the APKumW distribution can take for different values of its parameters. These different behaviours indicate the flexibility and adaptability for the APKumW to fit a variety of data shapes.

2.1 Special cases

Table 1 shows important special models of the APKumW distribution.

2.2 Expansion of the probability density function

Using the following power series expansion

$$\alpha^{-z} = \sum_{k=0}^{\infty} \frac{(-\log(\alpha))^k z^k}{k!}, \tag{9}$$
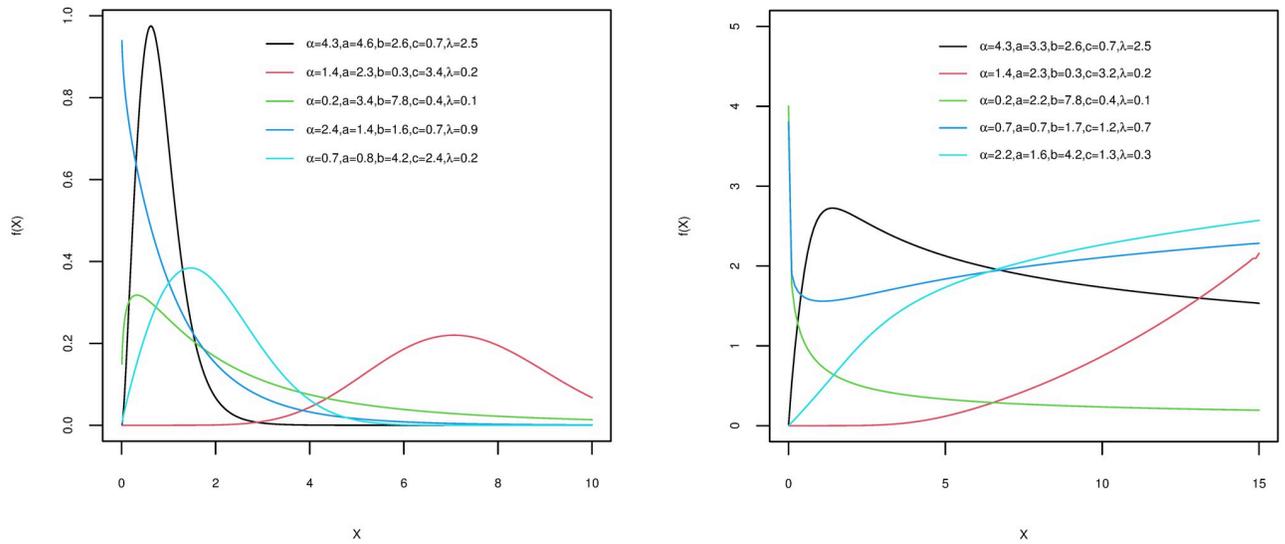
Fig 1. The APKumW pdf and hazard function for various values of its parameters.

the pdf in Eq (6) can be written as

$$f(x; \boldsymbol{\theta}) = \frac{\alpha}{\alpha - 1} \, abc\lambda^c x^{c-1} e^{-(\lambda x)^c} \left(1 - e^{-(\lambda x)^c}\right)^{a-1} \times$$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(log(\alpha)\right)^{k+1} \left[1 - \left(1 - e^{-(\lambda x)^c}\right)^a\right]^{bk+b-1}.$$

Then, the following binomial expansion

$$(1 - z)^{b-1} = \sum_{j=0}^{\infty} (-1)^j \binom{b-1}{j} z^j \qquad ; \qquad \text{for } |z| < 1 \text{ and } b > 0, \tag{10}$$

Table 1. Special models of the APKumW distribution.

α	a	b	c	λ	Resulting Distribution
1	–	–	–	–	KumW
–	–	–	1	–	APKumExp
–	–	–	2	–	APKum-Rayleigh
–	1	1	1	–	AP-Exponential
–	1	1	2	–	AP-Rayleigh
–	1	1	–	–	AP-Weibull
1	–	1	–	–	Expontiated Weibull
1	–	1	2	–	Expontiated Rayleigh
1	–	1	1	–	Expontiated Exponential
1	1	1	1	–	Exponential
1	1	1	1	–	Rayleigh
1	1	1	–	–	Weibull

is applied twice to obtain a useful expansion of the pdf of the APKumW as follows

$$f(x; \boldsymbol{\theta}) = \frac{\alpha}{\alpha - 1} abc\lambda^c \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{i+j+k}}{k!} \binom{b(k+1)-1}{j} \binom{a(j+1)-1}{i} (log(\alpha))^{k+1}$$

$$x^{c-1} e^{-(i+1)(\lambda x)^c}. \tag{11}$$

3 Properties of alpha power Kumaraswamy Weibull distribution

Some properties of the APKumW distribution are considered in the following as

3.1 Simulation, quantiles and median

To simulate a random variable from APKumW distribution, Eq (5) can be used to obtain

$$X = \frac{1}{\lambda} \left[- log \left\{ 1 - \left[1 - \left(1 - \frac{log(U(\alpha - 1) + 1)}{log(\alpha)} \right)^{\frac{1}{b}} \right]^{\frac{1}{a}} \right\}^{\frac{1}{c}}, \tag{12}$$

where U is a random variable follows a uniform $(0, 1)$ distribution. Also, the pth quantile function of the APKumW distribution for $0 < p < 1$, is given by

$$X_p = \frac{1}{\lambda} \left[- log \left\{ 1 - \left[1 - \left(1 - \frac{log(p(\alpha - 1) + 1)}{log(\alpha)} \right)^{\frac{1}{b}} \right]^{\frac{1}{a}} \right\}^{\frac{1}{c}}. \tag{13}$$

Consequently for $p = \frac{1}{2}$, the median for the APKumW can be obtained as

$$X_{0.5} = \frac{1}{\lambda} \left[- log \left\{ 1 - \left[1 - \left(1 - \frac{log(\alpha + 1) - log(2)}{log(\alpha)} \right)^{\frac{1}{b}} \right]^{\frac{1}{a}} \right\}^{\frac{1}{c}}. \tag{14}$$

3.2 Moments

The rth moment of a random variable X is given by

$$E(X^r) = \int_0^{\infty} x^r f(x; \boldsymbol{\theta}) dx.$$

Then, the rth moment of the APKumW is given from Eq (11) as

$$E(X^r) = \frac{\alpha}{\alpha - 1} abc\lambda^c \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{i+j+k}}{k!} \binom{b(k+1)-1}{j} \binom{a(j+1)-1}{i} (log(\alpha))^{k+1} \times$$

$$\int_0^{\infty} x^{r+c-1} e^{-(i+1)(\lambda x)^c} dx.$$

By letting $u = (i + 1)(\lambda x)^c$, then the rth moment can be obtained as

$$E(X^r) = \frac{\alpha}{\alpha - 1} \frac{ab}{\lambda^r} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{i+j+k}}{k!} \binom{b(k+1)-1}{j} \binom{a(j+1)-1}{i} (log(\alpha))^{k+1} \times$$

$$\left(\frac{1}{i+1} \right)^{\frac{r}{c}+1} \Gamma(\frac{r}{c} + 1), \tag{15}$$

where $\Gamma(.)$ is the gamma function. Subsequently, the mean and variance can be obtained by substituting $r = 1$ and $r = 2$ in Eq (15).

The moment generating function of a random variable X can be defined with the form

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x; \boldsymbol{\theta}) dx.$$

That is, using the following power series expansion for the exponential function

$$e^z = \sum_{l=0}^\infty \frac{(z)^l}{l!}, \tag{16}$$

the moment generating function of a random variable X whose pdf in Eq (6), can be obtained similarly as

$$M_x(t) = \frac{ab\alpha}{\alpha - 1} \sum_{l=0}^\infty \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty \frac{(-1)^{i+j+k}}{k!} \frac{1}{l!} \binom{b(k+1)-1}{j} \binom{a(j+1)-1}{i} (log(\alpha))^{k+1} \times$$
$$\left(\frac{t}{\lambda}\right)^l \left(\frac{1}{i+1}\right)^{\frac{l}{c}+1} \Gamma(\frac{l}{c}+1). \tag{17}$$

3.3 Rényi entropy

The Rényi entropy of a random variable X represents a measure of variation of the uncertainty and given by

$$RE_x(v) = \frac{1}{1-v} log\left(\int_{-\infty}^\infty [f(x; \boldsymbol{\theta})]^v dx\right) \qquad ; \qquad v > 0, \quad v \neq 0.$$

Then from Eq (6), we have

$$[f(x; \boldsymbol{\theta})]^v = \left(\frac{\alpha log(\alpha)}{\alpha - 1}\right)^v (abc\lambda^c)^v x^{v(c-1)} e^{-v(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{v(a-1)} \times$$
$$[1 - (1 - e^{-(\lambda x)^c})^a]^{v(b-1)} \alpha^{-v[1-(1-e^{-(\lambda x)^c})^a]^b}.$$

Applying Eqs (9) and (10) twice, we obtain

$$[f(x; \boldsymbol{\theta})]^v = \left(\frac{\alpha}{\alpha - 1} abc\lambda^c\right)^v x^{v(c-1)} \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty \frac{(-1)^{i+j+k}}{k!} \binom{b(k+v)-v}{j} \binom{a(j+v)-v}{i} \times$$
$$v^k (log(\alpha))^{v+k} e^{-(i+v)(\lambda x)^c}.$$

Then,

$$RE_x(v) = \frac{1}{1-v} log\left\{\left(\frac{\alpha}{\alpha - 1} abc\lambda^c\right)^v \sum_{k=0}^\infty \sum_{j=0}^\infty \sum_{i=0}^\infty \frac{(-1)^{i+j+k}}{k!} \binom{b(k+v)-v}{j} \binom{a(j+v)-v}{i} \times\right.$$
$$\left. v^k (log(\alpha))^{v+k} \int_0^\infty x^{v(c-1)} e^{-(i+v)(\lambda x)^c} dx\right\}.$$

By assuming $u = (i+v)(\lambda x)^c$, the Rényi entropy for the APKumW can be expressed as

$$
\begin{aligned}
RE_x(v) = \quad & \frac{v}{1-v}log\left(\frac{ab\alpha}{\alpha-1}\right) - log(\lambda c) + \\
& \frac{1}{1-v}log\left\{\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\frac{(-1)^{i+j+k}}{k!}\binom{b(k+v)-v}{j}\binom{a(j+v)-v}{i}\times \right. \\
& \left. v^k(log(\alpha))^{v+k}\left(\frac{1}{i+v}\right)^{\frac{v(c-1)}{c}+\frac{1}{c}}\Gamma\left(\frac{v(c-1)}{c}+\frac{1}{c}\right)\right\}.
\end{aligned}
\tag{18}
$$

3.4 Order statistics

Suppose that $F(x)$ and $f(x)$ are respectively the cdf and pdf of n independent and identically distributed random variables $X_1, X_2, \ldots X_n$ with $X_{1:n} < X_{2:n} < \ldots < X_{n:n}$ be their corresponding ordered statistics. Then, the pdf of the sth order statistic can be obtained as

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

Using the binomial theorem, we have

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

Substituting by Eqs (5) and (6) and using the binomial theorem, we get

$$
\begin{aligned}
f_{s:n}(x) = & \frac{n!}{(s-1)!(n-s)!}\sum_{w=0}^{n-s}\sum_{v=0}^{w+s-1}(-1)^{w+v+1}\binom{n-s}{w}\binom{w+s-1}{v}\frac{\alpha^{v+1}}{(1-\alpha)^{w+s}}log(\alpha)\times \\
& abc\lambda^c x^{c-1}e^{-(\lambda x)^c}(1-e^{-(\lambda x)^c})^{a-1}[1-(1-e^{-(\lambda x)^c})^a]^{b-1}\times \\
& \alpha^{-(v+1)[1-(1-e^{-(\lambda x)^c})^a]^b}.
\end{aligned}
$$

Then, using the series expansion in Eq (9) and applying the binomial theorem in Eq (10) twice, we obtain

$$
\begin{aligned}
f_{s:n}(x) = & \frac{n!}{(s-1)!(n-s)!}\sum_{w=0}^{n-s}\sum_{v=0}^{w+s-1}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\binom{n-s}{w}\binom{w+s-1}{v}\binom{b(k+1)-1}{j}\binom{a(j+1)-1}{i}\times \\
& \frac{(-1)^{w+v+k+j+i+1}}{k!}\frac{\alpha^{v+1}(v+1)^k}{(1-\alpha)^{w+s}}(log(\alpha))^{k+1}abc\lambda^c x^{c-1}e^{-(i+1)(\lambda x)^c}.
\end{aligned}
\tag{19}
$$

4 Parameter estimation for alpha power Kumaraswamy Weibull distribution

The maximum likelihood method is applied to obtain the estimation for the parameters of APKumW distribution. That is, if we have a random sample $x_1, x_2, \ldots, x_n$ from the APKumW distribution, with the unknown vector of parameter $\boldsymbol{\theta} = (a, b, c, \lambda, \alpha)$, then the log-likelihood

function (ℓ) can be defined as

$$\ell = n \log\left(\frac{\alpha\log(\alpha)}{\alpha-1}\right) + n \log(abc\lambda^c) + (c-1)\sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}(\lambda x_i)^c + (a-1)\sum_{i=1}^{n}\log(1 - e^{-(x_i\lambda)^c}) +$$
$$(b-1)\sum_{i=1}^{n}\log(1 - (1 - e^{-(x_i\lambda)^c})^a) - \log(\alpha)\sum_{i=1}^{n}[1 - (1 - e^{-(\lambda x_i)^c})^a]^b. \tag{20}$$

The associated nonlinear equations for the partial derivative of ℓ with respect to each parameter, are given as

$$\frac{\partial\ell}{\partial a} = \frac{n}{a} + \sum_{i=1}^{n}\log(1 - e^{-(\lambda x)^c})\left\{1 - \frac{(1 - e^{-(\lambda x)^c})^a}{1 - (1 - e^{-(\lambda x)^c})^a}\left[b(1 - \log(\alpha)[1 - (1 - e^{-(\lambda x)^c})^a]^b) - 1]\right]\right\}, \tag{21}$$

$$\frac{\partial\ell}{\partial b} = \frac{n}{b} + \sum_{i=1}^{n}\log[1 - (1 - e^{-(\lambda x)^c})^a]\left[1 - \log(\alpha)[1 - (1 - e^{-(\lambda x)^c})^a]^b\right], \tag{22}$$

$$\frac{\partial\ell}{\partial c} = \frac{n}{c}(1 + c\log(\lambda)) + \sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}(\lambda x_i)^c \log(\lambda x_i)\left\{1 - \frac{e^{-(\lambda x)^c}}{1 - e^{-(\lambda x)^c}}\left[a - 1 + \frac{a(1 - e^{-(\lambda x)^c})^a}{1 - (1 - e^{-(\lambda x)^c})^a} \times \right.\right.$$
$$\left.\left.(b(1 + \log(\alpha)[1 - (1 - e^{-(\lambda x)^c})^a]^b) - 1)\right]\right\}, \tag{23}$$

$$\frac{\partial\ell}{\partial\lambda} = \frac{nc}{\lambda} - c\lambda^{c-1}\sum_{i=1}^{n}x_i^c\left\{1 - \frac{e^{-(\lambda x)^c}}{1 - e^{-(\lambda x)^c}}\left[a - 1 - \frac{a(1 - e^{-(\lambda x)^c})^a}{1 - (1 - e^{-(\lambda x)^c})^a} \times \right.\right.$$
$$\left.\left.(b(1 - \log(\alpha)[1 - (1 - e^{-(\lambda x)^c})^a]^b) - 1)\right]\right\}, \tag{24}$$

and

$$\frac{\partial\ell}{\partial\alpha} = \frac{1}{\alpha}\left[\frac{n}{\log(\alpha)} - \frac{n}{\alpha-1} - \sum_{i=1}^{n}[1 - (1 - e^{-(\lambda x_i)^c})^a]^b\right]. \tag{25}$$

Then, the MLEs of the unknown parameters can be obtained by equating the equations from Eqs (21) to (25) to zero and solving them simultaneously. Particularly, a numerical iterative approach, such as the Newton-Raphson algorithm should be applied to solve these equations. Alternatively, any software like R, might be used to maximise Eq (20) directly and obtain the MLEs.

Studying survival times often results in the presence of censored observations, meaning there are incomplete observations of the period of interest. Right censoring technique is applied in medical studies when some patients lost to follow up and their exact occurrence time cannot be determined. The most common form of right censoring, which is encountered in survival analysis, is type I right censoring. A study of this type occurs when it is conducted over a specified period of time that will end before all units have failed. To illustrate, consider a study for a random sample of n patient in which, each patient is assigned a censoring time Y_i; $i = 1, \ldots, n$, that is the time between entry and the end of the study and where X_i; $i = 1, \ldots, n$, be the failure time of the ith patient. These X_i's and Y_i's are supposed to be independent and follow the APKumW distribution in Eq (6) and a non-informative distribution, respectively. For

$T_i = \min(X_i, Y_i)$, the pair (T_i, δ_i) is observed such that

$$\delta_i = \begin{cases} 1 & \text{if failure has occurred} \\ 0 & \text{if censoring has occurred} \end{cases}$$

Then, the log-likelihood function (ℓ) will be

$$\ell = \sum_{i=1}^{n} \delta_i \log[f(t_i)] + \sum_{i=1}^{n} (1 - \delta_i) \log[SF(t_i)], \tag{26}$$

where $f(.)$ and $SF(.)$ are respectively defined in Eqs (6) and (7). In order to obtain the MLEs, the log-likelihood in Eq (26) can be maximized numerically.

5 Simulation study

Some simulation studies are conducted to evaluate the performance of the MLEs for the five parameters of APKumW distribution. The simulation is considered over a number of iterations equal to $nsim = 1000$, for different sample sizes n with the following cases for the true parameters θ_{tr}

- **Case I**: $a = 0.5$, $b = 0.6$, $c = 1.3$, $\lambda = 0.2$, $\alpha = 0.1$

- **Case II**: $a = 0.8$, $b = 2.7$, $c = 3.6$, $\lambda = 2.4$, $\alpha = 1.2$

- **Case III**: $a = 1.3$, $b = 0.4$, $c = 1.6$, $\lambda = 0.5$, $\alpha = 0.2$

- **Case IV**: $a = 2.2$, $b = 1.6$, $c = 1.4$, $\lambda = 0.07$, $\alpha = 0.03$

The MLEs for each estimator $\hat{\theta}$ can be evaluated using an accuracy measurement, such as the root mean squared error (RMSE) that can be calculated as follows

$$RMSE(\hat{\theta}) = \sqrt{\frac{\sum_{i=1}^{nsim} (\hat{\theta}_i - \theta_{tr})^2}{nsim}} \tag{27}$$

All estimation results are obtained using the "optim" function in R software. Table 2 shows the results for the MLEs of the parameters of APKumW along with their corresponding RMSE. Generally, it can be seen from this table, that the MLEs are more closer to the true values of the parameters as the sample size increased. In addition, RMSE became smaller as sample size n increased, indicating that the estimates are consistent. These results demonstrate that maximum likelihood method is effective at estimating the parameters of the proposed distribution.

6 Applications

Six real datasets for cancer patients are fitted using the APKumW distribution. The results obtained using the APKumW distribution are compared against the corresponding ones achieved with the application of the following

The Weibull distribution with the following pdf

$$f(x) = \frac{c}{\lambda^c} x^{c-1} e^{-\left(\frac{x}{\lambda}\right)^c}.$$

Table 2. Simulation study: APKumW parameter estimates, together with the RMSE for three different cases with different sample sizes.

		Case I		Case II		Case III		Case IV	
		MLE	RMSE	MLE	RMSE	MLE	RMSE	MLE	RMSE
$n = 25$	a	0.7374	1.1311	1.1366	1.7467	3.1581	12.3097	4.1684	5.8308
	b	1.3078	1.7491	2.9682	4.1624	0.4752	0.6268	1.5157	1.8984
	c	2.2590	1.8346	6.1786	5.0080	2.0867	1.7713	2.2103	2.1365
	λ	0.4063	0.8406	2.9813	1.7717	1.0853	1.3514	0.1890	0.2221
	α	0.4441	1.7620	2.8268	8.8754	0.8573	1.8160	0.5149	1.0372
$n = 50$	a	0.5936	0.7352	0.8191	0.8192	2.2807	4.3226	3.7718	4.3232
	b	1.1829	1.3862	3.0076	3.9600	0.4590	0.4819	1.4895	1.8195
	c	1.8486	1.2120	5.6346	3.9938	1.7671	0.9930	1.7707	1.1602
	λ	0.3280	0.6997	2.7180	1.0426	0.8880	0.9771	0.1720	0.1899
	α	0.3669	1.0412	2.9617	6.8317	0.6719	1.3418	0.4554	0.9540
$n = 100$	a	0.5002	0.3808	0.7360	0.5757	1.7846	1.3917	3.4064	3.2416
	b	1.0501	0.9617	2.8792	2.8078	0.4423	0.4725	1.5827	1.5955
	c	1.7149	0.9487	5.2997	3.2607	1.5933	0.4265	1.5137	0.6138
	λ	0.2617	0.4597	2.5268	0.6980	0.7591	0.6132	0.1450	0.1497
	α	0.2640	0.6310	2.6456	5.7300	0.6315	1.2849	0.3413	0.7541
$n = 500$	a	0.4883	0.1427	0.7330	0.2806	1.5301	0.6641	2.6718	1.3610
	b	0.8331	0.5439	3.0105	1.8025	0.4243	0.3008	1.6417	0.9258
	c	1.4369	0.4402	4.4465	1.9672	1.5428	0.2977	1.3714	0.3012
	λ	0.2056	0.1635	2.4009	0.4645	0.6278	0.3305	0.0939	0.0577
	α	0.1791	0.2369	1.9772	2.8766	0.5035	0.8366	0.1442	0.3296
$n = 1000$	a	0.4891	0.1038	0.7486	0.2159	1.4622	0.4818	2.4720	0.8942
	b	0.7738	0.4584	2.8873	1.3060	0.4324	0.2165	1.6791	0.7932
	c	1.3883	0.3229	4.1520	1.4735	1.5478	0.2613	1.3679	0.2323
	λ	0.1986	0.1182	2.3915	0.3327	0.5743	0.2204	0.0809	0.0341
	α	0.1589	0.1963	1.5851	1.6256	0.4032	0.5784	0.0803	0.1853

The Exponenetiated generalized Weibull (EGW) distribution [46] with the following pdf

$$f(x) = ab \frac{c}{\lambda^c} x^{c-1} e^{-a\left(\frac{x}{\lambda}\right)^c} \left[1 - e^{-a\left(\frac{x}{\lambda}\right)^c}\right]^{b-1}.$$

The Beta Weibull (BW) distribution [4] with the following pdf

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{c}{\lambda^c} x^{c-1} e^{-b\left(\frac{x}{\lambda}\right)^c} \left[1 - e^{-\left(\frac{x}{\lambda}\right)^c}\right]^{a-1}.$$

The KumW distribution [21] with the following pdf

$$f(x) = ab \frac{c}{\lambda^c} x^{c-1} e^{-\left(\frac{x}{\lambda}\right)^c} \left(1 - e^{-\left(\frac{x}{\lambda}\right)^c}\right)^{a-1} \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^c}\right)^a\right]^{b-1}.$$

The exponentiated Kumaraswamy Weibull (EKumW) distribution [47] with the following pdf

$$f(x) = \alpha abc\lambda^c x^{c-1} e^{-(\lambda x)^c} (1 - e^{-(\lambda x)^c})^{a-1} [1 - (1 - e^{-(\lambda x)^c})^a]^{b-1} \times$$
$$\{1 - [1 - (1 - e^{-(\lambda x)^c})^a]^b\}^{\alpha-1}.$$

Table 3. MLEs, (SEs) for the parameters and associated goodness of fit statistics for the acute bone cancer data.

Distribution	MLE and SE	AIC	KS	P value
APKumW	α = 0.0046 (0.0055), a = 5.0887 (2.1344), b = 0.4137 (0.3181), c = 0.5358 (0.1424), λ = 1.3007 (0.7691)	291.7005	0.0680	0.8888
Weibull	c = 0.7656 (0.0568), λ = 2.9260 (0.4761)	326.8033	0.1887	0.0111
EGW	a = 2.7262 (4.3589), b = 80.5514 (119.2631), c = 0.2353 (0.0755), λ = 0.15070 (0.9666)	294.0796	0.0924	0.5612
BW	a = 59.2646 (35.3442), b = 62.3944 (32.4899), c = 0.1262 (0.0363), λ = 41.4347 (42.1657)	298.9643	0.0988	0.4747
KumW	a = 2.7498 (0.0095), b = 0.3506 (0.0423), c = 0.6483 (0.0035), λ = 0.3447 (0.0052)	311.4273	0.1470	0.0853
EKumW	a = 1.273 (0.4415), b = 1.9974 (1.0322), c = 0.4002 (0.0658), λ = 1.0345 (1.066), α = 5.4855 (2.0861)	302.2774	0.1168	0.272
APW	c = 0.9218 (0.078), λ = 0.0791 (0.0141), α = 0.0021 (0.001)	309.0348	0.1884	0.0112

https://doi.org/10.1371/journal.pone.0264229.t003

The alpha power Weibull (APW) distribution [32] with the following pdf

$$f(x; \boldsymbol{\theta}) = \begin{cases} \dfrac{\log(\alpha)}{\alpha - 1} \quad c\lambda x^{c-1} \quad e^{-\lambda x^c} \alpha^{1-e^{-\lambda x^c}} & ; \qquad \text{if } \alpha > 0, \alpha \neq 1 \\ c\lambda x^{c-1} e^{-\lambda x^c} & ; \qquad \text{if } \alpha = 1. \end{cases}$$

A variety of tools can be applied for comparing different competing distribution for a specific dataset and choosing the best model for the fitting. To investigate the goodness-of-fit for the compared distribution, Akaike Information Criterion (AIC) and Kolmogorov–Smirnov (KS) along with its P value are considered in order to choose the best distribution. The better distribution is which corresponds to the lower values of AIC, KS and highest P value of KS statistic. The plots of the estimated cdf for each of the distributions are compared with the plot of the empirical cdf. Also, the histogram of the observed frequencies is compared with the plots of the expected frequencies for each fitted distribution. The MLEs of the parameters for all the five datasets along with their SEs (in parentheses) and the corresponding goodness-of-fit criteria for all the competing models are respectively presented in Tables 3–7. Additionally,

Table 4. MLEs, (SEs) for the parameters and associated goodness of fit statistics for the head and Neck cancer data.

Distribution	MLE and SE	AIC	KS	P value
APKumW	a = 4.8428 (3.2155), b = 0.6523 (0.7809), c = 0.5719 (0.2531), λ = 0.026 (0.0106), α = 0.3173 (1.0202)	565.1112	0.0751	0.9492
Weibull	c = 0.9386 (0.1007), λ = 213.6881 (36.2325)	567.6877	0.1267	0.4435
EGW	a = 0.0784 (0.0131), b = 1.6582 (0.3456), c = 0.3437 (0.0022), λ = 0.0513 (0.0028)	602.3591	0.2687	0.0027
BW	a = 2.3728 (1.1206), b = 0.0759 (0.0119), c = 0.3999 (0.002), λ = 0.2225 (0.0025)	602.5257	0.3075	0.0003
KumW	a = 0.3532 (0.1743), b = 0.0782 (0.0122), c = 0.4911 (0.0025), λ = 1.2577 (0.0025)	604.0207	0.3142	0.0002
EKumW	a = 7.8983 (2.4909), b = 6.8318 (3.9697), c = 0.2203 (0.0377), λ = 0.0541 (0.0336), α = 1.8148 (1.1851)	566.0263	0.0973	0.7625
APW	c = 0.8779 (0.077), λ = 0.0105 (0.0089), α = 1.6918 (2.2826)	570.2769	0.1277	0.4342

https://doi.org/10.1371/journal.pone.0264229.t004

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

Distribution	MLE and SE	AIC	KS	P value
APKumW	a = 0.1192 (0.0576), b = 1.493 (0.2111), c = 12.8763 (0.0026), λ = 0.1888 (0.0025), α = 3.2018 (8.3932)	139.4392	0.0625	0.9976
Weibull	c = 2.4993 (0.3370), λ = 3.5183 (0.2316)	143.1159	0.1185	0.6284
EGW	a = 0.0814 (0.0147), b = 1.4848 (0.3171), c = 1.7819 (0.0025), λ = 0.7165 (0.0025)	149.7559	0.1495	0.3330
BW	a = 0.3617 (0.1295), b = 0.0431 (0.0071), c = 1.008 (0.0023), λ = 0.1524 (0.0023)	185.7482	0.3171	0.0006
KumW	a = 0.9562 (0.0734), b = 0.0838 (0.0133), c = 2.5959 (0.0083), λ = 1.3823 (0.0083)	146.7387	0.0975	0.8411
EKumW	a = 3.2765 (0.0021), b = 1.3117 (0.0677), c = 4.1917 (0.0021), λ = 0.2253 (0.002), α = 0.1367 (0.0218)	142.2623	0.1039	0.7806
APW	c = 2.101 (0.4359), λ = 0.1058 (0.0827), α = 5.3518 (6.7462)	143.3641	0.0919	0.888

https://doi.org/10.1371/journal.pone.0264229.t005

Table 6. MLEs, (SEs) for the parameters and associated goodness of fit statistics for the bladder cancer I data.

Distribution	MLE and SE	AIC	KS	P value
APKumW	a = 0.0692 (0.0859), b = 0.8434 (0.1146), c = 9.3634 (0.0079), λ = 0.3289 (0.0075), α = 8.4765 (35.1643),	100.1924	0.0806	0.9736
Weibull	c = 1.9570 (0.2816), λ = 2.1645 (0.1918)	106.7741	0.165	0.2753
EGW	a = 0.0478 (0.0141), b = 0.3927 (0.0748), c = 3.5985 (0.0036), λ = 1.2523 (0.0043)	104.5355	0.1490	0.4015
BW	a = 0.8648 (0.3929), b = 0.0587 (0.0101), c = 1.6055 (0.0016), λ = 0.3582 (0.0016)	110.6557	0.1995	0.1139
KumW	a = 0.2273 (0.0026), b = 0.0487 (0.0081), c = 1.2286 (0.0014), λ = 0.1863 (0.0013)	121.6362	0.2092	0.0855
EKumW	a = 0.2692 (0.1951), b = 0.1116 (0.0855), c = 3.7307 (0.9426), λ = 0.6668 (0.1644), α = 0.7241 (0.3537)	104.0068	0.0904	0.9302
APW	c = 1.6796 (0.3636), λ = 0.3915 (0.2099), α = 4.5688 (5.7946)	107.2964	0.1406	0.4751

Table 7. MLEs, (SEs) for the parameters and associated goodness of fit statistics for the bladder cancer II data.

Distribution	MLE and SE	AIC	KS	P value
APKumW	a = 1.4958 (0.8164), b = 0.3964 (1.1823), c = 1.0028 (0.4566), λ = 0.1139 (0.2304), α = 0.0168 (0.0409),	829.509	0.0370	0.9947
Weibull	c = 1.0478 (0.0676), λ = 9.5607 (0.8529)	832.1738	0.0700	0.5570
EGW	a = 0.1608 (0.0169), b = 1.4844 (0.185), c = 0.7023 (0.0026), λ = 0.4745 (0.0026)	841.2811	0.1091	0.0947
BW	a = 0.8742 (0.1428), b = 0.163 (0.0157), c = 0.8088 (0.0026), λ = 0.9306 (0.0026)	854.7262	0.1755	0.0007
KumW	a = 0.8027 (0.1796), b = 0.1707 (0.0164), c = 0.8367 (0.0026), λ = 1.079 (0.0026)	852.5814	0.1664	0.0017
EKumW	a = 1.2965 (0.6309), b = 0.7662 (0.7314), c = 0.6749 (0.1550), λ = 0.4087 (0.4881), α = 2.1286 (1.1644)	831.4566	0.0418	0.9786
APW	c = 0.9553 (0.088), λ = 0.1449 (0.0656), α = 2.1243 (1.8644)	834.9494	0.0725	0.5115

Figs 2–6 display on lefts the empirical and the fitted cdfs, and on rights the fitted pdfs with histogram of the observed datasets in Tables 3–7.

6.1 Acute bone cancer dataset

[48] considered a simulated data represents the survival times (in days) of 73 patients who diagnosed with acute bone cancer, as follows: 0.09, 0.76, 1.81, 1.10, 3.72, 0.72, 2.49, 1.00, 0.53,
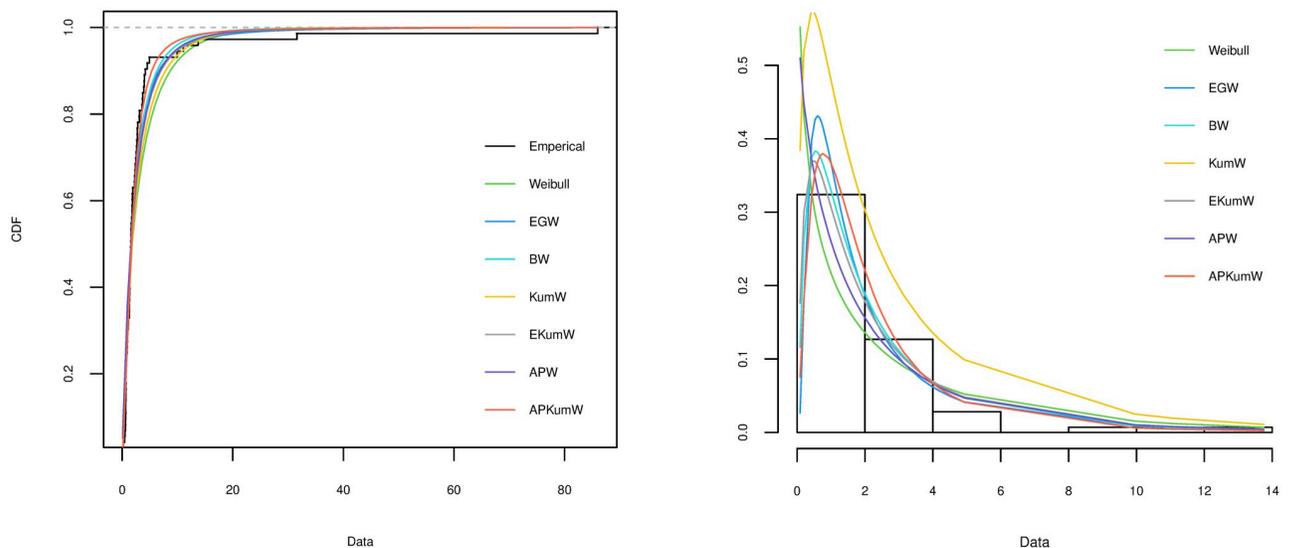

Fig 2. Theoretical and empirical cdf and pdf comparison of the acute bone cancer data.
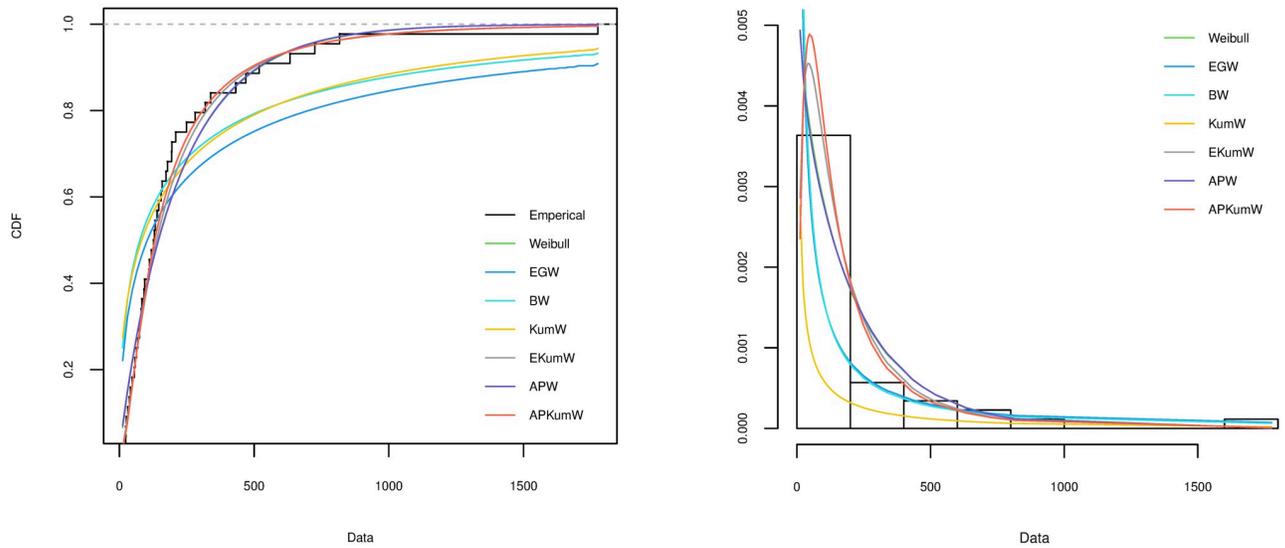
Fig 3. Theoretical and empirical cdf and pdf comparison of the head and Neck cancer data.

0.66, 31.61, 0.60, 0.20, 1.61, 1.88, 0.70, 1.36, 0.43, 3.16, 1.57, 4.93, 11.07, 1.63, 1.39, 4.54, 3.12, 86.01, 1.92, 0.92, 4.04, 1.16, 2.26, 0.20, 0.94, 1.82, 3.99, 1.46, 2.75, 1.38, 2.76, 1.86, 2.68, 1.76, 0.67, 1.29, 1.56, 2.83, 0.71, 1.48, 2.41, 0.66, 0.65, 2.36, 1.29, 13.75, 0.67, 3.70, 0.76, 3.63, 0.68, 2.65, 0.95, 2.30, 2.57, 0.61, 3.93, 1.56, 1.29, 9.94, 1.67, 1.42, 4.18, 1.37. This medical dataset is considered and analyzed using the APKumW and the competitive models.

6.2 Head and Neck cancer dataset

Survival time for 44 patients diagnosed by Head and Neck cancer disease from [49] and analyzed recently by [50] is considered. The dataset are: 12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35,
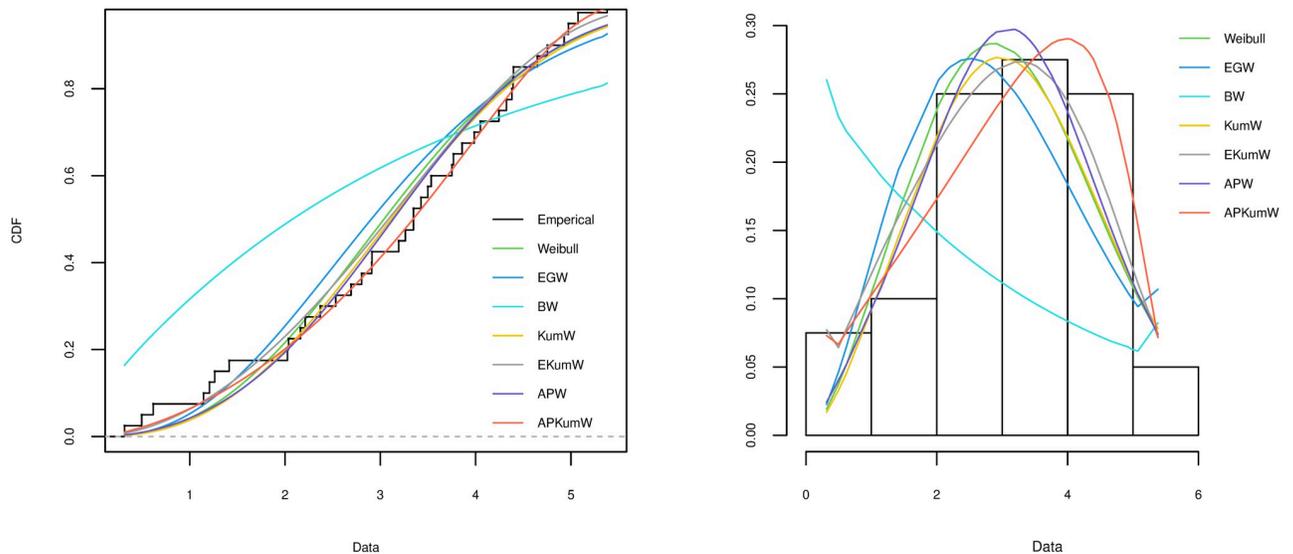

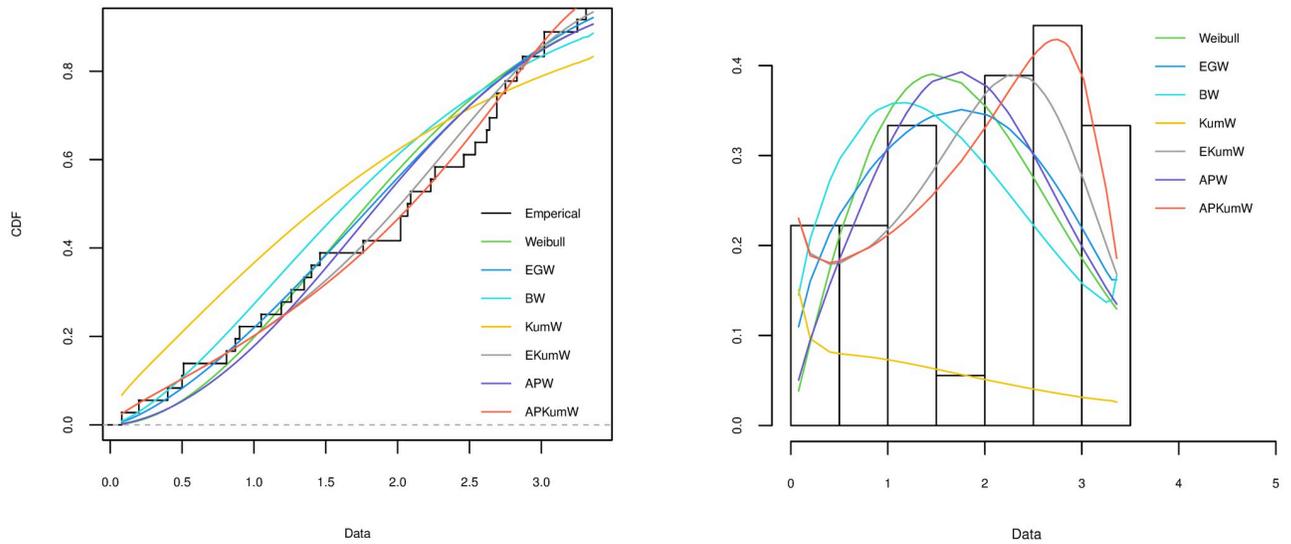**Fig 4. Theoretical and empirical cdf and pdf comparison of the blood cancer data.**

Fig 5. Theoretical and empirical cdf and pdf comparison of the bladder cancer I data.

47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 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, 1776.

6.3 Blood cancer dataset

This data consists of the life time (in years) of a 40 blood cancer (leukemia) patients from one of Ministry of health hospitals in Saudi Arabia reported in [51]. This actual data are: 0.315, 0.496, 0.616, 1.145, 1.208, 1.263, 1.414, 2.025, 2.036, 2.162, 2.211, 2.370, 2.532, 2.693, 2.805, 2.910, 2.912, 3.192, 3.263, 3.348, 3.348, 3.427, 3.499, 3.534, 3.767, 3.751, 3.858, 3.986, 4.049, 4.244, 4.323, 4.381, 4.392, 4.397, 4.647, 4.753, 4.929, 4.973, 5.074, 5.381.
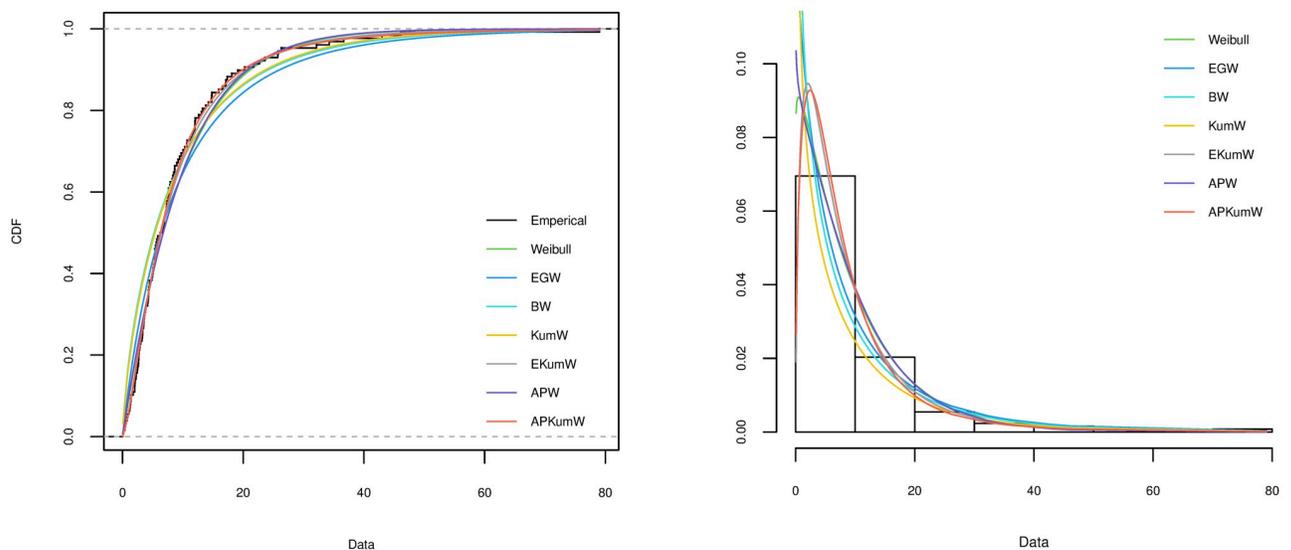

Fig 6. Theoretical and empirical cdf and pdf comparison of the bladder cancer II data.

6.4 Bladder cancer I dataset

The dataset on the remission times (in months) of a 36 bladder cancer patients reported in [52] and given by: 0.08, 0.2, 0.4, 0.5,0.51, 0.81, 0.87, 0.9, 1.05, 1.19, 1.26, 1.35, 1.4, 1.46, 1.76, 2.02, 2.02, 2.07, 2.09, 2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.02, 3.25, 3.31, 3.36, 3.36.

6.5 Bladder cancer II dataset

This application is from [53] and it is about the remission times (in months) of a 128 patients suffering from bladder cancer. This data has been analyzed recently in many papers, such as [54, 55]. The dataset values are as follows: 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

On the basis of the results presented in Tables 3–7, it can be seen that APKumW is the best model of all fitted distributions, resulting in the lowest values for AIC, KS and highest P values of KS statistic across all datasets. Besides, Figs 2–6 demonstrate the adequacy of the suggested APKumW distribution due to its closed fit to the observed cancer datasets. Thus, it can be concluded that the APKumW would be preferred to the other distributions applied to the considered cancer datasets.

6.6 Censored dataset

Listed below are the ordered remission times (in months) of a random sample of 137 bladder cancer patients from [53]: 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 24.80*, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 0.87*, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 10.86*, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 4.33*, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 3.02*, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 19.36*, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 4.65*, 6.76, 8.60*, 12.07, 21.73, 2.07, 3.36, 4.70*, 6.93, 8.65, 12.63, 22.69. The asterisk * means censored data.

Table 8 shows the MLEs, SEs of the unknown parameters of the APKumW distribution for the censored data obtained by maximizing the log-likelihood function in Eq (26). The table

Table 8. MLE, (SE) for the parameters and associated goodness of fit statistics for the censored data.

Distribution	MLE and SE	AIC
APKumW	a = 1.4053 (0.6525), b = 0.3253 (1.0591), c = 1.0518 (0.3953), λ = 0.1309 (0.2918), α = 0.0273 (0.0637),	846.8787
Weibull	c = 1.0164 (0.0655), λ = 0.1106 (0.0099)	847.7042
EKumW	a = 0.4527 (0.1008), b = 0.2853 (0.0353), c = 1.3493 (0.0028), λ = 0.1782 (0.0023), α = 0.9005 (0.1333)	875.5156

also displays the MLEs, SEs of the unknown parameters of the Weibull and exponentiated Kumaraswamy Weibull (EKumW) distributions based on the censored cancer data. As shown by the lowest AIC for the APKumW, it appears that the distribution can fit censored data well.

7 Conclusion

Choosing a suitable model for fitting survival data has been a major concern among researchers. One of the most popular distributions for life-time data is the Weibull distribution. In this paper, the Weibull distribution is extended to provide a new distribution called the APKumW to model life time data. It has different special cases which have been presented in the paper. A number of statistical characteristics of the proposed distribution have been studied, including survival and hazard functions, quantiles, moments, Rényi entropy and order statistics. Inference of parameters for an APKumW was obtained using the method of maximum likelihood. The estimates have been evaluated via different simulation studies. A good performance is observed when the parameters have been estimated using the maximum likelihood method. The applications of statistical distributions are essential for medical research and can have a crucial impact on public health, especially for cancer patients. Thus, the usefulness of this distribution is illustrated through its applications to some real datasets that describe the survival of some cancer patients, including both complete and censored cases. The results indicate the superior performance of the APKumW distribution compared to other competitive distributions by means of different goodness-of-fit criteria. Overall, it is hoped that the proposed APKumW distribution will provide an alternative to other existing distributions available for modeling positive skewed real data in survival analysis, especially for cancer research.

Acknowledgments

Author Contributions

Conceptualization: Hadeel S. Klakattawi.

Data curation: Hadeel S. Klakattawi.

Formal analysis: Hadeel S. Klakattawi.

Funding acquisition: Hadeel S. Klakattawi.

Investigation: Hadeel S. Klakattawi.

Methodology: Hadeel S. Klakattawi.

Project administration: Hadeel S. Klakattawi.

Resources: Hadeel S. Klakattawi.

Software: Hadeel S. Klakattawi.

Supervision: Hadeel S. Klakattawi.

Validation: Hadeel S. Klakattawi.

Visualization: Hadeel S. Klakattawi.

Writing – original draft: Hadeel S. Klakattawi.

Writing – review & editing: Hadeel S. Klakattawi.

References

1. Weibull W. A statistical distribution function of wide applicability. Journal of applied mechanics. 1951; 18(3):293–297. https://doi.org/10.1115/1.4010337

2. Mudholkar GS, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE transactions on reliability. 1993; 42(2):299–302. https://doi.org/10.1109/24.229504

3. Gupta RC, Gupta PL, Gupta RD. Modeling failure time data by Lehman alternatives. Communications in Statistics-Theory and methods. 1998; 27(4):887–904. https://doi.org/10.1080/03610929808832134

4. Famoye F, Lee C, Olumolade O. The beta-Weibull distribution. Journal of Statistical Theory and Applications. 2005; 4(2):121–136.

5. Eugene N, Lee C, Famoye F. Beta-normal distribution and its applications. Communications in Statistics-Theory and methods. 2002; 31(4):497–512. https://doi.org/10.1081/STA-120003130

6. Ghitany M, Al-Hussaini E, Al-Jarallah R. Marshall–Olkin extended Weibull distribution and its application to censored data. Journal of Applied Statistics. 2005; 32(10):1025–1034. https://doi.org/10.1080/02664760500165008

7. Marshall AW, Olkin I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika. 1997; 84(3):641–652. https://doi.org/10.1093/biomet/84.3.641

8. Ishaq AI, Abiodun AA. The Maxwell–Weibull Distribution in Modeling Lifetime Datasets. Annals of Data Science. 2020; 7:639–662. https://doi.org/10.1007/s40745-020-00288-8

9. Mansour MM, Butt NS, Ansari SI, Yousof HM, Ali MM, Ibrahim M. A new exponentiated Weibull distribution's extension: copula, mathematical properties and applications. Contributions to Mathematics. 2020; 1:57–66.

10. Vila R, Niyazi Çankaya M. A bimodal Weibull distribution: properties and inference. Journal of Applied Statistics. 2021; p. 1–19. https://doi.org/10.1080/02664763.2021.1931822

11. Cordeiro GM, de Castro M. A new family of generalized distributions. Journal of statistical computation and simulation. 2011; 81(7):883–898. https://doi.org/10.1080/00949650903530745

12. Cordeiro GM, Nadarajah S, Ortega EM. The Kumaraswamy Gumbel distribution. Statistical Methods & Applications. 2012; 21(2):139–168. https://doi.org/10.1007/s10260-011-0183-y

13. Saulo H, Leão J, Bourguignon M. The kumaraswamy birnbaum-saunders distribution. Journal of Statistical Theory and Practice. 2012; 6(4):745–759. https://doi.org/10.1080/15598608.2012.719814

14. Paranaíba PF, Ortega EM, Cordeiro GM, Pascoa MAd. The Kumaraswamy Burr XII distribution: theory and practice. Journal of Statistical Computation and Simulation. 2013; 83(11):2117–2143. https://doi.org/10.1080/00949655.2012.683003

15. Gomes AE, da Silva CQ, Cordeiro GM, Ortega EM. A new lifetime model: the Kumaraswamy generalized Rayleigh distribution. Journal of statistical computation and simulation. 2014; 84(2):290–309. https://doi.org/10.1080/00949655.2012.706813

16. Nassar MM. The Kumaraswamy Laplace Distribution. Pakistan Journal of Statistics and Operation Research. 2016; p. 609–624. https://doi.org/10.18187/pjsor.v12i4.1485

17. Usman RM, Haq M, Talib J. Kumaraswamy half-logistic distribution: properties and applications. J Stat Appl Probab. 2017; 6:597–609. https://doi.org/10.18576/jsap/060315

18. Ibrahim NA, Khaleel MA. Exponentiated Kumaraswamy Exponentiated Weibull Distribution with Application. Jurnal Karya Asli Lorekan Ahli Matematik Vol. 2018; 11(1):015–022.

19. George R, Thobias S. Kumaraswamy Marshall-Olkin Exponential distribution. Communications in Statistics-Theory and methods. 2019; 48(8):1920–1937. https://doi.org/10.1080/03610926.2018.1440594

20. Tahir MH, Cordeiro GM, Mansoor M, Zubair M, Alzaatreh A. The Kumaraswamy Pareto IV Distribution. Austrian Journal of Statistics. 2021; 50(5):1–22. https://doi.org/10.17713/ajs.v50i5.96

21. Cordeiro GM, Ortega EM, Nadarajah S. The Kumaraswamy Weibull distribution with application to failure data. Journal of the Franklin Institute. 2010; 347(8):1399–1429. https://doi.org/10.1016/j.jfranklin.2010.06.010

22. Al-Dayian G, El-Helbawy A, Rezk H. Statistical inference for a simple constant stress model based on censored sampling data from the Kumaraswamy Weibull distribution. International Journal of Statistics and Probability. 2014; 3(3):80. https://doi.org/10.5539/ijsp.v3n3p80

23. El-Helbawy AAA, Al-Dayian GR, Rezk HR. Bayesian Approach for Constant-Stress Accelerated Life Testing for Kumaraswamy Weibull Distribution with Censoring. Pakistan Journal of Statistics and Operation Research. 2016; 12(3):407–428. https://doi.org/10.18187/pjsor.v12i3.1171

24. Mandouh R. Bayesian inference from the kumaraswamy-weibull distribution with applications to real data. International Journal of Contemporary Mathematical Sciences. 2016; 11(3):119–129. https://doi.org/10.12988/ijcms.2016.51162

25. Fachini-Gomes JB, Ortega EM, Cordeiro GM, Suzuki AK. The Bivariate Kumaraswamy Weibull regression model: a complete classical and Bayesian analysis. Communications for Statistical Applications and Methods. 2018; 25(5):523–544. https://doi.org/10.29220/CSAM.2018.25.5.523

26. Ahmed MA, Mahmoud MR, ElSherpieny EA. The new Kumaraswamy Kumaraswamy Weibull distribution with application. Pakistan Journal of Statistics and Operation Research. 2016; p. 165–184. https://doi.org/10.18187/pjsor.v12i1.1129

27. Ahmed MA, Mahmoud MR, ElSherbini EA. The new Kumaraswamy Kumaraswamy family of generalized distributions with application. Pakistan Journal of Statistics and Operation Research. 2015; p. 159–180. https://doi.org/10.18187/pjsor.v11i2.969

28. Silva R, Gomes-Silva F, Ramos M, Cordeiro G, Marinho P, De Andrade TA. The exponentiated Kumaraswamy-G Class: General properties and application. Revista Colombiana de Estadística. 2019; 42 (1):1–33. https://doi.org/10.15446/rce.v42n1.66205

29. Khan MS, King R, Hudson IL. Transmuted Kumaraswamy Weibull Distribution with Covariates Regression Modelling to Analyze Reliability Data. Journal of Statistical Theory and Applications. 2020;. https://doi.org/10.2991/jsta.d.201016.003

30. Shaw WT, Buckley IR. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv preprint arXiv:09010434. 2009;.

31. Mahdavi A, Kundu D. A new method for generating distributions with an application to exponential distribution. Communications in Statistics-Theory and Methods. 2017; 46(13):6543–6557. https://doi.org/10.1080/03610926.2015.1130839

32. Nassar M, Alzaatreh A, Mead M, Abo-Kasem O. Alpha power Weibull distribution: Properties and applications. Communications in Statistics-Theory and Methods. 2017; 46(20):10236–10252. https://doi.org/10.1080/03610926.2016.1231816

33. Dey S, Sharma VK, Mesfioui M. A new extension of Weibull distribution with application to lifetime data. Annals of Data Science. 2017; 4(1):31–61. https://doi.org/10.1007/s40745-016-0094-8

34. Ramadan DA, Magdy WA. On the Alpha-Power Inverse Weibull Distribution. International Journal of Computer Applications. 2018; 975:8887.

35. Basheer AM. Alpha power inverse Weibull distribution with reliability application. Journal of Taibah University for Science. 2019; 13(1):423–432. https://doi.org/10.1080/16583655.2019.1588488

36. Ceren Ü, Cakmakyapan S,Gamze Ö. Alpha power inverted exponential distribution: Properties and application. Gazi University Journal of Science. 2018; 31(3):954–965.

37. Hassan A, Mohamd R, Elgarhy M, Fayomi A. Alpha power transformed extended exponential distribution: properties and applications. Journal of Nonlinear Sciences and Applications. 2018; 12(4):62–67. https://doi.org/10.22436/jnsa.012.04.05

38. Hassan AS, Elgarhy M, Mohamd RE, Alrajhi S. On the Alpha Power Transformed Power Lindley Distribution. Journal of Probability and Statistics. 2019; 2019. https://doi.org/10.1155/2019/8024769

39. Dey S, Ghosh I, Kumar D. Alpha-power transformed lindley distribution: properties and associated inference with application to earthquake data. Annals of Data Science. 2018; p. 1–28.

40. Dey S, Nassar M, Kumar D. Alpha power transformed inverse Lindley distribution: A distribution with an upside-down bathtub-shaped hazard function. Journal of Computational and Applied Mathematics. 2019; 348:130–145. https://doi.org/10.1016/j.cam.2018.03.037

41. ZeinEldin RA, Ahsan ul Haq M, Hashmi S, Elsehety M. Alpha power transformed inverse Lomax distribution with different methods of estimation and applications. Complexity. 2020; 2020. https://doi.org/10.1155/2020/4658596

42. Erdogan N, Bagci K, Arslan T, Celik HE. Alpha power Maxwell distribution: Properties and application. Journal of Mathematical Modeling. 2021; p. 1–14.

43. Ali M, Khalil A, Ijaz M, Saeed N. Alpha-Power Exponentiated Inverse Rayleigh distribution and its applications to real and simulated data. PloS one. 2021; 16(1):e0245253. https://doi.org/10.1371/journal.pone.0245253 PMID: 33444340

44. Baharith LA, Aljuhani WH. New Method for Generating New Families of Distributions. Symmetry. 2021; 13(4):726. https://doi.org/10.3390/sym13040726

45. Nassar M, Alzaatreh A, Abo-Kasem O, Mead M, Mansoor M. A new family of generalized distributions based on alpha power transformation with application to cancer data. Annals of Data Science. 2018; 5 (3):421–436. https://doi.org/10.1007/s40745-018-0144-5

46. Cordeiro GM, Ortega EM, da Cunha DC. The exponentiated generalized class of distributions. Journal of Data Science. 2013; 11(1):1–27. https://doi.org/10.6339/JDS.2013.11(1).1086

47. Eissa FH. The exponentiated Kumaraswamy–Weibull distribution with application to real data. International Journal of Statistics and Probability. 2017; 6(6):167–82. https://doi.org/10.5539/ijsp.v6n6p167

48. Mansour M, Yousof HM, Shehata WA, Ibrahim M. A new two parameter Burr XII distribution: properties, copula, different estimation methods and modeling acute bone cancer data. Journal of Nonlinear Science and Applications. 2020; 13(5):223–238. https://doi.org/10.22436/jnsa.013.05.01

49. Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. Journal of the American statistical Association. 1988; 83(402):414–425. https://doi.org/10.1080/01621459.1988.10478612

50. Sule I, Doguwa SI, Isah A, Jibril HM. Topp Leone Kumaraswamy-G Family of Distributions with Applications to Cancer Disease Data. Journal of Biostatistics and Epidemiology. 2020; 6(1):37–48.

51. Al-Saiary ZA, Bakoban RA. The Topp-Leone generalized inverted exponential distribution with real data applications. Entropy. 2020; 22(10):1144. https://doi.org/10.3390/e22101144 PMID: 33286913

52. Hibatullah R, Widyaningsih Y, Abdullah S. Marshall-Olkin extended power Lindley distribution with application. Jurnal Riset dan Aplikasi Matematika (JRAM). 2018; 2(2):84–92. https://doi.org/10.26740/jram.v2n2.p84-92

53. Lee ET, Wang J. Statistical methods for survival data analysis. vol. 476. John Wiley & Sons; 2003.

54. Hamdeni T, Gasmi S. The Marshall–Olkin generalized defective Gompertz distribution for surviving fraction modeling. Communications in Statistics-Simulation and Computation. 2020; p. 1–14. https://doi.org/10.1080/03610918.2020.1804937

55. Ijaz M, Mashwani WK, Göktaş A, Unvan YA. A novel alpha power transformed exponential distribution with real-life applications. Journal of Applied Statistics. 2021; p. 1–16.