Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/120232

Discrete Kumaraswamy Erlang-Truncated Exponential Distribution with Applications to Count Data

Hussein Eledum[1,*] *and Alaa R. El-Alosey* [2]

[1]Department of Statistics, Faculty of Science, University of Tabuk, Tabuk, KSA
[2]Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

Received: 30 Apr. 2022, Revised: 22 Oct. 2022, Accepted: 19 Nov. 2022
Published online: 1 May 2023

Abstract: This paper defines the discrete Kumaraswamy Erlang-truncated exponential distribution (DKw_ETE) using the general approach of discretizing a continuous distribution while retaining its survival function. The statistical properties of the DKw_ETE distribution, such as the quantile function, moments, moment generating function, Rényi entropy, and order statistics, are studied. The maximum likelihood (ML) method is utilized to estimate the model's parameters. In contrast, the stress-strength parameter is derived and estimated using the ML method. Finally, the proposed distribution's importance is explained by application to a real data set.

Keywords: Erlang-truncated exponential distribution, Discrete Kumaraswamy Erlang-truncated exponential distribution; survival function, maximum likelihood, Quantile functions, Order statistics, Stress-strength parameter.

1 Introduction

In many cases, the researcher is faced with data obtained from real-world phenomena that are discrete in nature, even for continuous variables. For example, in computer databases, the actual sample spaces are discrete; therefore, the corresponding random variables should be considered discrete distributions. In life testing trials for lifetime data sets, the life length of devices sometimes are impossible to be measured on a continuous scale, e.g. the lifetime of a switch is a discrete random variable in the case of an on/off switching mechanism. On the other hand, the reliability data is often expressed on a discrete scale, e.g. the number of runs, cycles, or shocks that a device can hold out before failing and the number of days of survival for lung cancer patients since starting therapy. In this context, some common discrete distributions are used as alternatives for continuous ones, e.g. geometric and negative binomial are considered alternatives for the exponential and gamma distributions, respectively. Unfortunately, these distributions are unsuitable in some cases because they have monotonic hazard functions. On the other hand, some discrete distributions used to model the count data, like Poisson, had suited only the integer and zero values. As a result, there is an urgent need to develop new discrete lifetime distributions that are more suitable for modelling different types of lifetime data sets.

In literature, diverse techniques have been developed to derive a discrete distribution based on the corresponding continuous one. One of these methods is to use the properties of a continuous distribution to infer similar properties of the discrete one. The other method is to study discrete lifetime as the integer part of a continuous lifetime [1,2].

Roy [3] deduced a relationship between different reliability measures and gained a unique determination of a bivariate geometric distribution based on a bivariate extension of a univariate characterizing property. This author concluded that a survival function could be used to explain the univariate geometric distribution as a discrete concentration of a corresponding exponential distribution. On the flip side, if the discretization of a continuous life distribution keeps the same functional form of the survival function, then most reliability measures and statistical properties of the distribution will remain unchanged. Consequently, Roy [4] introduced the discrete concentration concept and considered it a simple approach to generating a discrete life distribution model based on a continuous model. Many discrete distributions have been introduced using this idea, e.g. the discrete normal distribution [4], the discrete Burr and Pareto distributions [5], the discrete gamma distribution [6], the discrete generalized exponential

* Corresponding author e-mail: (heledum@yahoo.com) or (haladem@ut.edu.sa)

distribution of the second type [7], the exponentiated discrete Weibull distribution [8], the discrete Logistic distribution [9], the discrete Lindley distribution [10], the discrete generalized Rayleigh distribution [11], the discrete weighted exponential distribution [12], and the discrete additive Weibull geometric distribution [13].

The Erlang-truncated exponential (ETE) distribution proposed by El-Alosey [14] extends the standard one-parameter Exponential distribution. Many distributions have been proposed based on ETE, e.g. the Kumaraswamy Erlang-truncated exponential (Kw_ETE) [15], the Poisson exponentiated Erlang truncated exponential [16], Marshall–Olkin generalized Erlang-truncated exponential [17], the transmuted Erlang-truncated exponential (TETE) [18], the Extended Erlang-truncated exponential (EETE) [19], the gamma Log-logistic Erlang Truncated Exponential [20], and the beta Erlang-truncated exponential [21]. Recently, two discrete distributions have been introduced using the general approach of discretizing ETE distribution while retaining its survival function, including the discrete Erlang-truncated Exponential(DETE) [22] and the discrete extended Erlang-truncated Exponential (DEETE) [23].

This paper introduces a new discrete version of Kumaraswamy Erlang-truncated exponential distribution called discrete Kumaraswamy Erlang-truncated exponential distribution (DKw_ETE) using the general approach of discretizing a continuous distribution. We investigate some statistical properties of DKw_ETE and estimate its parameters using the maximum likelihood method. Moreover, We apply the proposed distribution to two real count data sets. The DKw_ETE distribution is also used to model some data sets concerning the death cases of COVID-19 collected from some Gulf Cooperation Council countries, involving Qatar, UAE and Bahrain.

The rest of this paper is structured as follows: The new proposed distribution is discussed in section 2, and some statistical properties such as quantile function, moment generating function, Entropy and order statistics are presented in section 3. Section refsec4 pertains to the maximum likelihood method to estimate DKw_ETE parameters. An application of the DKw_ETE to two right-skewed real count data sets and data sets regarding the death cases of COVID-19 for illustration is demonstrated in section 5. Finally, section 6 gives some concluding remarks.

The proposed distribution is motivated because it exhibits the four parameters that can be adapted to meet most over-dispersed or undispersed right-skewed count data sets.

1.1 Discretizing a continuous distribution

The idea of discretization of a given continuous random variable first proposed by [4]. This concept is illustrated as follows.

Suppose X be a continuous random variable with survival function $S_X(x)$, a discrete random variable Y can be defined as equal to $[X]$ that is floor of X that is largest integer less or equal to X. The probability mass function (*pmf*) of Y is then defined as

$$P(Y=y) = S_X(y) - S_X(y+1)$$

The *pmf* of the random variable Y thus defined may be viewed as a discrete concentration of the pdf of X, see [3,4].

1.2 The Kumaraswamy Erlang-Truncated Exponential distribution

The Kumaraswamy Erlang-truncated exponential distribution, which is a generalized Erlang-truncated exponential distribution (Kw_ETE), was researched by [15]. A non-negative random variable X has a Kw_ETE distribution with parameters $\alpha > 0, \beta > 0, \lambda > 0$ and $\theta > 0$, if the cdf is provided by

$$F_{Kw\_ETE}(x;\alpha,\ \beta,\ \lambda,\ \theta) = 1 - \left[1 - \left(1 - e^{-\beta(1-e^{-\lambda})x}\right)^{\alpha}\right]^{\theta},\ x > 0 \tag{1}$$

The parameters α, β and θ are shape parameters while λ is a scale parameter.
If $\theta = 1$, the EETE distribution is then obtained, and if, $\alpha = 1$ and $\theta = 1$ then the ETE distribution is obtained.
The corresponding pdf is as follows

$$\begin{aligned} f_{Kw\_ETE}(x;\alpha,\ \beta,\ \lambda,\ \theta) =& \alpha\theta\beta(1-e^{-\lambda})e^{-\beta(1-e^{-\lambda})x} \\ & \times \left[1 - e^{-\beta(1-e^{-\lambda})x}\right]^{\alpha-1} \\ & \times \left[1 - \left(1 - e^{-\beta(1-e^{-\lambda})x}\right)^{\alpha}\right]^{\theta-1},\ x > 0 \end{aligned} \tag{2}$$

The hazard rate functions of the Kw_ETE is

$$h_{EETE}(x;\alpha,\ \beta,\ \lambda) = \frac{\alpha\theta\beta(1-e^{-\lambda})e^{-\beta(1-e^{-\lambda})x}\left(1-e^{-\beta(1-e^{-\lambda})x}\right)^{\alpha-1}}{1-\left(1-e^{-\beta(1-e^{-\lambda})x}\right)^{\alpha}} \quad x>0 \tag{3}$$

The survival of the Kw_ETE is:

$$S_{Kw\_ETE}(x;\alpha,\ \beta,\ \lambda,\ \theta) = \left[1-\left(1-e^{-\beta(1-e^{-\lambda})x}\right)^{\alpha}\right]^{\theta} \quad x>0 \tag{4}$$

By using a series expansion

$$(1-z)^{a} = \sum_{i=0}^{\infty}(-1)^{i}\binom{a}{i}z^{i}$$

The survival of the Kw_ETE can be written on the form:

$$S_{Kw\_ETE}(x;\alpha,\ \beta,\ \lambda,\ \theta) = 1-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}e^{-(1-e^{-\lambda})\beta jx},\ x>0 \tag{5}$$

2 Discrete Kumaraswamy Erlang-Truncated Exponential distribution

This section pertains to deriving the proposed distribution discrete Kumaraswamy Erlang-truncated exponential (DKw_ETE) and exploring its probability mass and cumulative distribution functions. Moreover, some especial cases of the proposed DKw_ETE are derived.

2.1 Probability mass and cumulative distribution functions for DKw_ETE distribution

The proposed DKw_ETE distribution is produced by discretizing the re-parameterized version of the Kw_ETE distribution with parameters $\alpha>0, \beta>0, \lambda>0$ and $\theta>0$ using the concept of discretization given in subsection 1.1. Re-parameterization of Kw_ETE in Eq.(2) by setting $p = e^{-(1-e^{-\lambda})}$, we get the pmf of DKw_ETE distribution, as follows:

$$f_{DKw\_ETE}(y;\alpha,\ \beta,\ p,\ \theta) = \begin{cases} \left[1-\left(1-p^{\beta y}\right)^{\alpha}\right]^{\theta} - \left[1-\left(1-p^{\beta(y+1)}\right)^{\alpha}\right]^{\theta} \\ \sum_{i=0}^{\infty}\sum_{j=0}^{i\alpha}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}p^{\beta jy}\left(1-p^{\beta j}\right) \end{cases} \tag{6}$$

$$y=0,\ 1,\ 2,\ \ldots,\ \alpha\ >\ 0,\ \beta\ >\ 0,\ 0<p<1, \theta>0$$

When $\theta>0$ is an integer value, the above outside infinite sums stop at θ.
The cdf of DKW_ETE is given as:

$$F_{DKw\_ETE}(y;\alpha,\ \beta,\ p,\ \theta) = P(Y\leq y) = \begin{cases} 1-\left[1-\left(1-p^{\beta(y+1)}\right)^{\alpha}\right]^{\theta}, & y\geq 0 \\ 0, & y<0 \end{cases} \tag{7}$$

Figure 1 and Figure 2 respectively show the pmfs and cdfs of DKw_ETE distribution for different values of parameters $\alpha,\ \beta,\ p$ and θ.

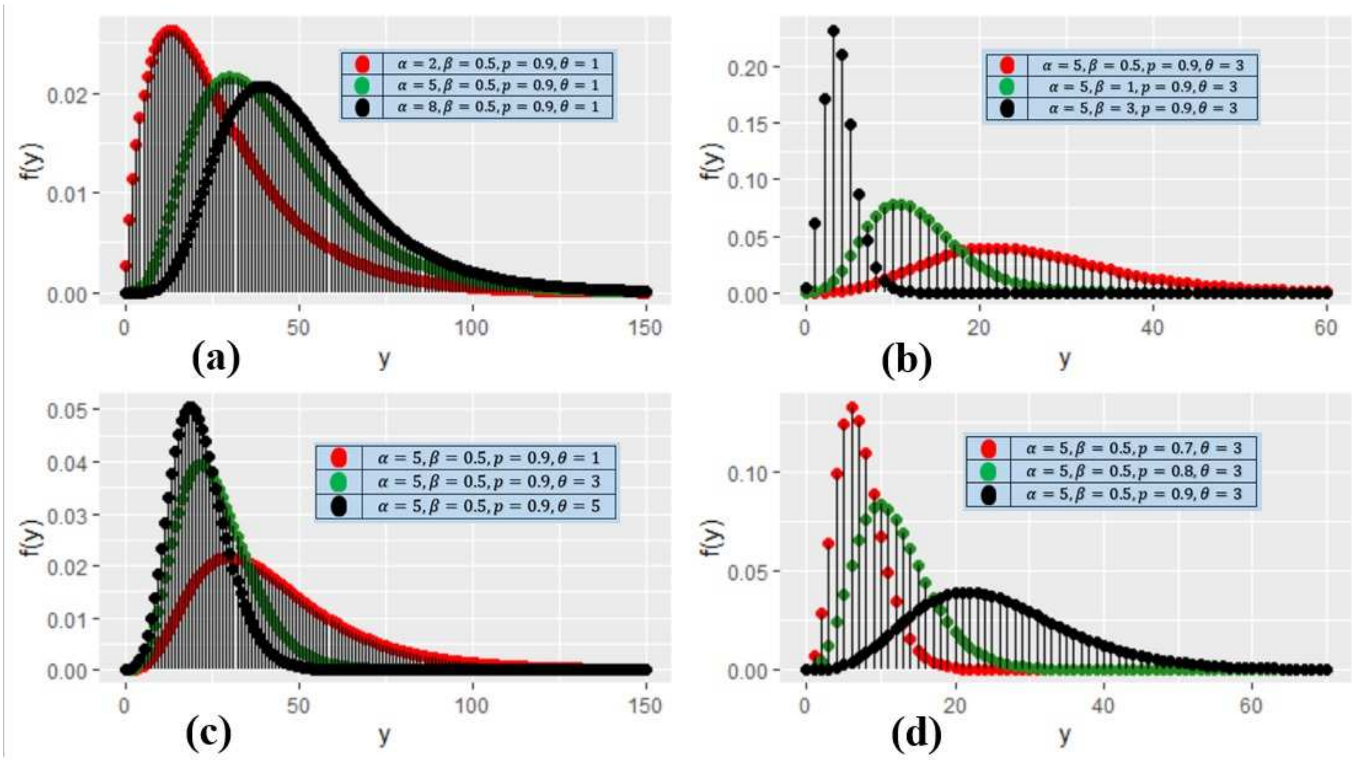

Fig. 1: pmfs of DKw_ETE distribution for combination values of parameters

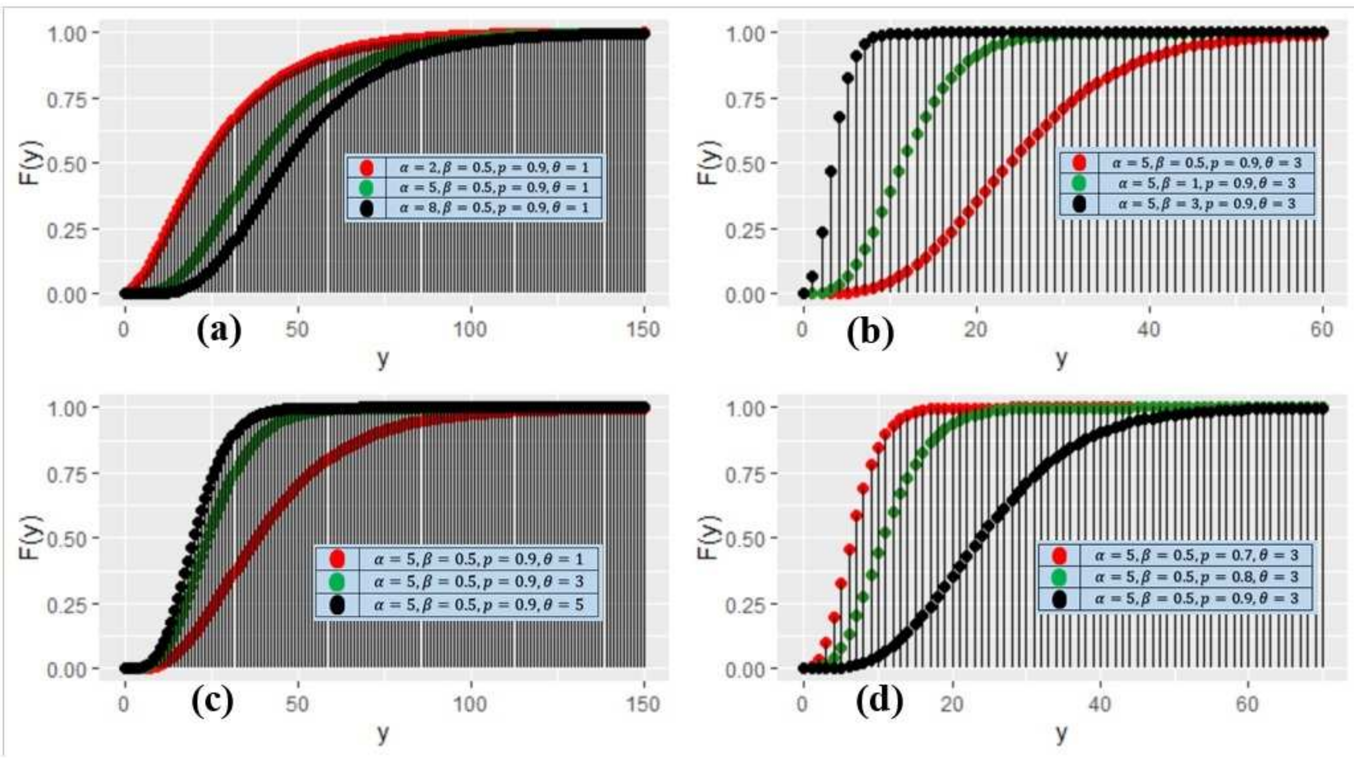

Fig. 2: cdfs of DKw_ETE distribution for combination values of parameters

From Figure 1, it can be seen that the proposed DKw_ETE distribution is a unimodal right-skewed. Moreover, with an increase in the parameter α, the coefficient of the skewness decreases (see Figure 1(a)); this is also observed from Figure 2(a); that is, the cdf approaches 1 faster for small values of α than large ones. The coefficient of skewness increases with β and θ (see Figures 1(b) and 1(c)). This is also noted from Figures 2(b) and 2(c); that is, the cdf approaches 1 faster for small values for both β and θ than large ones.

2.2 Some special cases of DKw_ETE

The DEETE can be concluded from Eq.(6) by taking $\theta = 1$ as follows [23]

$$f_{DEETE}(y;\alpha,\ \beta,\ p) = \left[1-p^{\beta(y+1)}\right]^{\alpha} - \left[1-p^{\beta y}\right]^{\alpha},\ \ y = 0,1,2,\ldots$$

The discrete generalized exponential distribution of the second type can be concluded from Eq.(6) by taking $\theta = 1$ and $\beta = 1$ as follows [7]

$$f_{\mathrm{DGE}_2}(y;\alpha,\beta) = \left[1-p^{y+1}\right]^{\alpha} - \left[1-p^{y}\right]^{\alpha},\ \ y = 0,1,2,\ldots$$

The pmf of the DETE distribution with parameters $\beta > 0,\ 0 < p < 1$ is explored, when $\alpha = \theta = 1$ in Eq.(6), as [22]

$$f_{DETE}(y;\ \beta,\ p) = p^{\beta y}\left(1-p^{\beta}\right),\ \ y = 0,1,2,\ldots.$$

The distribution of the random variable Y in Eq.(6), takes the form of geometric distribution, when $\alpha = \beta = \theta = 1$, as follows:

$$f_{Geom}(y;\ \ p) = p^{y}\left[1-p\right],\ \ y = 0,1,2,\ldots$$

2.3 Survival and Hazard function

The survival function of Y is given by:

$$S_{DKw\_ETE}(y;\alpha,\beta,p,\theta) = \left[1-\left(1-p^{\beta y}\right)^{\alpha}\right]^{\theta} \tag{8}$$

and the hazard function is:

$$\begin{aligned} h_{DKw\_ETE}(y;\alpha,\beta,p,\theta) =& \frac{f_{DKw\_ETE}(y;\alpha,\beta,p,\theta)}{S_{DKw\_ETE}(y;\alpha,\beta,p,\theta)} \\ =& 1-\left[\frac{1-\left(1-p^{\beta(y+1)}\right)^{\alpha}}{1-\left(1-p^{\beta y}\right)^{\alpha}}\right]^{\theta} \end{aligned}$$

Figure 3 illustrates the hazard function of DKw_ETE for different values of parameters $\alpha,\ \beta,\ p$ and θ. Based on Figure 3, it is observed that the hazard function of the proposed distribution is an increasing function and the rate of increase affected by the value each parameter takes.

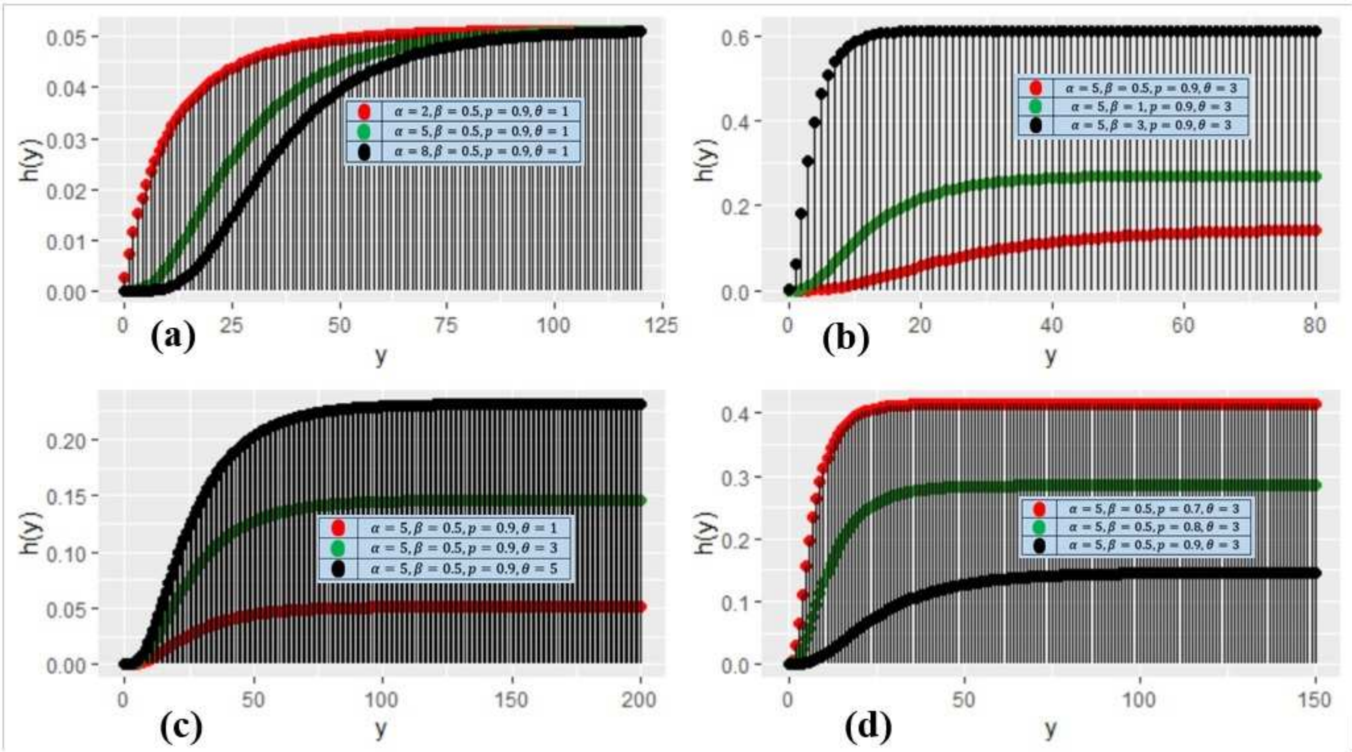

Fig. 3: Hazard function of DKw_ETE distribution for combination values of parameters

3 Distributional properties

In this section, the statistical properties of the DKw_ETE distribution are derived.

3.1 Quantile function

The quantile of order $0 < \gamma < 1$, can be obtained by inverting the cdf in Eq.(7) as

$$F_{Kw\_ETE}\left(y;\alpha,\ \beta,\ p,\ \theta\right) = 1-\left[1-\left(1-p^{\beta(y+1)}\right)^{\alpha}\right]^{\theta}$$

then

$$\begin{aligned} F^{-1}_{Kw\_ETE}\left(\gamma\right) &= min y \in R : F_{Kw\_ETE}\left(y\right) \geq \gamma \\ &= 1-\left[1-\left(1-p^{\beta(y+1)}\right)^{\alpha}\right]^{\theta} = \gamma \end{aligned}$$

Thus, The γ^{th} quantile is

$$Q(\gamma;\alpha,\beta,p,\theta) = \log_p\left\{1-\left[1-(1-\gamma)^{\frac{1}{\theta}}\right]^{\frac{1}{\alpha}}\right\}^{\frac{1}{\beta}} - 1 \tag{9}$$

Further the median of DKw_ETE obtained by substituting $\gamma = 1/2$ in Eq.(9) as follows:

$$Q_{\,0.5} = Q(\gamma;\alpha,\beta,p,\theta) = \log_p\left\{1-\left[1-\left(\frac{1}{2}\right)^{\frac{1}{\theta}}\right]^{\frac{1}{\alpha}}\right\}^{\frac{1}{\beta}} - 1$$

3.2 The moment generating function

In this subsection, the moment generating function of a random variable Y having a the DKw_ETE with parameters $(\alpha, \beta, p, \theta)$ is obtained.

$$M_Y(t) = E(e^{ty}) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}\frac{1-p^{\beta j}}{1-p^{\beta j}e^t}$$

Therefore,

$$M_Y(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}(1-p^{\beta j})p^{\beta jk}e^{tk}$$

where $\sum_{k=0}^{\infty} p^{\beta jk}e^{tk} = (1-p^{\beta j}e^t)^{-1}$ Thus we can calculate the first moment (the mean) of the DKw_ETE distribution as

$$E(Y) = \left.\frac{dM_Y(t)}{dt}\right|_{t=0} = \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+1}\binom{\theta}{i}\binom{i\alpha}{j}\frac{p^{\beta j}}{1-p^{\beta j}}$$

Also the r^{th} moment is

$$\begin{aligned}\mu_r = E(Y^r) = &\left.\frac{d^r M_Y(t)}{dt^r}\right|_{t=0}\\ =&\frac{d^r}{dt^r}\sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}\sum_{k=0}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}(1-p^{\beta j})p^{\beta jk}e^{tk}\\ =&\sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}\sum_{k=0}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}(1-p^{\beta j})k^r p^{\beta jk}\end{aligned}$$

Setting $r = 1$ we get the mean as follows

$$\begin{aligned}\mu_1 = E(Y) &= \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}(1-p^{\beta j})\sum_{k=0}^{\infty}kp^{\beta jk}\\ &= \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}\frac{p^{\beta j}}{(1-p^{\beta j})}\end{aligned} \tag{10}$$

where, $\sum_{k=0}^{\infty} kp^{\beta jk} = p^{\beta j}(1-p^{\beta j})^{-2}$
Setting $r = 2$ we obtain the second moment about the origin as

$$\begin{aligned}\mu_2 = E\left(Y^2\right) &= \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}\sum_{k=1}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}(1-p^{\beta j})k^2 p^{\beta jk}\\ &= \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}\frac{p^{\beta j}(1+p^{\beta j})}{(1-p^{\beta j})^2}\end{aligned}$$

The third moment about the origin is given by

$$\begin{aligned}\mu_3 = E\left(Y^3\right) &= \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}\sum_{k=1}^{\infty}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}(1-p^{\beta j})k^3 p^{\beta jk}\\ &= \sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}\frac{p^{\beta j}(1+4p^{\beta j}+p^{2\beta j})}{(1-p^{\beta j})^3}\end{aligned}$$

Thus, the variance of the DKw_ETE is given by

$$\begin{aligned}V(Y) =&\mu_2 - \mu_1^2\\ =&\sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}\frac{p^{\beta j}(1+p^{\beta j})}{(1-p^{\beta j})^2}\\ &-\left[\sum_{i=0}^{\alpha}\sum_{j=0}^{i\theta}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}\frac{p^{\beta j}}{(1-p^{\beta j})}\right]^2\end{aligned} \tag{11}$$

The skewness for the DKw_ETE can be calculated using the below formula

$$S_k = \frac{E\,(Y-\mu_1)^3}{\sigma^3} = \frac{\mu_3 - 3\mu_1\sigma^2 - \mu_1^3}{\sigma^3}$$

The mean, variance and coefficient of skewness for DKw_ETE distribution for different combinations of α, β, θ and p are computed in Table 1. From Table 1, it is observed that both mean and variance increase with α and p , while they decrease when β, and θ increase. The positive skewness decreases when α increases, while it increases with β and θ. Moreover, with the increase in the parameter p, the rate of increase in the mean and variance growing exponentially. As β and θ increase and α and p decrease, the $P(Y=0)$ approaches 1 and therefore, the mean and variance of the distribution approach zero, consequently, the distribution skews more extreme to the right. Therefore, we conclude that the four parameters of the proposed distribution can be adapted to meet most of the right-skewed count data sets.

Table 1: Mean and variance and skewness of DKw_ETE for combination values of parameters.

	p		0.1			0.5			0.9		
α	β	θ	**Mean**	**Variance**	**Skewness**	**Mean**	**Variance**	**Skewness**	**Mean**	**Variance**	**Skewness**
		1	0.81384	0.98235	1.58958	3.82843	10.4853	1.59292	27.97367	450.4996	1.60952
	0.5	3	0.15806	0.14781	2.2844	1.52091	1.83951	1.21436	12.78771	76.5538	1.26969
		5	0.04305	0.04169	4.59123	0.96314	0.91064	1.07916	9.11175	36.61563	1.13971
		1	0.21212	0.2163	2.27554	1.66667	2.66667	1.56495	13.73684	112.687	1.6082
2	1	3	0.00687	0.00684	11.9849	0.52053	0.48143	1.3263	6.14388	19.2001	1.26373
		5	0.00025	0.00025	63.5285	0.25407	0.22457	1.65757	4.30592	9.21508	1.1289
		1	0.002	0.002	22.3552	0.26984	0.27816	2.06793	4.24587	12.59081	1.59548
	3	3	0	0	11188.7	0.0129	0.0128	8.69385	1.71547	2.1987	1.21973
		5	0	0	5597164	0.00071	0.00071	37.5669	1.10352	1.08438	1.07666
		1	1.48049	1.20089	1.15938	6.08832	12.2685	1.32564	42.84325	527.4709	1.3389
	0.5	3	0.68734	0.36607	0.37545	3.4857	2.8207	0.7858	25.72091	118.5701	0.82089
		5	0.45668	0.27145	0.42392	2.77941	1.6382	0.59183	21.07419	67.39508	0.63853
		1	0.46407	0.37014	1.12885	2.79416	3.13039	1.28434	21.17162	131.9302	1.33795
5	1	3	0.06879	0.0643	3.42915	1.49302	0.76982	0.66878	12.61045	29.70503	0.8183
		5	0.01152	0.01138	9.15717	1.14038	0.4759	0.40526	10.2871	16.91126	0.63499
		1	0.005	0.00498	14.08579	0.57393	0.44365	1.03409	6.72388	14.73288	1.3279
	3	3	0	0	2836.925	0.116	0.10342	2.43773	3.87018	3.37446	0.79165
		5	0	0	568520.9	0.02742	0.02667	5.78929	3.09574	1.95288	0.59948
		1	1.86238	1.23093	1.14613	7.34208	12.79986	1.2539	51.09157	550.4641	1.26596
	0.5	3	1.06025	0.35509	0.35273	4.65718	3.1544	0.67889	33.42813	133.0006	0.70607
		5	0.84337	0.25436	0.023096	3.90785	1.90378	0.47995	28.49843	78.87248	0.51252
		1	0.65565	0.4177	0.71773	3.42108	3.26197	1.21885	25.29578	137.6785	1.26509
8	1	3	0.1852	0.15182	1.64423	2.07865	0.85011	0.61316	16.46406	33.31265	0.70408
		5	0.05993	0.05634	3.70886	1.70385	0.53717	0.42141	13.99921	19.78062	0.51009
	3	1	0.00798	0.00793	11.09394	0.79251	0.47728	0.72121	8.09859	15.3717	1.25596
	3	3	0	0	1404.896	0.28447	0.20688	1.00872	5.15469	3.77549	0.68345
		5	0	0	176227.7	0.12186	0.10706	2.31376	4.33307	2.27194	0.48532

3.3 Entropy

The most fundamental concept of information theory is entropy. Entropy is defined as an average amount of information per message. In Statistics, entropy is defined as a measure of uncertainty in a random variable. There are some types of entropy the most popular of them is the Rényi entropy, which has been computed as

$$I_R(\gamma) = \frac{1}{1-\gamma} log \left[\sum_{y=0}^{\infty} f^{\gamma}(y) \right]$$

where $\gamma > 0$ and $\gamma \neq 0$
the Renyi entropy for the a random variable Y having a *pmf* of the DKw_ETE distribution in Eq.(6) is given as

$$I_R(\gamma) = \frac{1}{1-\gamma}\log\left[\sum_{y=0}^{\infty} f^{\gamma}(y)\right]$$
$$=\frac{1}{1-\gamma}\log\left[\sum_{y=0}^{\infty}\left\{\sum_{i=0}^{\theta}\sum_{j=0}^{i\alpha}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}p^{\beta jy}\left(1-p^{\beta j}\right)\right\}^{\gamma}\right]$$

By using the complete multinomial expansions theorem [24], we get

$$\left\{\sum_{i=0}^{\theta}\sum_{j=0}^{i\alpha}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}p^{\beta jy}\left(1-p^{\beta j}\right)\right\}^{\gamma}$$
$$=\gamma!\sum_{p=1}^{M}\sum_{m_1=1}^{m-(k-1)}\sum_{m_2=m_1+1}^{m-(k-2)}\cdots\sum_{m_p=m_{p-1}+1}^{m}$$
$$\times\sum_{n_1=1}^{n-(k-1)}\sum_{n_2=1}^{n-n_1-(k-2)}\cdots$$
$$\sum_{n_{k-1}=1}^{n-n_1-n_2-\ldots-n_{k-2}-1}\sum_{n_p=n-n_1-n_2-\ldots-n_{k-1}}^{n-n_1-n_2-\ldots-n_{k-1}}$$
$$\times\prod_{i=1}^{k}\frac{\left[\sum j=0 i\alpha(-1)^{i+j}\binom{\alpha}{i}p^{\beta iy}\left(1-p^{\beta i}\right)\right]^{n_i}}{n_i!}$$

where $M = \begin{cases} n, & n < m \\ m, & n \geq m \end{cases}$
Therefore, the Renyi entropy measure can be written as

$$I_R(\gamma)$$
$$=\frac{1}{1-\gamma}\log\left\{\sum_{y=0}^{\infty}\gamma!\sum_{p=1}^{M}\sum_{m_1=1}^{m-(k-1)}\sum_{m_2=m_1+1}^{m-(k-2)}\cdots\sum_{m_p=m_{p-1}+1}^{m}\right.$$
$$\times\sum_{n_1=1}^{n-(k-1)}\sum_{n_2=1}^{n-n_1-(k-2)}\cdots\sum_{n_{k-1}=1}^{n-n_1-n_2-\ldots-n_{k-2}-1}\sum_{n_p=n-n_1-n_2-\ldots-n_{k-1}}^{n-n_1-n_2-\ldots-n_{k-1}}$$
$$\left.\times\prod_{i=1}^{k}\frac{\left[\sum_{j=0}^{i\alpha}(-1)^{i+j}\binom{\theta}{i}\binom{i\alpha}{j}p^{\beta jy}\left(1-p^{\beta j}\right)\right]^{n_i}}{n_i!}\right\}$$

When $\alpha = \theta = \beta = 1$ the same result as the geometric distribution achieved, when $\alpha = \theta = 1$ we get the Rényi entropy of DETE [22] and when $\alpha = 1$ we obtain the Rényi entropy of DEETE [23].

3.4 Order statistics

The most popular techniques used in nonparametric statistics are order statistics, defined as sample values placed in ascending order. Therefore, the present subsection pertains to deriving some order statistics for the DKw_ETE distribution. Let $F_i(y;\alpha,\beta,p,\theta)$ and $f_i(y;\alpha,\beta,p,\theta)$ be the cdf and pmf of the $i-th$ order statistic of a random sample of size n from DKw_ETE (α,β,p,θ).

We know that

$$F_i(y;\alpha,\beta,p,\theta) = \sum_{k=i}^{n}\binom{n}{k}\left[F(y;\alpha,\beta,p,\theta)\right]^{k} \times\left[1-F(y;\alpha,\beta,p,\theta)\right]^{n-k} \tag{12}$$

using the binomial expansion for $[1-F(y;\alpha,\beta,p,\theta)]^{n-k}$, we get the following result:

$$\begin{aligned}
&F_i(y;\alpha,\beta,p,\theta)\\
&=\sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j\left[F(y;\alpha,\beta,p,\theta)\right]^{k+j}\\
&=\sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j\left[1-\left[1-\left(1-p^{\beta(y+1)}\right)^{\alpha}\right]^{\theta}\right]^{(k+j)}\\
&=\sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{(k+j)}\sum_{h=0}^{(\theta l)}\sum_{g=0}^{(\alpha h)}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}\binom{\theta l}{h}\binom{\alpha h}{g}(-1)^{j+l+h+g}p^{\beta g(y+1)}
\end{aligned}$$

The corresponding *pmf* of the *ith* order statistic

$$f_i(y;\alpha,\beta,p,\theta)=F_i(y;\alpha,\beta,p,\theta)-F_i(y-1;\alpha,\beta,p,\theta)$$

for an integer value of y, $f_i(y;\alpha,\beta,p,\theta)$ can written as

$$\begin{aligned}
&f_i(y;\alpha,\beta,p,\theta)\\
&=\sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{(k+j)}\sum_{h=0}^{(\theta l)}\sum_{g=0}^{(\alpha h)}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}\binom{\theta l}{h}\binom{\alpha h}{g}(-1)^{j+l+h+g+1}p^{\beta g y}(1-p^{\beta g})\\
&=\sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{(k+j)}\sum_{h=0}^{(\theta l)}\sum_{g=0}^{(\alpha h)}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}\binom{\theta l}{h}\binom{\alpha h}{g}(-1)^{j+l+h+g+1}f_{DKw_ETE}(y;\beta g,\ p)
\end{aligned}$$

where, $f_{DKw\_ETE}(y;\beta g,\ p)$ is the pmf of the DKw_ETE distribution with parameters βg and p. Since $f_i(y;\alpha,\ \beta,p,\ \theta)$ is a linear combination of a finite number of DKw_ETE $(\alpha,\beta g,\ p)$, some properties of order statistics can be computed from the corresponding DKw_ETE distribution, such as moments. For example, the mean of the i^{th} order statistic is obtained by

$$\mu_{i:n}==\sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{(k+j)}\sum_{h=0}^{(\theta l)}\sum_{g=0}^{(\alpha h)}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}\binom{\theta l}{h}\binom{\alpha h}{g}(-1)^{j+l+h+g+1}\frac{p^{\beta g}}{(1-p^{\beta g})}$$

4 Maximum likelihood estimation

In this section we use the maximum likelihood estimation (MLE) to estimate parameters of the proposed distribution DKw_ETE.
Suppose $Y_1,Y_2,\dots,Y_n$ be a random sample of size n having the DKw_ETE distribution. The log-likelihood of the DKw_ETE distribution is

$$\ell=\sum_{j=1}^{n}\log\left\{[1-(1-\Psi)^{\alpha}]^{\theta}-[1-(1-\Lambda)^{\alpha}]^{\theta}\right\}\tag{13}$$

where $\Lambda=p^{\beta y_j}$ and $\psi=p^{\beta(y_j+1)}$
differentiating the log-likelihood in Eq.(13) partially with respect to the shape parameters α,β,θ and p to obtain the likelihood equations as

$$\begin{aligned}
\frac{\partial\ell}{\partial\alpha}=\theta\sum_{j=1}^{n}\Bigg\{&\frac{(1-\Psi)^{\alpha}[1-(1-\Psi)^{\alpha}]^{\theta-1}\log(1-\Psi)}{[1-(1-\Psi)^{\alpha}]^{\theta}-[1-(1-\Lambda)^{\alpha}]^{\theta}}\\
&-\frac{(1-\Lambda)^{\alpha}[1-(1-\Lambda)^{\alpha}]^{\theta-1}\log(1-\Lambda)}{[1-(1-\Psi)^{\alpha}]^{\theta}-[1-(1-\Lambda)^{\alpha}]^{\theta}}\Bigg\}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\ell}{\partial\beta}=\alpha\theta\sum_{j=1}^{n}\Bigg\{&\frac{y_j(1-\Psi)^{\alpha-1}[1-(1-\Psi)^{\alpha}]^{\theta-1}\log p}{[1-(1-\Psi)^{\alpha}]^{\theta}-[1-(1-\Lambda)^{\alpha}]^{\theta}}\\
&-\frac{(y_j+1)(1-\Lambda)^{\alpha-1}[1-(1-\Lambda)^{\alpha}]^{\theta-1}\Lambda\log p}{[1-(1-\Psi)^{\alpha}]^{\theta}-[1-(1-\Lambda)^{\alpha}]^{\theta}}\Bigg\}
\end{aligned}$$

$$\frac{\partial \ell}{\partial \theta} = \sum_{j=1}^{n} \left\{ \frac{[1-(1-\Psi)^{\alpha}]^{\theta} \log[1-(1-\Psi)^{\theta}]}{[1-(1-\Psi)^{\alpha}]^{\theta} - [1-(1-\Lambda)^{\alpha}]^{\theta}} - \frac{[1-(1-\Lambda)^{\alpha}]^{\theta} \log[1-(1-\Lambda)^{\alpha}]}{[1-(1-\Psi)^{\alpha}]^{\theta} - [1-(1-\Lambda)^{\alpha}]^{\theta}} \right\}$$

$$\frac{\partial \ell}{\partial p} = \frac{\alpha\beta\theta}{p} \sum_{j=1}^{n} \left\{ \frac{y_j \Lambda (1-\Lambda)^{\alpha-1} [1-(1-\Lambda)^{\alpha}]^{\theta-1}}{[1-(1-\Psi)^{\alpha}]^{\theta} - [1-(1-\Lambda)^{\alpha}]^{\theta}} - \frac{(y_j+1)\Psi(1-\Psi)^{\alpha-1}[1-(1-\Psi)^{\alpha}]^{\theta-1}}{[1-(1-\Psi)^{\alpha}]^{\theta} - [1-(1-\Lambda)^{\alpha}]^{\theta}} \right\}$$

Now let

$$\frac{\partial \ell}{\partial \alpha} = 0, \quad \frac{\partial \ell}{\partial \beta} = 0, \quad \frac{\partial \ell}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial \ell}{\partial p} = 0$$

and solve the resulting nonlinear system of equations, to get MLEs

$$\hat{\tau} = (\,\hat{\alpha},\, \hat{\beta}, \hat{\theta}\,,\, \hat{p})^{T} \quad \text{of} \quad \tau = (\,\alpha,\, \beta, \theta\,,\, p)^{T}$$

Since the MLE of the vector of unknown parameters $\tau = (\alpha, \beta, \lambda, \theta)^T$ cannot be derived in closed forms, it is therefore hard to derive the exact distribution of the MLEs. Moreover, the second partial derivatives can also be obtained.

$$\frac{\partial^2 \ell}{\partial \alpha^2}, \quad \frac{\partial^2 \ell}{\partial \beta^2}, \quad \frac{\partial^2 \ell}{\partial \theta^2} \qquad \text{and} \qquad \frac{\partial^2 \ell}{\partial p^2}$$

It is known that the asymptotic distribution of the MLEs $\hat{\tau}$ are given by [25].

$$(\hat{\tau} - \tau) \to N\left(0,\, I^{-1}(\tau)\right)$$

where $I^{-1}(\tau)$ is the inverse of the Fisher's information matrix of the unknown parameters τ and written as

$$I_{Y(\alpha,\beta,\theta,p)}(\tau) = \begin{bmatrix} -E\left(\frac{\partial^2\ell}{\partial\alpha^2}\right) & -E\left(\frac{\partial^2\ell}{\partial\alpha\partial\beta}\right) & -E\left(\frac{\partial^2\ell}{\partial\alpha\partial\theta}\right) & -E\left(\frac{\partial^2\ell}{\partial\alpha\partial p}\right) \\ -E\left(\frac{\partial^2\ell}{\partial\beta\partial\alpha}\right) & -E\left(\frac{\partial^2\ell}{\partial\beta^2}\right) & -E\left(\frac{\partial^2\ell}{\partial\beta\partial\theta}\right) & -E\left(\frac{\partial^2\ell}{\partial\beta\partial p}\right) \\ -E\left(\frac{\partial^2\ell}{\partial\theta\partial\alpha}\right) & -E\left(\frac{\partial^2\ell}{\partial\theta\partial\beta}\right) & -E\left(\frac{\partial^2\ell}{\partial\theta^2}\right) & -E\left(\frac{\partial^2\ell}{\partial\theta\partial p}\right) \\ -E\left(\frac{\partial^2\ell}{\partial p\partial\alpha}\right) & -E\left(\frac{\partial^2\ell}{\partial p\partial\beta}\right) & -E\left(\frac{\partial^2\ell}{\partial p\partial\theta}\right) & -E\left(\frac{\partial^2\ell}{\partial p^2}\right) \end{bmatrix}$$

Further, the Fisher's information matrix can be computed by using the approximation

$$I_Y(\hat{\tau}) = \begin{bmatrix} -E\left(\frac{\partial^2\ell}{\partial\alpha^2}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\alpha\partial\beta}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\alpha\partial\theta}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\alpha\partial p}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} \\ -E\left(\frac{\partial^2\ell}{\partial\beta\partial\alpha}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\beta^2}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\beta\partial\theta}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\beta\partial p}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} \\ -E\left(\frac{\partial^2\ell}{\partial\theta\partial\alpha}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\theta\partial\beta}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\theta^2}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial\theta\partial p}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} \\ -E\left(\frac{\partial^2\ell}{\partial p\partial\alpha}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial p\partial\beta}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial p\partial\theta}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} & -E\left(\frac{\partial^2\ell}{\partial p^2}\right)\Big|_{(\hat{\alpha},\hat{\beta},\hat{\theta},\hat{p})} \end{bmatrix}$$

5 Application

This section provides an application of the proposed distribution DKw_ETE. Therefore, to examine the goodness of fit for DKw_ETE and to assess its performance compared with some related distributions, DKw_ETE is fitted to two right-skewed (slightly and extremely) real lifetime count data sets. The first data includes the accidents that occurred to some women working on Shells for some weeks (AWS), this data set is reported by [26]. The second data is adduced by [27] represents the number of outbreaks of strikes in the UK coal mining industries in four successive week periods during 1948-59 (OSU). Table 2 demonstrates the summary statistics of the two data sets. Further, DKw_ETE is utilized to model the number of death cases of COVID-19 on some certain successive days during a given period of time. MLE is used to obtain the parameters estimate of DKw_ETE. The Statistical programming language R was used to do all of the calculations.

Table 2: Summary Statistics for AWS and OSU data sets.

Data set	n	**Min**	**Max**	**Mean**	**Variance**	**Skewness**
AWS data set	647	0	5	0.46522	0.69190	2.1212
OUS data set	156	0	4	0.99359	0.74189	0.80231

5.1 The first data set (AWS)

The AWS data set explains the accidents to 647 women working on Shells for 5 weeks. This data is extremely right-skewed (2.1212) and over-dispersed since the sample variance (0.69190) is greater than the respective sample mean (0.46522). The AWS data set was recently used by [7,23,28] in the application of generalization of the geometric distribution, discrete Extended Erlang-Truncated Exponential, and a discrete generalized exponential distribution of a second type, respectively. To compare the proposed DKw_ETE distribution with its related distributions involving discrete Extended Erlang-Truncated Exponential DEETE, Generalized Rayleigh distribution DGR, discrete Burr DBD, and discrete generalized exponentiated distribution of a second type DGE_2, we used -log-likelihood (-log(L)), and the χ^2 (chi square) statistic as criterions for comparison. Table 3 demonstrates the AWS observed values and the corresponding expected values computed using each distribution, while Table 4 explains the MLE estimators, -LogL, and the χ^2 statistics with the corresponding p-values.

Table 3: The observed and expected values of AWS data set.

Count	**Observed**	**DDKw_ETE**	**DEETE**	**DGR**	**DBD**	$\mathbf{DGE_2}$
0	447	447.53	446.87	448.12	447.44	446.92
1	132	133.20	133.64	122.60	142.01	133.63
2	42	44.53	44.44	53.00	35.97	44.43
3	21	14.99	15.03	18.15	12.65	15.02
4	3	5.05	5.17	4.41	5.77	5.16
>5	2	1.70	1.85	0.72	3.16	1.84
Total	647	647	647	647	647	647

Table 4: Parameters estimates, -Log L, χ^2 statistic and p-value for the selected distributions of the AWS data set.

Model	**MLEs**	**-Log L**	χ^2 **stat**	**p-value**
DKw_ETE	$\hat{\alpha}$=0.845, $\hat{\beta}$=1.6, $\hat{p}$=0.289, θ=0.547	592.160	3.393	0.6397
DEETE	$\hat{\alpha}$=0.898, $\hat{\beta}$=2.082, $\hat{p}$=0.594	592.183	3.445	0.6317
DGR	$\hat{\alpha}$=0.2196, $\hat{p}$=0.8123	592.544	6.172	0.2899
DBD	$\hat{\alpha}$=1.642, $\hat{\theta}$=0.1841	597.955	8.979	0.1099
DGE_2	$\hat{\alpha}$=0.898, $\hat{\theta}$=0.3379	592.183	3.448	0.6312

From the results in Table 4, it can be seen that the values of -logL for the proposed DKw_ETE is (592.16), which is the most minimum among the other related distributions (the smaller the better). On the other hand, this value together with the values of χ^2 statistic and their corresponding p-values demonstrate that the proposed DKw_ETE is the most appropriate model to fit the AWS data set. Moreover, all the studied distributions are appropriate for fitting the AWS data set.

5.2 The second data set (OSU)

The second data set OSU shows 156 outbreaks of strikes in the UK coal mining industries distributed among 4 successive week periods during 1948-59. This data is slightly right-skewed (0.80231), and it was utilized by [6,11,23] in the application of the discrete Gamma distribution, discrete generalized Rayleigh distribution, and discrete Extended Erlang-Truncated Exponential distribution, respectively. Table 5 shows the OSU data set observed values, and the corresponding expected values computed using the proposed DKw_ETE and the other related distributions, whilst Table 6 demonstrates the MLE estimators, -LogL, the χ^2 statistics, and its corresponding p-values.

Table 5: The observed and expected values of OSU data set.

Count	Observed	DKw_ETE	DEETE	DGR	DBD	DGE_2
0	46	46.34	46.08	47.50	47.89	46.08
1	76	74.73	75.75	69.32	80.87	75.75
2	24	27.19	26.11	31.82	18.28	26.11
3	9	6.42	6.50	6.67	6.07	6.50
>4	1	1.32	1.56	0.69	2.89	1.56
Total	156	156	156	156	156	156

Table 6: Parameters estimates, -Log L, χ^2 statistic and p-value for the selected distributions of the OSU data set.

Model	MLEs	-Log L	χ^2 stat	p-value
DKw_ETE	$\hat{\alpha}$=3.62, $\hat{\beta}$=0.301, $\hat{p}$=0.035, θ=1.66	187.498	1.489	0.8284
DEETE	$\hat{\alpha}$=4.799, $\hat{\beta}$=1.302, $\hat{p}$=0.318	187.534	1.335	0.8554
DGR	$\hat{\alpha}$=0.9415, $\hat{p}$=0.7172	188.329	3.5623	0.4685
DBD	$\hat{\alpha}$=4.6543, $\hat{\theta}$=0.5941	192.210	4.816	0.3067
DGE_2	$\hat{\alpha}$=4.7994, $\hat{\theta}$=0.2247	187.534	1.3347	0.8555

Based on the outputs in Table 6, by looking at the values of χ^2 statistic and their corresponding p-values (¿0.05) we can conclude that all the studied distributions are appropriate to fit the OSU data set, while the minimum value of -logL (187.498) for the proposed DKw_ETE explains that DKw_ETE is the most convenient model to fit this data sets.

5.3 Application of DKw_ETE in COVID-19 number of death cases

The DKw_ETE is utilized to model the number of death cases of COVID-19 in Qatar, UAE, and Bahrain that have occurred in a successive days during a given period. This application was adopted by [23] in the study of DEETE distribution (see: Table 8). Table 7 shows the summary statistics of the three data sets. Based on Table 7 we observe that the three data sets are right-skewed and overdispersed.

The discrete random variable X which represents number of death cases of COVID-19 occurred in a successive days having a DKw_ETE distribution with the pmf written as

$$f_{DKw\_ETE}(x;\alpha,\beta,p,\theta) = \left[1-\left(1-p^{\beta x}\right)^{\alpha}\right]^{\theta} - \left[1-\left(1-p^{\beta(x+1)}\right)^{\alpha}\right]^{\theta} \tag{14}$$

where $x = 0,1,2,\dots, \alpha,\ \beta,\ \theta > 0$ and $0 < p < 1$.

Table 7: Summary Statistics for death cases of COVID-19 data sets.

Data set	n	Min	Max	Mean	Variance	Skewness
Qatar data set	277	0	6	0.85198	1.53236	1.5914
UAE data set	306	0	11	1.82026	4.19382	1.7736
Bahrain data set	283	0	7	1.20495	2.14934	1.3944

Table 9 gives the MLE estimators, -LogL, and the χ^2 statistics with the corresponding p-values for Qatar, UAE, and Bahrain data sets. From the results in Table 9, by looking at -LogL values and p-values, it is clear that the proposed DKw_ETE distribution is more appropriate for fitting the death cases of COVID-19 data set than the DEETE distribution in three countries.

Table 8: Number of death cases of COVID-19 and the corresponding number of days in which they occurred until 30 Nov 2021 in Qatar, UAE and Bahrain

Qatar data set		**UAE data set**		**Bahrain data set**	
# deaths	**# days**	**# deaths**	**# days**	**# deaths**	**# days**
0	156	0	93	0	124
1	58	1	70	1	67
2	31	2	67	2	41
3	18	3	31	3	31
4	9	4	17	4	7
5	4	5	8	5	9
6	0	6	7	6	2
7	1	7	4	>7	2
8	0	8	3		
9	0	9	3		
10	0	10	0		
>11	1	>11	3		
Total	278		306		283

Table 9: Parameters estimates, -Log L, χ^2 statistic and p-value for DKw_ETE and DEETE distributions of the COVID-19 death cases data sets in the three countries.

		Qatar COVID-19 death cases data set					
		MLEs					
Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{p}$	θ	**-Log L**	χ^2 **stat**	**p-value**
DKw_ETE	0.5065	1.3723	0.0367	0.1588	353.29	1.7692	0.9396
DEETE	0.8659	2.0393	0.7042	-	353.50	2.1638	0.9040
		UAE COVID-19 death cases data set					
DKw_ETE	1.3301	0.5939	0.4132	0.9861			0.552
DEETE	1.3283	2.5176	0.8141	-	557.90	9.6952	0.558
		Bahrain COVID-19 death cases data set					
DKw_ETE	1.1210	0.1704	0.1843	2.3492	429.14	8.4242	0.297
DEETE	1.1456	1.4931	0.6456	-	429.30	8.5281	0.288

6 Conclusions

In this paper, the discrete Kumaraswamy Erlang-truncated exponential distribution (DKw_ETE) is introduced by discretizing the Kumaraswamy Erlang-truncated exponential distribution. Some statistical properties of DKw_ETE are derived, and the Maximum likelihood method is used to estimate the model's parameters. The pmf of DKw_ETE is unimodal right-skewed, while the hazard function is increasing. The proposed DKw_ETE is applied to two real count data sets as well as used to model some data sets regarding the death cases of COVID-19. The main findings have been concluded from the application results are; DKw_ETE is the most convenient model to fit both data sets than other related distributions; further, DKw_ETE is more appropriate for fitting the death cases of COVID-19 than DEETE distribution. We recommend the proposed distribution for modelling the right-skewed real-life count data sets adequacy because the four parameters make DKw_ETE more flexible to suit most types of count data sets.

References

[1] C. Bracquemond , O. A Gaudoin. survey on discrete lifetime distribution. *International Journal of Reliability, Quality and Safety Engineering*, **10(01)**, 69-98 (2003).
[2] S. Chakraborty. Generating discrete analogues of continuous probability distributions-A survey of methods and constructions. *Journal of Statistical Distributions and Applications*, **2(1)**, 1-30 (2015).
[3] D. Roy. Reliability measures in the discrete bivariate set-up and related characterization results for a bivariate geometric distribution. *Journal of Multivariate Analysis*, **46(2)**, 362-373 (1993).

[4] D. Roy. The discrete normal distribution. *Communications in Statistics - Theory and Methods*, **32(10)**, 1871-1883 (2003).
[5] H. Krishna and P. S. Pundir. Discrete Burr and discrete Pareto distribution. *Statistical Methodology*, **6(1)**, 177–188 (2009).
[6] S. Chakraborty and D. Chakravarty. Discrete Gamma Distribution: Properties and parameter estimations.*Communications in statistics-Theory and Methods*, **41**, 3301-3324 (2012).
[7] V. Nekoukhou, M. H. Alamatsaz and H. Bidram. Discrete generalized exponential distribution of a second type, Statistics. *A Journal of Theoretical and Applied Statistics*, **47(4)**, 876-887 (2013).
[8] V. Nekoukhou and H. Bidram. The exponentiated discrete Weibull distribution. *Statistics and Operations Research Transactions*, **39(1)**, 127-146 (2015).
[9] S. Chakraborty and D. Chakravarty. A new discrete probability distribution with integer support on $(-\infty,\infty)$. *Communications in Statistics-Theory and Methods*, **45(52)**, 492-505 (2016).
[10] T. Hussain, M. Aslam and M. Ahmad. A Two Parameter Discrete Lindley Distribution. *Revista Colombiana de Estadística*, **39(1)**, 45-61 (2016).
[11] M. H. Alamatsaz, S. Dey, T. Dey and S. S. Harandi. Discrete generalized Rayleigh distribution. *Pakistan journal of statistics*, **32(1)**, 1-20(2016).
[12] P. Khongthip, M. Patummasut and W. Bodhisuwan. The discrete weighted exponential distribution and its applications. *Songklanakarin Journal of Science & Technology*, **40(5)**, 1105-1114 (2018).
[13] K. Jayakumar and M. G. Babu. Discrete Additive Weibull Geometric Distribution. *Journal of Statistical Theory and Applications*, **18(1)**, 33-45 (2019).
[14] A. R. El-Alosey. Random Sum of New Type of Mixtures of Distributions. *International Journal of Statistics and Systems*, **2(1)**, 49-57 (2007).
[15] S. Nasiru, A. Luguterah and M. M. Iddrisu. Generalized Erlang-truncated Exponential Distribution. *Advances and Applications in Statistics*, **48(4)**, 273-301 (2016).
[16] S. Nasiru, M. Atem and K. Nantomah. Poisson exponentiated erlang-truncated exponential distribution. *Journal of Statistics Applications and Probability*, **7**, 245-261 (2018).
[17] I. E. Okorie, A. C. Akpanta and J. Ohakwe. Marshall-Olkin generalized Erlang-truncated exponential distribution: Properties and applications. *Cogent Mathematics*, **4(1)**, 1285093 (2017).
[18] I. E. Okorie, A. C. Akpanta and J. Ohakwe. Transmuted Erlang-truncated exponential distribution. *Economic Quality Control*, **31(2)**, 71-84 (2016).
[19] I. E. Okoriea, ; A. C. Akpantab, J. Ohakwec and D. C. Chikezieb. The Extended Erlang-Truncated Exponential distribution: properties and application to rainfall data. **Heliyon**, *3* (2017).
[20] H. Jimoh, A. O. Oluyede, D. Manduku and B. Makubate. The Gamma Log-Logistic Erlang Truncated Exponential Distribution with Applications. *Afrika Statistika*, **14(4)**, 2141-2164 (2019).
[21] M. Shrahili, I. Elbatal, I. Muhammad and M. Muhammad. Properties and applications of beta Erlang-truncated exponential distribution. *Journal of Mathematics and Computer Science*, **22**, 16-37 (2021).
[22] A. R. El-Alosey. Discrete Erlang-Truncated Exponential Distribution. *International Journal of Statistics and Applied Mathematics*, **6(1)**, 230-236 (2021).
[23] A. R. El-Alosey and H. Eledum. Discrete Erlang-Truncated Exponential Distribution. *Applied Mathematics & Information Sciences*, **16(1)**, 1-12 (2022).
[24] N.Y. Ma. Complete multinomial expansions. *Applied Mathematics and Computation*, **124**, 365-370 (2001).
[25] J.F. Lawless. *Statistical Models and Methods for Lifetime Data*; John Wiley and Sons: New York, USA, 2002.
[26] P.C. Consul and G. C. Jain. A Generalization of the Poisson Distribution. *Technometrics*, **15**, 791-799 (1973).
[27] M. S. Ridout and P. Besbeas. An empirical model for underdispersed count data. *Statistical Modelling*, **4(1)**, 77-89 (2004).
[28] E. Gomez-Deniz. Another generalization of the geometric distribution. *Test*, **19(2)**, 399-415 (2010).