ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2018-61093

A Discrete Quasi Akash Distribution with Applications

Kesikli Quasi Akash Dağılımı ve Uygulamaları

Berhane ABEBE,[a]
Rama SHANKER[a]

[a]Department of Statistics
College of Science,
Eritrea Institute of Technology,
Asmara, Eritrea

Received: 15.04.2018
Received in revised form: 24.09.2018
Accepted: 15.11.2018
Available Online: 07.12.2018

Correspondence:
Berhane ABEBE
Department of Statistics
College of Science,
Eritrea Institute of Technology,
Asmara, Eritrea
brhaneaah@gmail.com

ABSTRACT A two-parameter discrete quasi Akash Distribution (DQAD) which includes one parameter discrete Akash distribution as a particular case has been obtained using infinite series method of discretization from quasi Akash distribution. Its moments about origin and central moments have been obtained. Behaviors of statistical constants including coefficients of variation, skewness, kurtosis and index of dispersion have been discussed. Maximum likelihood estimation has been discussed for estimating its parameters. Finally, applications of the proposed distribution have been explained using two real datasets form thunderstorm events and the goodness of fit has been compared to other discrete distributions.

Keywords: Quasi Akash distribution, discretization, moment generating function, moments, maximum likelihood estimation, goodness of fit.

ÖZET Özel durum olarak tek parametreli kesikli Akash dağılımını kapsayan iki parametreli kesikli quasi Akash dağılımı (DQAD), quasi Akash dağılımının kesikli hale getirilmesinde sonsuz seriler yöntemi kullanılarak elde edilmiştir. Dağılıma ait merkezi momentler elde edilmiştir. Yayılım indeksi, çarpıklık, basıklık ve değişim katsayını kapsayan istatistiksel sabitlerin özellikleri tartışılmıştır. Dağılım parametrelerini tahmin etmek için en çok olabilirlik tahmini tartışılmıştır. Son olarak, önerilen dağılımın uygulamaları fırtına olaylarını içeren iki gerçek veri seti kullanılarak açıklanmış ve uyum iyiliği diğer kesikli dağılımlarla karşılaştırılmıştır.

Anahtar Kelimeler: Quasi Akash dağılımı; kesikli hale getirme; moment üreten fonksiyon; momentler; en çok olabilirlik tahmini; uyum iyiliği

In the last few decades, many efforts have been done to derive a discrete analogue of continuous distribution. The main reasons for discretizing continuous distributions are (i) to derive alternative discrete distributions to the classical discrete distributions commonly used in the analysis of count data, failure data and reliability data (ii) the discrete analogue of continuous distributions avoids the use of continuous distributions in the case of strictly discrete data.

In many practical life situations, the observed values are measured on discrete scales or even if they are measured on continuous scales, they are measured to two or three decimal places and do not contain all points in the interval. For instance in case of lifetime data (waiting time or survival time), even if the measurements are taken on a continuous scale, the observations may be recorded in a way that makes a discrete distribution more appropriate model. Again, in survival analysis, it is most common to use continuous distributions to model discrete data. The discretization of a continuous distribution acts as a subterfuge to avoid the use of continuous distribution to model discrete time data. It has been pointed by Lai (2013) that discretization of a continuous lifetime model is an interesting and intuitively appealing approach to derive a discrete lifetime model corresponding to the continuous one.[1] It has been observed that in real world the original variables may be continuous in nature but discrete by observation and , therefore, it is reasonable and convenient to model the situation by an appropriate discrete distribution generated from the underlying continuous distribution preserving one or more important characteristics including probability density function (pdf), moment generating function (mgf), moments, hazard rate function, mean residual life function etc., of the continuous distribution.

A number of methods are available in Statistics literature to derive a discrete analogue of continuous distribution. One of the first proposed discretization methods is based on the definition of pmf that depends on an infinite series. The method of discretization by an infinite series has been proposed by Good (1953) who has proposed the discrete Good distribution to model the population frequencies of species and discussed the estimation of parameters.[2] A random variable Y is said to have a discrete Good distribution if its pmf can be expressed as

$$P(Y=y)=\frac{\alpha^{y}y^{\beta}}{\sum_{j=1}^{\infty}\alpha^{j}j^{\beta}}\ ;y=0,1,2,....,\beta\in R \text{ and } \alpha\in(0,1) \tag{1.1}$$

The method of infinite series is characterized by the following definition

Definition 1.1: Let X be a continuous random variable having pdf $f_X(x)$ and parameter with support on R. Then the corresponding discrete random variable Y has pmf given by

$$P(Y=y)=P(y;\theta)=\frac{f_X(y;\theta)}{\sum_{j=-\infty}^{\infty}f_X(j;\theta)}\ ;y\in Z,\theta>0 \tag{1.2}$$

where θ may be the vector of parameters indexing the distribution of X.

This method of discretizing a continuous distribution has been studied by several researchers including Kulasekara and Tonkyn (1992), Doray and Luong (1997), Sato *et al* (1999), Nekoukhou *et al* (2012), are some among others, who proposed a version of the method when the continuous random variable of interest is defined on R_+ [3,4,5,6]. Thus, if the continuous random variable X is defined on R_+, the pmf of corresponding discrete random variable Y can be defined as

$$P(Y=y)=P(y;\theta)=\frac{f_X(y;\theta)}{\sum_{j=0}^{\infty}f_X(j;\theta)}\ ;y\in Z_+ \tag{1.3}$$

Using (1.3), Berhane and Shanker (2018) introduced a discrete Akash distribution (DAD) defined by its pmf[7]

$$P_1(y;\theta)=\frac{\left(e^{\theta}-1\right)^{3}}{e^{\theta}\left(e^{2\theta}-e^{\theta}+2\right)}\left(1+y^{2}\right)e^{-\theta y};\ y=0,1,2,...,\theta>0 \tag{1.4}$$

Various statistical properties, estimation of parameter and applications of the DAD for count data have been discussed by Berhane and Shanker (2018).[7] Note that DAD is a discrete analogue of a continuous Akash distribution proposed by Shanker (2015) having pdf and cdf[8]

$$f_1(x;\theta)=\frac{\theta^2}{\theta+2}\left(1+x^2\right)e^{-\theta x}\;;\quad x>0,\quad \theta>0 \tag{1.5}$$

$$F_1(x;\theta)=1-\left[1+\frac{\theta x(\theta x+2)}{\theta^2+2}\right]e^{-\theta x};\; x>0,\theta>0 \tag{1.6}$$

Shanker (2017) obtained Poisson–Akash distribution (PAD), a Poisson mixture of Akash distribution, having pmf[9]

$$P_2(x;\theta)=\frac{\theta^3}{\theta^2+2}.\frac{x^2+3x+\left(\theta^2+2\theta+3\right)}{(\theta+1)^{x+3}};\; x=0,1,2,...\theta>0 \tag{1.7}$$

Important statistical and mathematical properties, estimation of parameter using the method of moments and the method of maximum likelihood along with applications of PAD to model count data have been studied by Shanker (2017).[9]

Shanker *et al* (2018) introduced a quasi Poisson-Akash distribution (QPAD) defined by its pmf[10]

$$P_3(x;\theta)=\frac{\theta^2}{\alpha\theta+2}\frac{\left(x^2+3x\right)\theta+\left(\alpha\theta^2+2\alpha\theta+2\theta+\alpha\right)}{(\theta+1)^{x+3}};\; x=0,1,2,...\theta>0,\alpha\theta+2>0 \tag{1.8}$$

It should be noted that QPAD is a Poisson mixture of quasi Akash distribution (QAD) proposed by Shanker (2016) and defined by its pdf and cdf[11]

$$f_2(x;\theta)=\frac{\theta^2}{\alpha\theta+2}\left(\alpha+\theta x^2\right)e^{-\theta x}\;;\quad x>0,\quad \theta>0\quad \alpha\theta+2>0 \tag{1.9}$$

$$F_2(x;\theta,\alpha)=1-\left[1+\frac{\theta x(\theta x+2)}{\alpha\theta+2}\right]e^{-\theta x};\; x>0,\theta>0,\alpha\theta+2>0 \tag{1.10}$$

Shanker (2016) has discussed its various mathematical and statistical properties including its shapes for varying values of parameters, moments based measures, hazard rate function, mean residual life function, stochastic ordering, mean deviations, order statistics, Renyi entropy measure, Bonferroni and Lorenz curves and stress-strength reliability along with estimation of parameter and applications of the distribution for modeling lifetime data from biomedical sciences and engineering.[11]

A second common method of discretization of a continuous distribution is based on the survival function of the original continuous distribution and was proposed by Nakagawa and Osaki (1975) and has the interesting feature of preserving the original survival function on its integer part for the generated pmf.[12] According to kemp (2004), a discrete random variable Y corresponding to a continuous random variable X can be defined as follows:[13]

Definition 1.2: Let X be a continuous random variable having survival function $S_X(x)=1-F_X(x)=P(X\geq x)$. Then, the discrete random variable $Y=[X]$ has pmf given by

$$P(Y=y)=P(y;\theta)=S_X(y;\theta)-S_X(y+1;\theta);\ y=0,1,2,... \quad , \tag{1.11}$$

where $Y=[X]=$ largest integer less than or equal to X.

Note that the resulting pmf will be in closed form if the original survival function has closed form. Using the method of survival function of distribution, Nakagawa and Osaki (1975) proposed a discrete Weibull distribution, a discrete analogue of Weibull distribution, and studied its properties, estimation of parameters and applications.[12]

In the present paper, a discrete quasi Akash distribution (DQAD), a discrete analogue of continuous Quasi Akash distribution introduced by Shanker (2016) has been investigated using a discretization method based on an infinite series.[11] Some distributional properties including moment generating function, moments and moments based measures, behaviors of coefficient of variation, skewness, kurtosis and index of dispersion of DQAD have been discussed. The estimation of parameters of the distribution has been discussed using the method of maximum likelihood. The goodness of fit of DQAD has been carried out using real datasets from thunderstorm events and the fit has been compared with one parameter discrete distributions including discrete Akash distribution and Poisson-Lindley distribution (PLD) and two-parameter quasi Poisson-Akash distribution (QPAD).

A DISCRETE QUASI AKASH DISTRIBUTION

Using equation (1.3), the pmf of the discrete random variable Y corresponding to a continuous random variable X following QAD (1.5) can be obtained as

$$P_4(y;\theta)=\frac{\left(e^{\theta}-1\right)^3}{e^{\theta}\left(\alpha\left(e^{\theta}-1\right)^2+\theta\left(e^{\theta}+1\right)\right)}\left(\alpha+\theta y^2\right)e^{-\theta y};\ y=0,1,2,...,\theta>0,\ \alpha\theta+2>0 \tag{2.1}$$

We would call this distribution, a discrete quasi Akash Distribution (DQAD). It can be easily verified that DAD (1.4) is a particular case of DQAD for $\alpha=\theta$.

The corresponding cdf and the survival function of DQAD can be obtained as

$$F_3(y;\theta,\alpha)=1-\frac{\left[\alpha\left(e^{\theta}-1\right)^2+\theta\left\{(y+1)\left(e^{\theta}-1\right)+1\right\}+\theta e^{\theta}\right]}{\alpha\left(e^{\theta}-1\right)^2+\theta\left(e^{\theta}+1\right)}e^{-\theta(y+1)};\ y=0,1,2,...,\theta>0,\ \alpha\theta+2>0\,, \tag{2.2}$$

$$S(y;\theta,\alpha)=\frac{\left[\alpha\left(e^{\theta}-1\right)^2+\theta\left\{(y+1)\left(e^{\theta}-1\right)+1\right\}+\theta e^{\theta}\right]}{\alpha\left(e^{\theta}-1\right)^2+\theta\left(e^{\theta}+1\right)}e^{-\theta(y+1)};\ y=0,1,2,...,\theta>0,\ \alpha\theta+2>0 \tag{2.3}$$

The behaviors of the pmf and cdf of DQAD for varying values of its parameter θ and α have been shown graphically in figures 1 and 2, respectively.

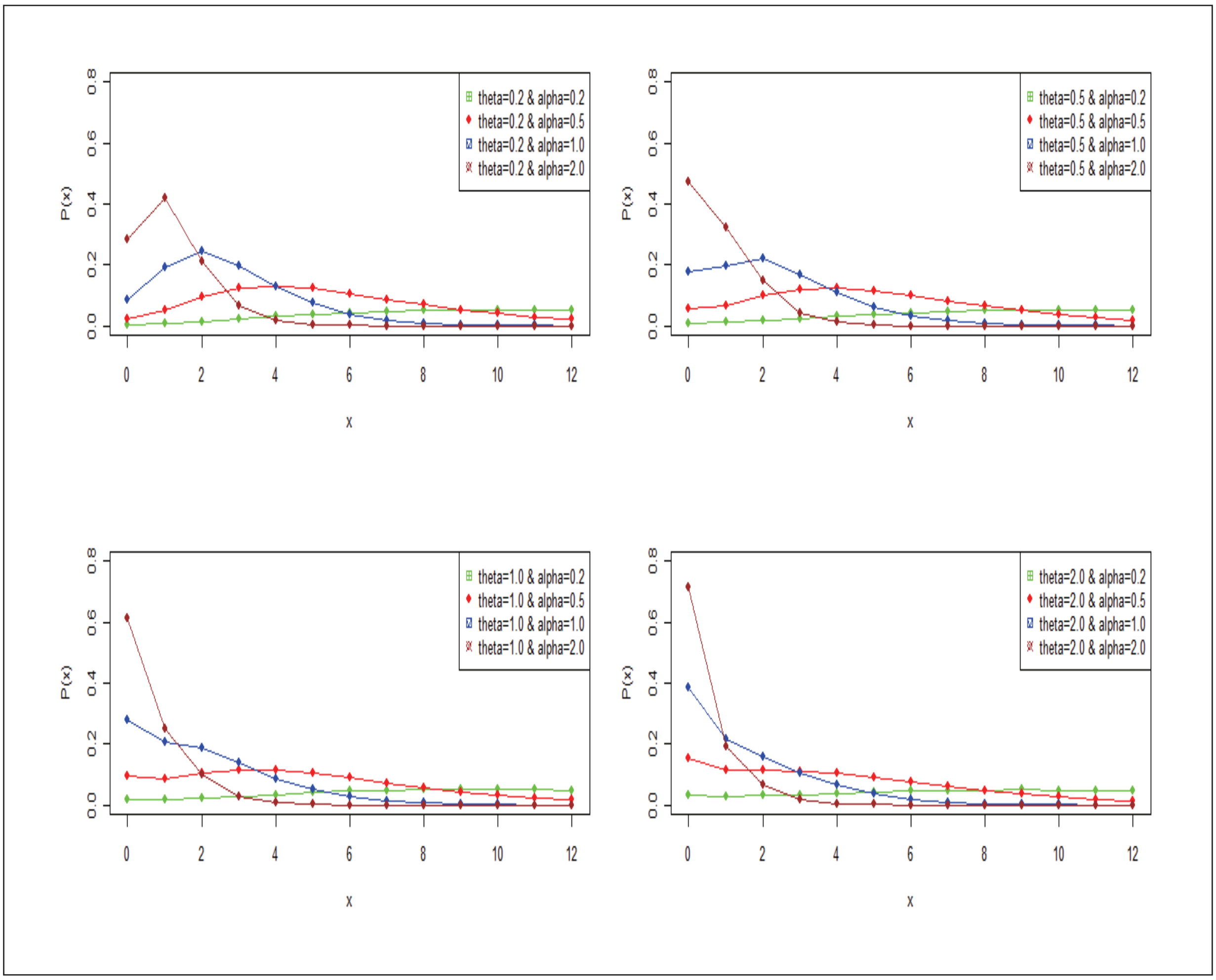

FIGURE 1: Behaviour of pmf of DQAD for varying values of the parameters θ and α

Since $\dfrac{P_4(y+1;\theta,\alpha)}{P_4(y;\theta,\alpha)} = \left[1+\dfrac{2y+1}{1+y^2}\right]e^{-\theta}$ is a decreasing function for $y \geq 1$, $P_4(y;\theta,\alpha)$ is log-concave and therefore, the DQAD has an increasing hazard rate and unimodality. The interrelationship between log-concavity, unimodality and increasing hazard rate of discrete distributions are available in Grandell (1997).[14] Further, $\left[P_4(y;\theta,\alpha)\right]^2 \geq P_4(y-1;\theta,\alpha)\cdot P_4(y+1;\theta,\alpha)$ for $y \geq 2$, which implies unimodality, by theorem 3 of Keilson and Gerber (1971).[15]

MOMENTS AND ASSOCIATED MEASURES

The probability generating function (pgf) and the moment generating function (mgf) of DQAD can be obtained as

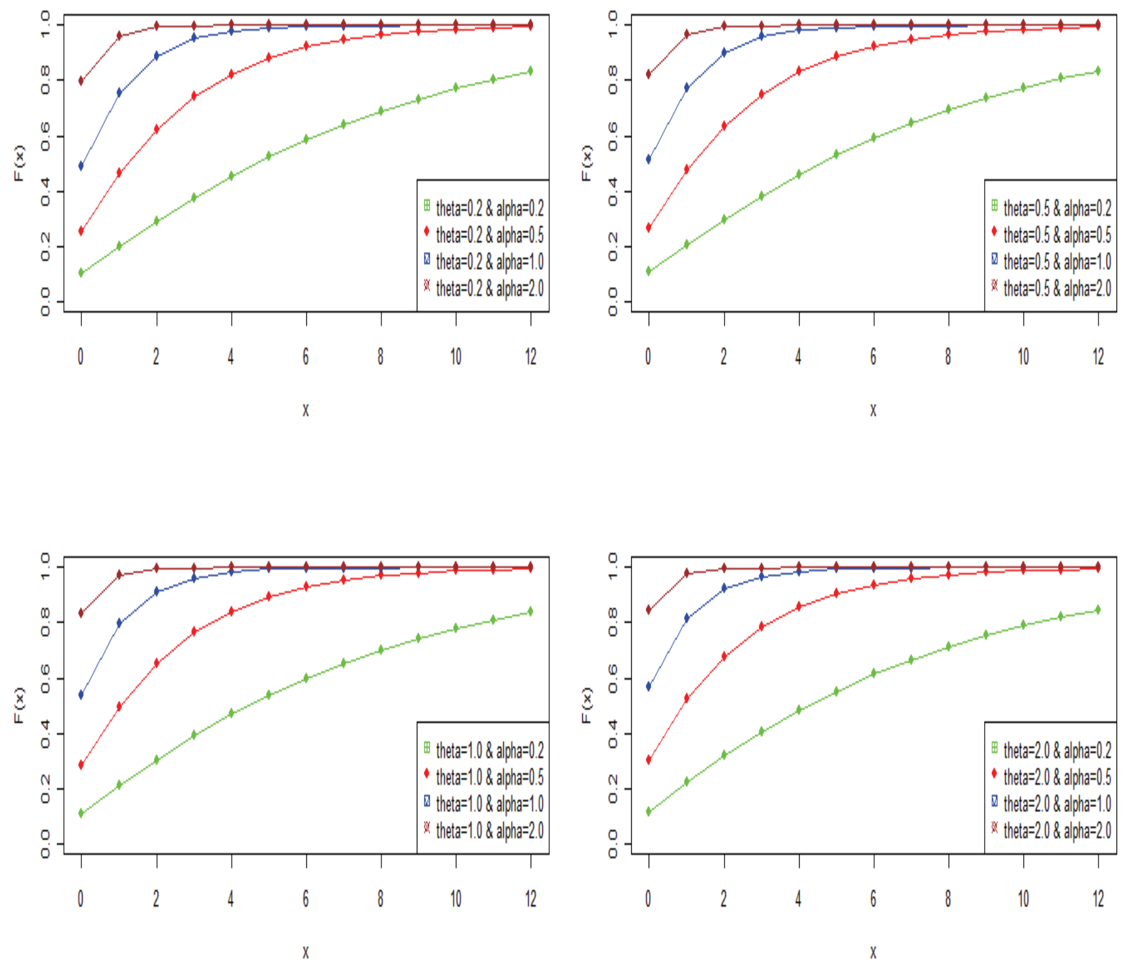

FIGURE 2: Behaviour of cdf of DQAD for varying values of the parameters θ and α.

$$G(t)=\frac{\left(e^{\theta}-1\right)^{3}}{\alpha\left(e^{\theta}-1\right)^{2}+\theta\left(e^{\theta}+1\right)}\left[\frac{\alpha\left(e^{\theta}-t\right)^{2}+t\theta\left(e^{\theta}+t\right)}{\left(e^{\theta}-t\right)^{3}}\right],\text{for } t\neq e^{\theta}, \tag{3.1}$$

and

$$M(t)=\frac{\left(e^{\theta}-1\right)^{3}}{e^{\theta}\left(\alpha\left(e^{\theta}-1\right)^{2}+\theta\left(e^{\theta}+1\right)\right)}\left[\frac{\alpha\left(1-e^{-(\theta-t)}\right)^{2}+\theta e^{-(\theta-t)}\left(1+e^{-(\theta-t)}\right)}{\left(1-e^{-(\theta-t)}\right)^{3}}\right],\text{for } t\neq\theta \tag{3.2}$$

It can be easily verified that the function in (3.2) is infinitely differentiable with respect to *t*, since it involves exponential terms of its argument. This means that one can derive all moments about origin μ_r' , $r\geq 1$ of DQAD from its mgf.

The first four moments about origin of DQAD can thus be obtained as

$$\mu_1' = \frac{(\theta+\alpha)e^{2\theta} + (4\theta-2\alpha)e^{\theta} + \theta + \alpha}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}\left(e^{\theta}-1\right)}$$

$$\mu_2' = \frac{(\theta+\alpha)e^{3\theta} + (11\theta-\alpha)e^{2\theta} + (11\theta-\alpha)e^{\theta} + \theta + \alpha}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}\left(e^{\theta}-1\right)^2}$$

$$\mu_3' = \frac{(\theta+\alpha)e^{4\theta} + (26\theta+2\alpha)e^{3\theta} + (66\theta-6\alpha)e^{2\theta} + (26\theta+2\alpha)e^{\theta} + \theta + \alpha}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}\left(e^{\theta}-1\right)^3}$$

$$\mu_4' = \frac{(\theta+\alpha)e^{5\theta} + (9\alpha+57\theta)e^{4\theta} + (302\theta-10\alpha)e^{3\theta} + (302\theta-10\alpha)e^{2\theta} + (9\alpha+57\theta)e^{\theta} + \theta + \alpha}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}\left(e^{\theta}-1\right)^4}$$

Using the relationship $\mu_r = E\left(Y-\mu_1'\right)^r = \sum_{k=0}^{r}\binom{r}{k}\mu_k'\left(-\mu_1'\right)^{r-k}$ between central moments and moments about origin, the central moments of DQAD are obtained as

$$\mu_2 = \frac{e^{\theta}\left\{\begin{array}{l}\left(\alpha^2+\theta\alpha\right)e^{4\theta} + \left(8\alpha\theta-4\alpha^2\right)e^{3\theta} + \left(6\alpha^2-14\theta\alpha+4\theta^2\right)e^{2\theta} + \left(4\theta^2-4\alpha^2\right)e^{\theta} \\ +\left(\alpha^2+4\theta^2+5\alpha\theta\right)\end{array}\right\}}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}^2\left(e^{\theta}-1\right)^2}$$

$$\mu_3 = \frac{\left(\begin{array}{l}\left(\alpha^3+\alpha\theta^2\right)e^{8\theta} + \left(18\theta\alpha^2-\alpha\theta^2-5\alpha^3\right)e^{7\theta} + \left(9\alpha^3-51\theta\alpha^2+12\alpha\theta^2\right)e^{6\theta} \\ +\left(4\theta^3+10\theta\alpha^2-5\alpha\theta^2-5\alpha^3\right)e^{5\theta} + \left(75\theta\alpha^2-5\alpha^3+8\theta^3+16\alpha\theta^2\right)e^{4\theta} \\ +\left(9\alpha^3+24\theta^3-51\alpha\theta^2-66\alpha\theta^2-66\alpha\theta^2\right)e^{3\theta} + \left(7\theta\alpha^2+7\theta\alpha^2+20\alpha\theta^2\right)e^{2\theta} \\ +\left(6\theta\alpha^2+9\alpha\theta^2+\alpha^3+4\theta^3\right)e^{\theta}\end{array}\right)}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}^3\left(e^{\theta}-1\right)^3}$$

$$\mu_4 = \frac{\left(\begin{array}{l}\left(\alpha^4+\alpha^3\theta\right)e^{11\theta} + \left(46\alpha\theta^3-\alpha^2\theta^2-\alpha^4\right)e^{10\theta} + \left(\alpha\theta^3-27\alpha^4-85\alpha^3\theta+63\alpha^2\theta^2\right)e^{9\theta} \\ +\left(132\alpha^4-400\alpha^3\theta+132\alpha^2\theta^2+40\alpha\theta^3\right)e^{8\theta} + \left(1490\alpha^3\theta-628\alpha^2\theta^2+232\alpha\theta^3+4\theta^4-294\alpha^4\right)e^{7\theta} \\ +\left(378\alpha^4+64\theta^4-1828\theta\alpha^3+234\theta^2\alpha^2+56\alpha\theta^3\right)e^{6\theta} + \left(734\theta\alpha^3+570\alpha^2\theta^2-486\alpha\theta^3-294\alpha^4+188\theta^4\right)e^{5\theta} \\ +\left(320\theta\alpha^3-124\alpha^2\theta^2-104\alpha\theta^3+132\alpha^4+208\theta^4\right)e^{4\theta} + \left(188\theta^4-27\alpha^4-355\alpha^3\theta-468\theta^2\alpha^2+48\theta^3\alpha\right)e^{3\theta} \\ +\left(200\theta^3\alpha+70\theta\alpha^3+207\theta^2\alpha^2-\alpha^4+64\theta^4\right)e^{2\theta} + \left(\alpha^4+4\theta^4+7\alpha^3\theta+15\alpha^2\theta^2+13\alpha\theta^3\right)e^{\theta}\end{array}\right)}{\left\{\alpha\left(e^{\theta}-1\right)^2 + \theta\left(e^{\theta}+1\right)\right\}^4\left(e^{\theta}-1\right)^4}$$

The coefficient of variation (C.V), coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis (β_2) and index of dispersion (γ) of DQAD can be expressed as

$$C.V=\frac{\sigma}{\mu_1'}=\frac{\sqrt{e^{\theta}\left\{\begin{array}{l}\left(\alpha^2+\theta\alpha\right)e^{4\theta}+\left(8\alpha\theta-4\alpha^2\right)e^{3\theta}+\left(6\alpha^2-14\theta\alpha+4\theta^2\right)e^{2\theta}+\left(4\theta^2-4\alpha^2\right)e^{\theta}\\+\left(\alpha^2+4\theta^2+5\alpha\theta\right)\end{array}\right\}}}{(\alpha+\theta)e^{2\theta}+(4\theta-2\alpha)e^{\theta}+\theta+\alpha}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{(\mu_2)^{3/2}}=\frac{\left(\begin{array}{l}\left(\alpha^3+\alpha\theta^2\right)e^{8\theta}+\left(18\theta\alpha^2-\alpha\theta^2-5\alpha^3\right)e^{7\theta}+\left(9\alpha^3-51\theta\alpha^2+12\alpha\theta^2\right)e^{6\theta}\\+\left(4\theta^3+10\theta\alpha^2-5\alpha\theta^2-5\alpha^3\right)e^{5\theta}+\left(75\theta\alpha^2-5\alpha^3+8\theta^3+16\alpha\theta^2\right)e^{4\theta}\\+\left(9\alpha^3+24\theta^3-51\alpha\theta^2-66\alpha\theta^2-66\alpha\theta^2\right)e^{3\theta}+\left(7\theta\alpha^2+7\theta\alpha^2+20\alpha\theta^2\right)e^{2\theta}\\+\left(6\theta\alpha^2+9\alpha\theta^2+\alpha^3+4\theta^3\right)e^{\theta}\end{array}\right)}{\left(e^{\theta}\left\{\begin{array}{l}\left(\alpha^2+\theta\alpha\right)e^{4\theta}+\left(8\alpha\theta-4\alpha^2\right)e^{3\theta}+\left(6\alpha^2-14\theta\alpha+4\theta^2\right)e^{2\theta}\\+\left(4\theta^2-4\alpha^2\right)e^{\theta}+\left(\alpha^2+4\theta^2+5\alpha\theta\right)\end{array}\right\}\right)^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^{2}}=\frac{\left(\begin{array}{l}\left(\alpha^4+\alpha^3\theta\right)e^{11\theta}+\left(46\alpha\theta^3-\alpha^2\theta^2-\alpha^4\right)e^{10\theta}+\left(\alpha\theta^3-27\alpha^4-85\alpha^3\theta+63\alpha^2\theta^2\right)e^{9\theta}\\+\left(132\alpha^4-400\alpha^3\theta+132\alpha^2\theta^2+40\alpha\theta^3\right)e^{8\theta}+\left(1490\alpha^3\theta-628\alpha^2\theta^2+232\alpha\theta^3+4\theta^4-294\alpha^4\right)e^{7\theta}\\+\left(378\alpha^4+64\theta^4-1828\theta\alpha^3+234\theta^2\alpha^2+56\alpha\theta^3\right)e^{6\theta}+\left(734\theta\alpha^3+570\alpha^2\theta^2-486\alpha\theta^3-294\alpha^4+188\theta^4\right)e^{5\theta}\\+\left(320\theta\alpha^3-124\alpha^2\theta^2-104\alpha\theta^3+132\alpha^4+208\theta^4\right)e^{4\theta}+\left(188\theta^4-27\alpha^4-355\alpha^3\theta-468\theta^2\alpha^2+48\theta^3\alpha\right)e^{3\theta}\\+\left(200\theta^3\alpha+70\theta\alpha^3+207\theta^2\alpha^2-\alpha^4+64\theta^4\right)e^{2\theta}+\left(\alpha^4+4\theta^4+7\alpha^3\theta+15\alpha^2\theta^2+13\alpha\theta^3\right)e^{\theta}\end{array}\right)}{\left[e^{\theta}\left\{\begin{array}{l}\left(\alpha^2+\theta\alpha\right)e^{4\theta}+\left(8\alpha\theta-4\alpha^2\right)e^{3\theta}+\left(6\alpha^2-14\theta\alpha+4\theta^2\right)e^{2\theta}\\+\left(4\theta^2-4\alpha^2\right)e^{\theta}+\left(\alpha^2+4\theta^2+5\alpha\theta\right)\end{array}\right\}\right]^2}$$

$$\gamma=\frac{\sigma^2}{\mu_1'}=\frac{e^{\theta}\left\{\begin{array}{l}\left(\alpha^2+\theta\alpha\right)e^{4\theta}+\left(8\alpha\theta-4\alpha^2\right)e^{3\theta}+\left(6\alpha^2-14\theta\alpha+4\theta^2\right)e^{2\theta}\\+\left(4\theta^2-4\alpha^2\right)e^{\theta}+\left(\alpha^2+4\theta^2+5\alpha\theta\right)\end{array}\right\}}{\left\{(\theta+\alpha)e^{2\theta}+(4\theta-2\alpha)e^{\theta}+\theta+\alpha\right\}\left\{\alpha\left(e^{\theta}-1\right)^2+\theta\left(e^{\theta}+1\right)\right\}\left(e^{\theta}-1\right)}$$

The following tables 1, 2, 3, 4, 5 and 6 summarize the behavior of mean, variance, coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of DQAD for varying values of the parameters θ and α.

TABLE 1: Mean of DQAD for varying values of parameters θ and α .

θ \ α	0.2	0.5	1.0	2.0	3.0	4.0	5.0
0.2	14.7738	14.4520	13.9584	13.1050	12.3930	11.7901	11.2729
0.5	5.7355	5.3905	4.9264	4.2687	3.8251	3.5056	3.2645
1	2.6814	2.3235	1.9381	1.5220	1.3013	1.1645	1.0715
2	1.1199	0.8027	0.5737	0.4007	0.3291	0.2900	0.2653
3	0.5879	0.3494	0.2229	0.1444	0.1154	0.1003	0.0911
4	0.3131	0.1601	0.0945	0.0580	0.0452	0.0387	0.0347
5	0.1564	0.0724	0.0407	0.0240	0.0183	0.0155	0.0138

TABLE 2: Variance of DQAD for varying values of parameters θ and α .

θ \ α	0.2	0.5	1.0	2.0	3.0	4.0	5.0
0.2	76.2384	77.8252	79.8573	82.2208	83.0780	83.0112	82.3746
0.5	12.6237	13.2355	13.6829	13.5792	13.0206	12.3745	11.7518
1	3.3793	3.5835	3.5169	3.1114	2.7559	2.4866	2.2820
2	0.9468	0.8997	0.7405	0.5507	0.4547	0.3979	0.3605
3	0.4576	0.3492	0.2455	0.1652	0.1323	0.1146	0.1034
4	0.2586	0.1558	0.0973	0.0610	0.0477	0.0408	0.0366
5	0.1401	0.0708	0.0410	0.0245	0.0187	0.01580	0.0140

TABLE 3: C.V of DQAD for varying values of parameters θ and α .

θ \ α	0.2	0.5	1.0	2.0	3.0	4.0	5.0
0.2	0.5910	0.6104	0.6402	0.6919	0.7355	0.7728	0.8051
0.5	0.6194	0.6749	0.7509	0.8632	0.9434	1.0033	1.0501
1	0.6856	0.8147	0.9676	1.1589	1.2757	1.3541	1.4099
2	0.8688	1.1847	1.4999	1.8521	2.0489	2.1752	2.6230
3	1.1507	1.6912	2.2231	2.8134	3.1509	3.3731	3.5313
4	1.6239	2.4645	3.3010	4.2607	4.8323	5.2206	5.5039
5	2.3927	3.6740	4.9729	6.5106	7.4590	8.1206	8.6134

TABLE 4: Skewness of DQAD for varying values of parameters θ and α .

θ \ α	0.2	0.5	1.0	2.0	3.0	4.0	5.0
0.2	1.1255	1.0936	1.0631	1.0515	1.0720	1.1062	1.1461
0.5	1.0659	1.0136	1.0209	1.1301	1.2527	1.3630	1.4586
1	0.9710	0.9834	1.1411	1.4338	1.6438	1.7963	1.9105
2	0.9005	1.2377	1.6741	2.1775	2.4579	2.6338	2.7520
3	0.9595	1.6836	2.3717	3.0976	3.4922	3.7411	3.9115
4	1.3339	2.4011	3.3831	4.4542	5.0683	5.4749	5.7656
5	2.1288	3.5930	5.0022	6.6234	7.6043	8.2805	8.7798

TABLE 5: Kurtosis of DQAD for varying values of parameters θ and α .

θ \ α	0.2	0.5	1.0	2.0	3.0	4.0	5.0
0.2	4.9347	4.8533	4.7542	4.6553	4.6421	4.6808	4.7520
0.5	4.7988	4.6169	4.5152	4.6519	4.9514	5.2937	5.6381
1	4.5321	4.3438	4.5637	5.4151	6.2576	6.9856	7.5983
2	4.1397	4.6292	6.0795	8.5435	10.2775	11.5043	12.3912
3	3.7986	5.8038	9.1002	13.8928	17.0323	19.1868	20.7265
4	4.0260	8.4824	14.9126	24.3392	30.8064	35.4751	38.9829
5	6.3473	15.3388	28.2938	48.2803	62.9261	74.0981	82.8908

TABLE 6: Index of Dispersion of DQAD for varying values of parameters θ and α .

θ \ α	0.2	0.5	1.0	2.0	3.0	4.0	5.0
0.2	5.1604	5.3851	5.7211	6.2740	6.7036	7.0408	7.3073
0.5	2.2010	2.4553	2.7775	3.1811	3.4040	3.5300	3.5999
1	1.2603	1.5423	1.8146	2.0443	2.1178	2.1353	2.1298
2	0.8454	1.1208	1.2908	1.3745	1.3816	1.3721	1.3587
3	0.7784	0.9995	1.1015	1.1433	1.1462	1.1417	1.1357
4	0.8258	0.9726	1.0296	1.0526	1.0556	1.0546	1.0526
5	0.8956	0.9775	1.0063	1.0184	1.0205	1.0207	1.0202

It is obvious from above tables that the mean, variance, and index of dispersion of DQAD are decreasing for increasing values of the parameters θ and α, while coefficient of variation, skewness and kurtosis are increasing for increasing values of parameters θ and α . Since the index of dispersion of DQAD is sometimes greater than 1 and less than 1 for some values of parameters θ and α, it is a suitable model for both over-dispersed and under-dispersed data.

MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS

Let $\left(y_1, y_2, y_3, \ldots, y_n\right)$ be a random sample from DQAD (2.1). The likelihood function, L of (2.1) is given by

$$L = \left(\frac{\left(e^{\theta}-1\right)^3}{e^{\theta}\left(\alpha\left(e^{\theta}-1\right)^2+\theta\left(e^{\theta}+1\right)\right)} \right)^n \prod_{i=1}^{n}\left(\alpha+\theta y_i^2\right)e^{-\theta y_i} .$$

The natural log likelihood function is thus obtained as

$$L = n\ln\left(\frac{\left(e^{\theta}-1\right)^3}{e^{\theta}\left(\alpha\left(e^{\theta}-1\right)^2+\theta\left(e^{\theta}+1\right)\right)} \right) + \sum_{i=1}^{n}\ln\left(\alpha+\theta y_i^2\right) - n\theta\,\overline{y}$$

$$= 3n\ln\left(e^{\theta}-1\right) - n\theta - n\ln\left(\alpha(e^{\theta}-1)^2+\theta(e^{\theta}+1)\right) + \sum_{i=1}^{n}\ln\left(\alpha+\theta y_i^2\right) - n\theta\,\overline{y} .$$

The maximum likelihood estimates (MLE) of $(\hat{\theta},\hat{\alpha})$ of parameters (θ,α) of DQAD are the solution of the following two non-linear equations

$$\frac{\partial \ln L}{\partial \theta}=\frac{3ne^{\theta}}{e^{\theta}-1}-\frac{n\left(2\alpha e^{\theta}\left(e^{\theta}-1\right)+\left(e^{\theta}+1\right)+\theta e^{\theta}\right)}{\alpha\left(e^{\theta}-1\right)^{2}+\theta\left(e^{\theta}+1\right)}-n\left(\overline{y}+1\right)+\sum_{i=1}^{n}\frac{y_i^2}{\alpha+\theta y_i^2}=0$$

$$\frac{\partial \ln L}{\partial \alpha}=-\frac{n\left(e^{\theta}-1\right)^{2}}{\alpha\left(e^{\theta}-1\right)^{2}+\theta\left(e^{\theta}+1\right)}+\sum_{i=1}^{n}\frac{1}{\alpha+\theta y_i^2}=0$$

These two log-likelihood equations do not seem to be solved directly, because they cannot be expressed in closed forms. However, the MLE's $(\hat{\theta},\hat{\alpha})$ of (θ,α) can be computed directly by solving the log-likelihood equation using Newton-Raphson iteration method available in R-software till sufficiently close estimates of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

GOODNESS OF FIT

In this section, the goodness of fit of the DQAD has been discussed with two count datasets from thunderstorm events available in Falls *et al* (1971) and Carter (2001) and the fit has been compared with discrete Akash distribution (DAD), Poisson-Akash distribution (PAD), Poisson-Lindley distribution (PLD) introduced by Sankaran (1970) and Quasi Poisson -Akash distribution (QPAD).[16, 17,18] It is obvious from the goodness of fit given in tables 7 and 8 that in table 7 DQAD is the best distribution whereas in table 8 DAD is the best distribution.

TABLE 7: Observed and expected number of days that experienced X thunderstorms event at Cape Kennedy, Florida for 11 year period of record for the month of June, January 1957 to December 1967.

	Observed	Expected Frequency			
	Frequency	DAD	PLD	QPAD	DQAD
0	187	186.2	185.3	184.6	187.0
1	77	77.7	83.5	83.4	76.9
2	40	40.5	35.9	36.8	40.1
3	17	16.9	15.0	15.4	17.0
4	6	6.0	6.1	6.1	6.1
5	2	2.0	2.5	2.3	2.0
6	1	0.7	1.7	1.4	0.9
Total	330	330.0	330.0	330.0	330.0
ML estimate $(\hat{\alpha})$				0.9462	0.9462
ML estimate $(\hat{\theta})$		1.5673	1.8042	2.5402	1.5451
χ^2		0.0269	1.42	1.0321	0.0004
d.f.		3	3	2	2
P-value		0.9999	0.8400	0.9049	1.0000

TABLE 8: Observed and expected number of days that experienced X thunderstorms event at Cape Kennedy, Florida for 11 year period of record for the month of July, January 1957 to December 1967.

	Observed	Expected Frequency			
	Frequency	DAD	PLD	QPAD	DQAD
0	177	177.9	177.7	172.0	174.2
1	80	81.8	88.0	91.4	84.0
2	47	47.0	41.5	44.1	48.7
3	26	21.6	18.9	19.5	21.8
4	9	8.4	8.4	8.0	8.2
5	2	4.3	6.5	5.0	4.1
Total	341	341.0	341.0	341.0	341.0
ML estimate $(\hat{\alpha})$				0.5585	1.2739
ML estimate $(\hat{\theta})$		1.47054	1.583536	2.4935	1.5217
χ^2		1.1680	5.1470	4.2323	1.2414
d.f.		3	3	2	2
P-value		0.8833	0.3538	0.3755	0.8712

CONCLUDING REMARKS

In this paper, a discrete quasi Akash distribution (DQAD), a discrete counterpart of the continuous quasi Akash Distribution has been proposed. Some statistical properties as well as the estimation of its parameters using maximum likelihood method have been studied. Applications of DQAD have been discussed with two examples of observed real datasets from thunderstorms events and the goodness of fit of DQAD has been found quite satisfactory over other discrete distributions including PLD and QPAD and competing well with DAD.

Acknowledgements

Authors are highly grateful to the anonymous reviewers for their fruitful comments which enhanced both the quality and the presentation of the paper.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Rama Shanker;* ***Design:*** *Berhane Abebe;* ***Control/Supervision:*** *Rama Shanker;* ***Data Collection and/or Processing:*** *Berhane Abebe;* ***Analysis and/or Interpretation:*** *Rama Shanker;* ***Literature Review:*** *Berhane Abebe;* ***Writing The Article:*** *Berhane Abebe;* ***Critical Review:*** *Rama Shanker;* ***References and Fundings:*** *Berhane Abebe;* ***Materials:*** *Rama Shanker*

REFERENCES

1. Lai CD. Issues concerning constructions of discrete lifetime models. Qual Techno Quant Manag 2013;10(2):251-62.
2. Good LJ. The population frequencies of species and the estimation of population parameters. Biometrika 1953;40(3/4):237-64.
3. Kulasekara KB, Tonkyn DW. A new discrete distribution, with application to survival, dispersal and dispersion. Commun Stat Simul Comput 1992;21(2):499-518.
4. Doray LG, Luong A. Efficient estimators for the good family. Commun Stat Simul Comput 1997;21:499-518.
5. Sato H, Ikota M, Aritoshi S, Masuda H. A new defect distribution in meteorology with a consistent discrete exponential formula and its applications. IEEE Trans Semicond Manufactur 1999;12(4):409-18.
6. Nekoukhou VM, Alamatsaz MH, Bidram H. Discrete generalized exponential distribution. Communications in Statistics-Theory & Methods 2012;41(11):2000-13.
7. Berhane A, Shanker R. A discrete Akash distribution with applications. Turkiye Klinikleri J Biostat 2018;10(1):1-12.
8. Shanker R. Akash distribution and its applications. International Journal of Probability and Statistics 2015;4(3):65-75.
9. Shanker R. The discrete poisson-Akash distribution. International Journal of Probability and Statistics 2017;6(1):1-10.
10. Shanker R, Shukla KK, Shanker R. A quasi Poisson-Akash distribution and its applications to ecology. International Journal of Statistics and Applied Mathematics 2018;3(3):111-9.
11. Shanker R. A quasi Akash distribution. Assam Statistical Review 2016;29-30(1-2):135-60.
12. Nakagawa T, Osaki S. The discrete Weibull distribution. IEEE Trans Reliability 1975;R-24(5):300-1.
13. Kemp AW. Classes of discrete lifetime distributions. Commun Stat Theory Methods 2004;33(12):3069-93.
14. Grandell J. Mixed Poisson Processes. 1st ed. London: Chapman & Hall; 1997. p.268.
15. Keilson J, Gerber H. Some resilts for discrete unimodality. J Am Stat Assoc 1971;66(334):386-9.
16. Falls LW, Wildford WO, Carter MC. Probability distributions for thunderstorms activity at Cape Kennedy, Florida. Journal of Applied meteorology 1971;10:97-104.
17. Carter MC. A model for thunderstorms activity: use of the compound negative binomial-positive binomial distribution. Journal of the Royal Statistical Society Series C (Applied Statistics) 2001;21(2):196-201.
18. Sankaran M. The discrete poisson-lindley distribution. Biometrics 1970;26(1):145-9.