Austrian Journal of Statistics
2025, Volume 54, 51–78.
http://www.ajs.or.at/
doi:10.17713/ajs.v54i5.2019

On an Alternative to Discrete Pareto Model

K. Jayakumar
Department of Statistics
University of Calicut

Jiji Jose
Department of Statistics
University of Calicut

Abstract

Here we introduce a novel and versatile discrete distribution as an alternative to discrete Pareto distribution and derive some noteworthy distributional properties. The proposed distribution has some interesting characteristics, such as decreasing hazard rate, unimodality, overdispersion and infinite divisibility. Some characterizations are also obtained. We also consider different estimation procedures for estimating the model parameters, namely: the maximum likelihood, least squares, weighted least squares, Cramer-von Mises methods and compare them using a comprehensive simulation study. Lastly, to illustrate the practical utility of the proposed model, we analyze four real datasets.

Keywords: characterizations, discrete Pareto, infinite divisibility, maximum likelihood estimation, simulation, unimodality.

1. Introduction

The Pareto model is a widely recognized continuous distribution that has been extensively employed over the past several decades for data modeling across various domains, notably in income and inequality of wealth, engineering, actuarial science and biological sciences. Because of its versatility in data modeling, numerous statisticians have dedicated their efforts to developing and thoroughly examining various generalizations of the Pareto distribution. A broad spectrum of socio-economic variables exhibit heavy-tailed distributions that are suitably approximated by Pareto distributions. Various forms of the Pareto distribution and their generalizations exist in the literature. For instance, the name generalized Pareto distribution (GPD) was first used by Pickands III (1975) when making statistical inferences about the upper tail of a distribution function. Pareto distribution is a special case of the GPD. Pareto distribution has been extended to the transformed Pareto distribution, otherwise called the Burr distribution. To add flexibility to Pareto distribution, various generalizations of the distribution have been derived. Stoppa (1990) proposed a three-parameter generalization of the Pareto distribution. The four parameter beta Pareto (BP) distribution was studied by Akinsete, Famoye, and Lee (2008). Arnold (2008) studied various properties of Pareto distribution and its extensions using transformations of variables. Jayakumar, Krishnan, and Hamedani (2020) introduced a generalization of Pareto distribution and studied its properties. Across literature, numerous applications of the Pareto distribution can be observed in diverse fields such as social studies, economics, and physics, see Jayakumar and Mathew (2008),Cifarelli, Gupta, and Jayakumar (2010) and Tahir, Cordeiro, Mansoor, Zubair, and Alzaatreh (2021).

We know that in numerous scenarios, data must be recorded on a discrete scale rather than a continuous one. For instance, this might include the number of failures for a specific engineering system per week, the daily count of COVID- 19 cases, the annual tally of fires in a country, or the number of kidney cysts observed in patients undergoing steroid treatment. Another example is studying the multi-system in- flammatory syndrome in children (MISC) associated with COVID-19, where various body parts can become inflamed, such as the brain, heart, eyes, lungs, or skin. So here we highlight some discrete counter parts of Pareto model exist in the literature. For instance, Krishna and Pundir (2009) developed the discrete Burr and discrete Pareto distributions for reliability modeling. Buddana and Kozubowski (2014) studied a generalization of the discrete Pareto distribution. Prieto, Gómez-Déniz, and Sarabia (2014) constructed another discrete generalized Pareto distribution, used to model the number of road crashes on blackspots. Para and Jan (2016) introduced discrete three-parameter Burr type XII and discrete Lomax distributions for modeling count data of kidney cysts using steroids. Ghosh (2020) developed the discrete Pareto type IV distribution. Jayakumar and Jose (2022) developed a discrete new generalized Pareto distribution. Thus discrete Pareto models have lots of application in many real life situations.

The aim of this paper is to propose a three parameter alternative discrete Pareto (ADP) distribution for modeling count data in real life situations. This model is shown to perform better than the traditional discrete Pareto models with respect to four real data applications in the field of biology, medical science, reliability engineering and epidemology.

Let Y be a non negative integer random variable (rv) having Probability Mass Function (PMF)

$$f_Y(y;\lambda,\beta,v) = v^{\log(1+\frac{\lambda}{\beta}y)} - v^{\log(1+\frac{\lambda}{\beta}(y+1))} \tag{1}$$

where $\lambda > 0, \beta > 0, 0 < v < 1$ and $y \in \mathbf{N_0}$. It can be easily verified that Equation (1) is a proper PMF. We call the rv following Equation (1) as alternative discrete Pareto distribution and denote it as ADP(λ, β, v).

The main motivations behind proposing the ADP distribution are as follows:

1. There are many data sets with discrete and heavy tailed property (for example, number of insurance claims and length of stay in hospitals), but not many discrete heavy tailed distributions are available in the literature. The model is specifically designed for modeling the probability distribution of count data with heavy tails.

2. We know that there are several distributions that are frequently used in practice, such as exponential, Erlang and Weibull distributions, in the continuous cases, for modeling system reliability. But a discrete approach can be appropriate for modeling the evolution of the systems over time. The discrete model can be used to model reliability in series system.

3. Infinite divisibility have a permanent position in the theory of probability. This is mainly due to its importance in solving the general central limit problem and its applications to stochastic processes with stationary independent increments. The model we propose is infinitely divisible and has compound Poisson representation.

4. In every discipline where the current discrete Pareto models fails, including reliability engineering, epidemiology, insurance, medicine and biology a new discrete Pareto model is much needed.

5. To achieve consistently superior fitting performance compared to other existing discrete Pareto distributions as well as compared to other popular discrete distributions documented in the literature.

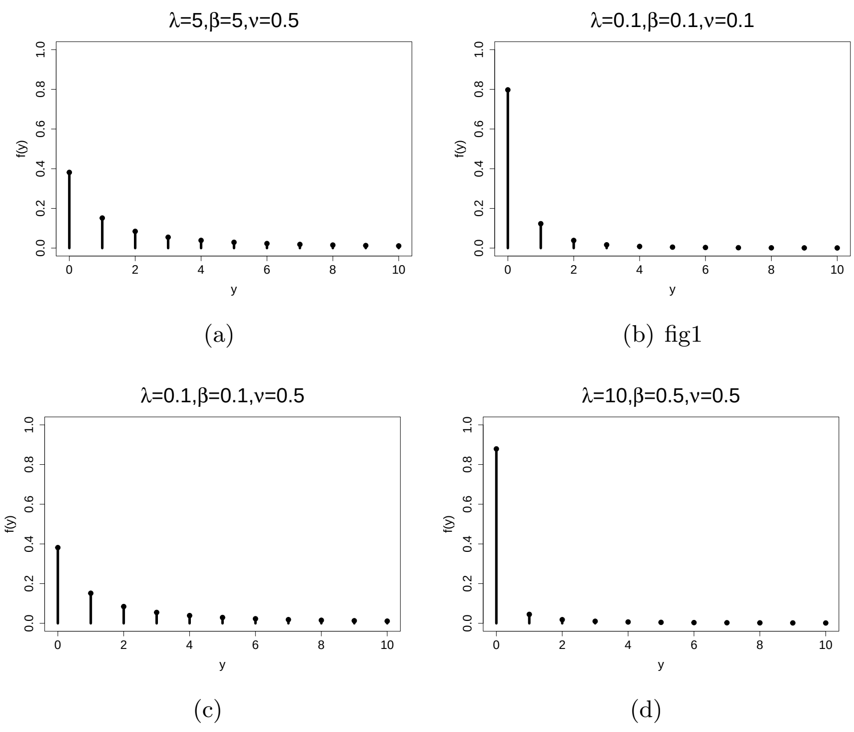
The remaining parts of this article are as follows: In Section 2, the ADP distribution is presented by reparameterizing one of its parameter v. Some statistical characteristics such as limiting behaviour, stochastic ordering, failure rate, infinite divisibility and characterizations of the ADP distribution are derived in Section 3. The ADP parameters are estimated by utilizing different methods of estimation such as maximum likelihood method, least squares method, weighted least squares and Cramer-von Mises estimation and testing of hypothesis in Section 4. A simulation study is performed to examine how the estimators perform with regard to estimating the parameters of the proposed model in Section 5. In Section 6, the goodness-of-fit of the proposed model to four real datasets is demonstrated. In Section 7, discussion with other method is done. Finally, Section 8 provides some conclusions.

2. Structure of ADP distribution

We start with an alternative parameterization of the ADP model, which helps to study some of its properties in an easy way. In 1, reparameterizing, $v = exp(-\alpha)$, we get

$$f_Y(y; \lambda, \beta, \alpha) = (1 + \frac{\lambda}{\beta} y)^{-\alpha} - (1 + \frac{\lambda}{\beta}(y + 1))^{-\alpha} \quad (2)$$

where $\lambda > 0, \beta > 0, \alpha > 0$ and $y \in \mathbf{N_0}$ and in this case we write $Y \stackrel{d}{=} ADP(\lambda, \beta, \alpha)$. The interpretation of the parameters of the ADP distribution is as follows. For this model, λ and β are the scale parameters and α is the shape parameter. From the plots in Figure 1, the following observations can be made: a) the distribution is always unimodal and the unique mode at 0. b) positively skewed.



(a) (b) fig1

(c) (d)

Figure 1: The PMF plots of ADP(λ, β, v) for different values of λ, β and v

The Cumulative Distribution Function (CDF) and Survival Function (SF) of ADP(λ, β, v) distribution can be obtained as

$$F_Y(y; \lambda, \beta, v) = 1 - v^{\log(1+\frac{\lambda}{\beta}(y+1))} \tag{3}$$

and

$$S_Y(y; \lambda, \beta, v) = v^{\log(1+\frac{\lambda}{\beta}(y+1))} \tag{4}$$

respectively, and $y \in \mathbf{N_0}$.

- If $\lambda = \beta = 1$, ADP reduces to Discrete Pareto distribution studied in Krishna and Pundir (2009).
- If $\beta = 1$, ADP reduces to Discrete Lomax distribution discussed in Prieto *et al.* (2014).

3. Some distributional properties

3.1. Limiting behaviour

The limiting behaviour of ADP(λ, β, v) distribution corresponding to various parameter choices at the boundary are:

- $lim_{\lambda \to 0,\infty} f_Y(y; \lambda, \beta, v) = 0$
- $lim_{\beta \to 0} f_Y(y; \lambda, \beta, v) = 0$, $lim_{\beta \to \infty} f_Y(y; \lambda, \beta, v) = \infty$
- $lim_{v \to 0,1} f_Y(y; \lambda, \beta, v) = 0$
- As $y \to \infty$, $f_Y(y; \lambda, \beta, v) \to 0$
- As $y \to \infty$, $\frac{F_Y(t+y;\lambda,\beta,v)}{1-F_Y(y;\lambda,\beta,v)}$ becomes 1, which implies ADP is a long tailed distribution.

3.2. Stochastic ordering

Stochastic ordering serves as a fundamental tool for assessing the comparative behaviours of random variables. Numerous stochastic orders exist, each with diverse applications. Theorem 1 and Corollary 1 (below) present results regarding the stochastic orderings of ADP distribution. The orders under consideration here are the stochastic order $\leq_{st}$ and the expectation order $\leq_E$.

Theorem 1. *The ADP(λ, β, v) has the following properties.*

- *Suppose $Y_1 \sim ADP(\lambda_1, \beta, v)$ and $Y_2 \sim ADP(\lambda_2, \beta, v)$. If $\lambda_1 < \lambda_2$, then $Y_2 \leqslant_{st} Y_1$.*
- *Suppose $Y_1 \sim ADP(\lambda, \beta_1, v)$ and $Y_2 \sim ADP(\lambda, \beta_2, v)$. If $\beta_1 < \beta_2$, then $Y_1 \leqslant_{st} Y_2$.*
- *Suppose $Y_1 \sim ADP(\lambda, \beta, v_1)$ and $Y_2 \sim ADP(\lambda, \beta, v_2)$. If $v_1 < v_2$, then $Y_1 \leqslant_{st} Y_2$.*

Proof. Follows easily and hence is omitted. □

Corollary 1. *From Theorem 1, the expectation ordering are as follows:*

- *Suppose* $Y_1 \sim ADP\,(\lambda_1, \beta, v)$ *and* $Y_2 \sim ADP\,(\lambda_2, \beta, v)$.
  If $\lambda_1 < \lambda_2$, *then* $Y_2 \leqslant_E Y_1$.

- *Suppose* $Y_1 \sim ADP\,(\lambda, \beta_1, v)$ *and* $Y_2 \sim ADP\,(\lambda, \beta_2, v)$.
  If $\beta_1 < \beta_2$, *then* $Y_1 \leqslant_E Y_2$.

- *Suppose* $Y_1 \sim ADP\,(\lambda, \beta, v_1)$ *and* $Y_2 \sim ADP\,(\lambda, \beta, v_2)$.
  If $v_1 < v_2$, *then* $Y_1 \leqslant_E Y_2$.

3.3. Failure rates

The failure rate is given by

$$\begin{aligned} r_Y(y;\lambda,\beta,v) =& \frac{f_Y(y;\lambda,\beta,v)}{S_Y(y;\lambda,\beta,v)} \\ =& v^{\Phi(y)} - 1 \end{aligned} \tag{5}$$

where $\Phi(y) = \log\left[\frac{1+\frac{\lambda}{\beta}y}{1+\frac{\lambda}{\beta}(y+1)}\right]$. Note that $r_Y(1;\lambda,\beta,v) = r_Y(0;\lambda,\beta,v) \Leftrightarrow \lambda/\beta = 0$. But this cannot happen since $\lambda > 0$ and $\beta < \infty$. Then $r_Y(y;\lambda,\beta,v) < r_Y(y-1;\lambda,\beta,v)$ for all choices of $\lambda/\beta = 1, \lambda/\beta > 1,$ and $\lambda/\beta < 1$. Thus ADP(λ,β,v) has decreasing failure rate for all y.

The reverse failure rate is given by

$$r^*_Y(y;\lambda,\beta,v) = \frac{v^{\log(1+\frac{\lambda}{\beta}y)} - v^{\log(1+\frac{\lambda}{\beta}(y+1))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(y+1))}}. \tag{6}$$

Since failure rate is decreasing, the reverse failure rate is also decreasing.
In discrete distributions, the failure rate $r_Y(y;\lambda,\beta,v)$ is a conditional probability with unity as its upper bound. Xie, Gaudoin, and Bracquemond (2002) highlighted that labeling this as the failure rate function could contribute to the existing confusion in the industry, where failure rate and failure probability are occasionally conflated. To address this issue, they introduced a second rate of failure $r^{**}_Y(y;\lambda,\beta,v)$ with the same monotonicity as $r_Y(y;\lambda,\beta,v)$. For ADP(λ,β,v) we have,

$$\begin{aligned} r^{**}_Y(y;\lambda,\beta,v) &= -log\left[\frac{S_Y(y;\lambda,\beta,v)}{S_Y(y+1;\lambda,\beta,v)}\right] \\ &= -\log\left[v^{\log\left[\frac{1+\frac{\lambda}{\beta}(y+1)}{1+\frac{\lambda}{\beta}(y+2)}\right]}\right]. \end{aligned} \tag{7}$$

According to Kemp (2004), one can establish the following relationships for discrete distributions, which are applicable to the ADP distribution is given below:
DFR (Decreasing Failure Rate)⇒ DFRA (Decreasing Failure Rate Average) ⇒
NWU ((New Worse than Used) ⇒ NWUE ((New Worse than Used in Expectation) ⇒ IMRL (Increasing Mean Residual Life).

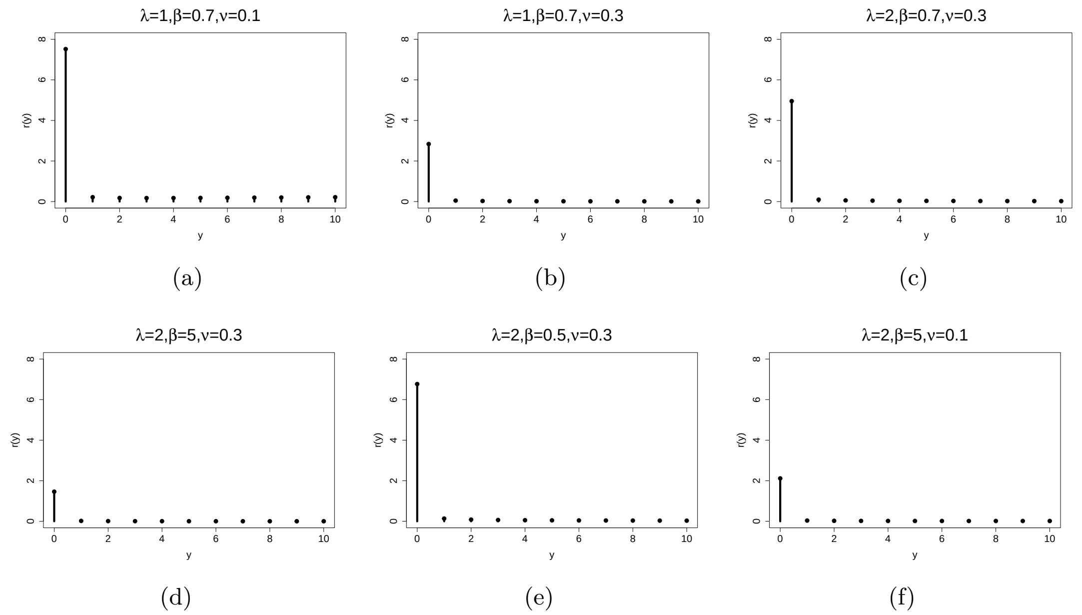


Figure 2: The failure rate plots of ADP(λ, β, v) for different values of λ, β and v

3.4. Moments

The r^{th} moment of the ADP(λ, β, v) is given by

$$\begin{aligned} E\left(Y^r\right) &= \sum_{y=0}^{\infty} y^r f_Y(y; \lambda, \beta, v) \\ &= \sum_{y=1}^{\infty} \left(y^r - (y-1)^r\right) S_Y(y; \lambda, \beta, v) \\ &\leq r \sum_{y=1}^{\infty} y^{r-1} v^{\log\left(1+\left(\frac{(y+1)\lambda}{\beta}\right)\right)} \\ &\leq r \sum_{y=1}^{\infty} y^{r-1} \left(\frac{1}{\frac{\lambda}{\beta} y^{\alpha-r+1}}\right). \end{aligned}$$

This is finite when $\alpha > r$.
Also,

$$\begin{aligned} E\left(Y^r\right) &= \sum_{y=0}^{\infty} y^r f_Y(y; \lambda, \beta, v) \\ &= \sum_{y=1}^{\infty} \frac{\left(y^r - (y-1)^r\right)}{\left(1 + \left(\frac{(y+1)\lambda}{\beta}\right)\right)^{\alpha}} \\ &> \sum_{y=1}^{\infty} \frac{y^{r-1}}{\left(1 + \left(\frac{(y+1)\lambda}{\beta}\right)\right)^{\alpha}} \\ &> \frac{1}{1+\lambda/\beta} \sum_{y=1}^{\infty} \left(\frac{1}{\frac{\lambda}{\beta} y^{\alpha-r+1}}\right). \end{aligned}$$

That is $E(Y^r)$ is finite if $\alpha > r$.
Hence $\frac{1}{1+\lambda/\beta} \sum_{y=1}^{\infty} \left(\frac{1}{\frac{\lambda}{\beta} y^{\alpha-r+1}}\right) < E(Y^r) \leq r \sum_{y=1}^{\infty} y^{r-1} \left(\frac{1}{\frac{\lambda}{\beta} y^{\alpha-r+1}}\right)$.

In particular, the mean lifetime of ADP(λ, β, v) can be obtained as

$$\begin{aligned}\mu(\lambda, \beta, v) &= \sum_{y=1}^{\infty} v^{\log(1+\frac{\lambda}{\beta}(y+1))} \\ &= \sum_{y=1}^{\infty} \frac{1}{(1+\frac{\lambda}{\beta}(y+1))^{\alpha}}\end{aligned}$$

where $v = e^{-\alpha}$. This series is convergent if $\alpha > 1$. Consequently, we have the following theorem:

Theorem 2. *$E(Y^r)$ exist iff $exp(-r) > v$.*

Proof. Follows from the previous discussion. □

3.5. Discrete aging intensity and Discrete alternative aging intensity

The aging intensity L is defined as the ratio of instantaneous failure rate to failure rate average. The discrete aging intensity and discrete alternative aging intensity of ADP(λ, β, v) are

$$\begin{aligned}L_Y(y; \lambda, \beta, v) &= y\left[1 - \frac{\log S(y-1; \lambda, \beta, v)}{\log S(y; \lambda, \beta, v)}\right] \\ &= y\left[1 - \frac{\log\left(1+\frac{\lambda}{\beta}y\right)}{\log(1+\frac{\lambda}{\beta}(y+1))}\right]\end{aligned} \tag{8}$$

and

$$\begin{aligned}L_Y^*(y; \lambda, \beta, v) &= \frac{log \frac{\log S(y;\lambda,\beta,v)}{\log S(y-1;\lambda,\beta,v)}}{\log(y/y-1)} \\ &= \frac{\log\left(\log(1+\frac{\lambda}{\beta}(y+1))\right) - \log\left(\log(1+\frac{\lambda}{\beta}y)\right)}{\log y - \log(y-1)}\end{aligned} \tag{9}$$

respectively.

Remark 1. *If for a discrete rv Y, discrete aging intensity L has the form in Equation (8) for $k = 1, 2, 3, \ldots$, $\lambda, \beta > 0$, then Y follows ADP(λ, β, v) distribution.*

From Table 1, it is noted that the ADP model can be used as a probability tool for modeling positively skewed data with leptokurtic shape. Moreover, it can be used to model over dispersed data.

Table 1: Mean, Variance, Skewness and Kurtosis for various values of λ, β and v

Parameter	Mean	Variance	Skewness	Kurtosis
$\lambda = 0.5$ $\beta = 0.5$ $v = 0.1$	0.4076	1.6119	49.9028	79.9414
$\lambda = 0.5$ $\beta = 0.5$ $v = 0.5$	2.6584	21.9575	6.2451	9.3161
$\lambda = 0.5$ $\beta = 0.5$ $v = 0.9$	1.6975	19.9875	10.1187	12.9989
$\lambda = 1.0$ $\beta = 0.5$ $v = 0.5$	1.8510	15.7551	10.1603	14.0555
$\lambda = 0.1$ $\beta = 5.0$ $v = 0.5$	2.7641	36.6448	4.9357	6.7892

3.6. Theorems related to ADP distribution

Theorem 3. *Let $Y_i's$ $(i = 1, 2, 3, \dots n)$ be non-negative independent and identically distributed (iid) distributed integer valued rvs, and $Z = \min_{1\le i\le n} Y_i$. Then Y is ADP(λ, β, v^n) iff Y_i is ADP (λ, β, v).*

Proof. Let $Y_i (i = 1, 2, 3, \dots n)$ be iid ADP (λ, β, v). Then $S_Y(y; \lambda, \beta, v) = v^{\log(1+\frac{\lambda}{\beta}(y+1))}; y = 0, 1, 2, 3, \dots$

Consider

$$S_Z(z; \lambda, \beta, v) = P(Z \geqslant z) = [P(Y_1 \geqslant z)]^n = (v^n)^{\log(1+\frac{\lambda}{\beta}(z+1))}$$

Thus, $Z \sim$ ADP (λ, β, v^n) for all $z = 0, 1, 2, 3, \dots$
Conversely, let $S_Z(z; \lambda, \beta, v) = v^n \log\left(1 + \frac{\lambda}{\beta}(z+1)\right); z = 0, 1, 2, ...$
We know that,

$$S_Y(y; \lambda, \beta, v) = P[Y_1 \geqslant y] = [P[Z \geqslant y]]^{1/n} = v^{\log(1+\frac{\lambda}{\beta}(y+1))}; y = 0, 1, 2, 3, \dots$$

Hence the theorem. □

This property is useful as it will allow modeling reliability of a series system with identical components having the ADP distribution.

Theorem 4. *Let $Y_i's$ $(i = 1, 2, 3, \dots n)$ be non-negative iid integer valued random variables, and $Z = \min_{1\le i\le n} Y_i$. Then Z is ADP (λ, β, v) if $Y_i's$ are ADP(λ, β, v_i) where*

$v = \prod_{i=1}^{n} v_i$.

Proof. Follows easily. □

Theorem 5. *If* $Y \sim \text{ADP}(\lambda, \beta, v)$*, then* $Z = \log(\frac{\beta}{\lambda}Y + 1) \sim Geo(v)$.

Proof.

$$\begin{aligned} &P(Z \geqslant z) = P[\log(\tfrac{\beta}{\lambda}Y + 1) \geqslant z] \\ &= P\left[Y \geqslant ((e^z - 1)\beta/\lambda)\right] \\ &= v^{\log(1+\lambda/\beta((e^z-1)\beta/\lambda))} \\ &= v^z, \quad z = 0, 1, 2, 3, \ldots \end{aligned}$$

which is the SF of Geometric r.v. This completes the proof. □

Theorem 6. *If* $Y \sim$ *Geo* (v)*, then* $Z = [\frac{\beta}{\lambda}(e^Y - 1) - 1] \sim ADP(\lambda, \beta, v)$.

Proof. Consider

$$\begin{aligned} &P(Z \geqslant z) \\ &= P\left[\left[\tfrac{\beta}{\lambda}(e^Y - 1) - 1\right] \geqslant z\right] \\ &= P\left[e^Y - 1 \geqslant \tfrac{\lambda}{\beta}(z+1)\right] \\ &= P\left[e^Y \geqslant 1 + \tfrac{\lambda}{\beta}(z+1)\right] \\ &= P[Y \geqslant \log(1 + \tfrac{\lambda}{\beta}(z+1))] \\ &= v^{\log(1+\frac{\lambda}{\beta}(z+1))} \end{aligned}$$

Hence the theorem. □

Theorem 7. *If* Y_1 *and* Y_2 *are independent and follow Geo*(v) *and* $ADP(\lambda, \beta, v)$ *and* $\lambda/\beta = 1$*, respectively, then* $Z = min(Y_1, Y_2) \sim NGDP(v, \alpha)$ *of Bhati and Bakouch (2019).*

Proof. The SF of $Y_1 \sim$ Geo (v) and $Y_2 \sim$ ADP (λ, β, v) are v^y and $v^{log((y+2))}$ respectively. Hence the SF of $min(Y_1, Y_2)$ is

$$\begin{aligned} S_Z(z) &= P(min(Y_1, Y_2) \geqslant z) \\ &= v^z * v^{log(z+2)} \\ &= \frac{v^z}{(z+2)^\alpha} \end{aligned}$$

where $\alpha = -lnv$ and this is the SF of NGDP(v, α). □

Theorem 8. *If* $Y \sim$ *ADP* (λ, β, v)*, then*

$$\frac{P(Y > t)}{\left(1 + \frac{\lambda}{\beta}(1+t)\right)^{-\alpha}} \to 1 \ as \ t \to \infty.$$

Proof. As $t \to \infty$, we consider $a = a(t)$ be a unique integer such that $a(t) \leq t \leq a(t) + 1$. As a consequence,

$$S_Y(a(t); \lambda, \beta, \alpha) \geqslant P(Y > t) \geqslant S_Y(a(t) + 1; \lambda, \beta, \alpha).$$

Therefore,

$$\left[\frac{(1 + \lambda/\beta(1+t))}{(1 + \lambda/\beta(1 + a(t))}\right]^\alpha \geqslant \frac{P(Y > t)}{[1 + \lambda/\beta(1+t)]^{-\alpha}} \geqslant \left[\frac{1 + \lambda/\beta(1+t)}{1 + \lambda/\beta(2 + a(t)}\right]^\alpha.$$

The term in the middle is bounded by two terms which converges to 1 as $t \to \infty$, since $t/a(t) \to 1$. □

3.7. Quantile function and random number generation

Definition 1. The point y_q is known as the q^{th} quantile of a discrete random variable Y if it satisfies $P(Y \leqslant y_q) \geqslant q$ and $P(Y \geqslant y_q) \geqslant 1-q$, that is, $F_Y(y_q-1;\lambda,\beta,\alpha) < q \leqslant F_Y(y_q;\lambda,\beta,\alpha)$.

Theorem 9. *The q^{th} quantile of the ADP distribution is*

$$y_q = \frac{\beta}{\lambda}\left[(1-q)^{-1/\alpha}-1\right]-1, \text{ where } \alpha = -lnv. \tag{10}$$

Proof.

$$P(Y \leqslant y_q) \geqslant q$$

implies

$$y_q \geqslant \frac{\beta}{\lambda}[(1-q)^{-1/\alpha}-1]-1$$

and
$P(Y \geqslant y_q) \geqslant 1-q$ gives

$$y_q \leqslant \frac{\beta}{\lambda}[(1-q)^{-1/\alpha}-1].$$

Combining these we get,

$$\frac{\beta}{\lambda}[(1-q)^{-1/\alpha}-1]-1 < y_q \leqslant \frac{\beta}{\lambda}[(1-q)^{-1/\alpha}-1],$$

and therefore, y_q is an integer given by

$$y_q = \frac{\beta}{\lambda}\left[(1-q)^{-1/\alpha}-1\right]-1.$$

Hence the proof. □

A random integer can be generated from the suggested model using the conventional inverse transformation technique. Suppose q represents a randomly selected number from a uniform distribution across the unit interval. When λ, β and α are fixed, a number adhering to the ADP distribution can be obtained by evaluating the expression on the right side of Equation (10).

3.8. Infinite divisibility

Considering $f_Y^2(y;\lambda,\beta,v) - f_Y(y-1;\lambda,\beta,v) \times f_Y(y+1;\lambda,\beta,v)$ for ADP distribution, we get

$$\begin{aligned} &v^{2\log(1+\frac{\lambda}{\beta}y)} + v^{2\log(1+\frac{\lambda}{\beta}(y+1))} - v^{\log(1+\frac{\lambda}{\beta}y)+\log(1+\frac{\lambda}{\beta}(y+1))} - v^{\log(1+\frac{\lambda}{\beta}(y-1))+\log(1+\frac{\lambda}{\beta}(y+1))} \\ &+ v^{\log(1+\frac{\lambda}{\beta}(y-1))+\log(1+\frac{\lambda}{\beta}(y+2))} - v^{\log(1+\frac{\lambda}{\beta}y)+\log(1+\frac{\lambda}{\beta}(y+2))}. \end{aligned} \tag{11}$$

The expression given in Equation (11) can be negative for $\lambda > 0, \beta > 0, 0 < v < 1$ and for all $y \in \mathbf{N_0}$ and also $f_Y(1;\lambda,\beta,v) \neq 0, f_Y(0;\lambda,\beta,v) \neq 0$. According to Steutel and van Harn (2003), a discrete model with PMF $f_Y(y;\lambda,\beta,v)$ is log convex if $f_Y^2(y;\lambda,\beta,v) < f_Y(y-1;\lambda,\beta,v)f_Y(y+1;\lambda,\beta,v)$. The ADP distribution satisfies $f_Y^2(y;\lambda,\beta,v) < f_Y(y-1;\lambda,\beta,v)f_Y(y+1;\lambda,\beta,v)$ and is log convex. As a direct consequence of log convexity, ADP distribution is infinitely divisible, see Steutel and van Harn (2003), for details.
Moreover, as any infinitely divisible distribution defined on non-negative integers is a compound Poisson distribution (see Karlis and Xekalaki (2005)), we conclude that the ADP distribution is a compound Poisson distribution. Additionally, the concept of infinitely divisible

distribution holds significance across various statistical domains, such as stochastic processes and actuarial statistics. In cases where a distribution G is infinitely divisible, it implies that for any integer $n \geq 2$, there exist a distribution G_n such that G is n-fold convolution of G_n, namely $G = G_n^{*n}$.
Moreover, in the case of infinitely divisible distributions, a maximum limit for the variance can be determined when $\lambda > 0, \beta > 0, 0 < v < 1$ (see, Johnson and Kotz (1982), page 75). This upper bound for ADP distribution is

$$V(Y) \leqslant \frac{f_Y(1;\lambda,\beta,v)}{f_Y(0;\lambda,\beta,v)} = \frac{v^{log(1+\lambda/\beta)} - v^{log(1+2\lambda/\beta)}}{1 - v^{log(1+\lambda/\beta)}}.$$

3.9. Model diagnostics

Consider the SF of ADP distribution

$$S_Y(y;\lambda,\beta,\alpha) = [1 + \frac{\lambda}{\beta}(y+1)]^{-\alpha}.$$

$$\text{That is,} \quad \log\left[[S_Y(y;\lambda,\beta,\alpha)]^{-1/\alpha} - 1\right] = \log(\frac{\lambda}{\beta}) + \log(y+1),$$

which can be written as

$$z = b + u \tag{12}$$

where $z = \log\left[[S_Y(y;\lambda,\beta,\alpha)]^{-1/\alpha} - 1\right], b = \log(\frac{\lambda}{\beta})$ and $u = \log(y+1)$. Note that Equation (12) presents a crucial means for assessing model adequacy. By calculating the empirical SF, $S_Y(y;\lambda,\beta,\alpha)$, from the dataset and plotting z against u, one can determine the suitability of the ADP distribution as a model for the given data. A nearly straight line plot indicates that the ADP distribution adequately represents the data. The parameters λ, β, and α can be estimated by considering Equation (12) as a linear regression equation and first estimating the constant b by using the ordinary least squares method.

3.10. Mean of excess over threshold and tail value at risk

Significant focus has been dedicated to quantifying operational risk in the last few decades. In this section, we derive several risk measures for the ADP(λ,β,α) distribution, including mean excess over the threshold (MT) and tail value at risk (TV). The MT of the random variable Y following ADP(λ,β,α) distribution is given as

$$\begin{aligned} m(y) &= \mathrm{E}(Y - y \mid Y \geq y) \\ &= [1 + \frac{\lambda}{\beta}(y+1)]^{-\alpha} \sum_{j=y}^{\infty} [1 + \frac{\lambda}{\beta}(j+1)]^{-\alpha}; \quad y \in \mathbb{N}_0. \end{aligned}$$

Based on an arbitrary choice of the parameters, we can get $m(0) \leq m(y)$, then the ADP distribution belongs to a class of discrete distributions having new worse than used in expectation, for more details refer Marshall and Proschan (1970). Hence, the ADP distribution plays an vital role in the reliability theory. Regarding the TV, for any quantile y_q, the TV of ADP(λ,β,α) distribution can be derived as

$$\begin{aligned} &\mathrm{TV}\, q(Y) \\ &= \mathbf{E}\,(Y \mid Y \geq y_q) \\ &= y_q + \tfrac{1}{1-F_Y(y_q;\lambda,\beta,\alpha)} \textstyle\sum_y^{y_q} (y - y_q)\, f_Y(y;\lambda,\beta,\alpha) \\ &= y_q + \tfrac{1}{1-F_Y(y_q;\lambda,\beta,\alpha)} \left(\mu_1' - y_q + \textstyle\sum_{y=0}^{y_q-1} (y_q - y)\, f_Y(y;\lambda,\beta,\alpha)\right) \\ &= y_q + [1 + \tfrac{\lambda}{\beta}(y+1)]^{-\alpha} \left(\mu_1' - y_q + \textstyle\sum_{y=0}^{y_q-1} (y_q - y)\, [1 + \tfrac{\lambda}{\beta}y]^{-\alpha} - [1 + \tfrac{\lambda}{\beta}(y+1)]^{-\alpha}\right); \\ &y \in \mathbb{N}_0. \end{aligned} \tag{13}$$

3.11. Entropy

The degree of uncertainty linked with a rv Y is termed as Entropy. This concept finds broad applicability across various domains, including information theory, econometrics, survival analysis, computer science, and quantum information. For further insights, refer to Rényi (1961). The Rényi and Shannon entropy for the rv Y is

$$I_k(Y) = \frac{1}{1-k}\log\sum_{y=0}^{\infty}\left([1+\frac{\lambda}{\beta}y]^{-\alpha} - [1+\frac{\lambda}{\beta}(y+1)]^{-\alpha}\right)^k; \quad y \in \mathbb{N}_0$$

and

$$I(Y) = -\sum_{y=0}^{\infty}\left([1+\tfrac{\lambda}{\beta}y]^{-\alpha} - [1+\tfrac{\lambda}{\beta}(y+1)]^{-\alpha}\right) \log\left([1+\tfrac{\lambda}{\beta}y]^{-\alpha} - [1+\tfrac{\lambda}{\beta}(y+1)]^{-\alpha}\right); y \in \mathbb{N}_0,$$

respectively, where $k \in (0,\infty)$ and $k \neq 1$. It is notable that the Shannon entropy can be derived as a specific case of the Rényi entropy if $k \to 1$, that is, $I(Y) = \lim_{k\to 1} I_k(Y)$.

3.12. L-moments

Given any random variables $Y_1, Y_2, \ldots, Y_n$, the order statistics $Y_{1:n} \leq Y_{2:n} \leq \ldots \leq Y_{n:n}$ are also rv. Let the rv Y follows ADP (λ, β, v) distribution. Then the CDF of the i^{th} order statistics is

$$\begin{aligned} F_{i:n}(y;\lambda,\beta,v) &= \sum_{j=i}^{n}\binom{n}{j}[F_i(y;\lambda,\beta,v)]^j\,[1-F_i(y;\lambda,\beta,v)]^{n-j} \\ &= \sum_{j=i}^{n}\sum_{k=0}^{n-j}\sum_{m=0}^{j+k}\Psi_{(n,j)}^{(m,k)}\,[1-F_i(y;\lambda,\beta,mv)]\,, \end{aligned}$$

where $\Psi_{(n,j)}^{(m,k)} = (-1)^{k+m}\binom{n}{j}\binom{n-j}{j}\binom{j+k}{m}$. The corresponding PMF of the i^{th} order statistics is given by

$$\begin{aligned} f_{i:n}(y;\lambda,\beta,v) &= F_{i:n}(y;\lambda,\beta,v) - F_{i:n}(y-1;\lambda,\beta,v) \\ &= \sum_{j=i}^{n}\sum_{k=0}^{n-j}\sum_{m=0}^{j+k}\Psi_{(n,j)}^{(m,k)}\left(F_i(y-1;\lambda,\beta,mv) - F_i(y;\lambda,\beta,mv)\right). \end{aligned}$$

Thus, the r^{th} Root Mean Square (RMS) of $Y_{i:n}$ can be computed as:

$$E\left(Y_{i:n}^r\right) = \sum_{y=0}^{\infty}\sum_{j=i}^{n}\sum_{k=0}^{n-j}\sum_{m=0}^{j+k}\Psi_{(n,j)}^{(m,k)}y^r\left(F_i(y-1;\lambda,\beta,mv) - F_i(y;\lambda,\beta,mv)\right). \quad (14)$$

Similar to the conventional ordinary moments, L-moments can be derived. However, L-moments can also be estimated using a linear combination of order statistics. Whenever the distribution has a defined mean, L-moments exist as well. Explicit expressions for L-moments, represented as infinite weighted linear combinations of the means of appropriate ADP order statistics, can be derived. These L-moments can be expressed as a linear function of the expected order statistics and based on Equation (14), can be defined by,

$$\Gamma_z = \frac{1}{z}\sum_{i=0}^{z-1}(-1)^i\binom{z-1}{i}E\left(Y_{z-i:z}\right); z = 1,2,3,\ldots \quad (15)$$

According to Equation (15), some statistical measures can be derived like mean, Skewness (Sk), and Kurtosis (Ku) where mean $= \Gamma_1$, Sk $= \frac{\Gamma_3}{\Gamma_2}$, and Ku $= \frac{\Gamma_4}{\Gamma_2}$.

3.13. Characterizations

In order to comprehend the patterns exhibited by data generated through a specific process, it is essential to articulate its behavior through an appropriate probability distribution. Hence, characterizing a distribution becomes a crucial challenge in applied sciences, as investigators are keen to ascertain whether their model adheres to the correct distribution. In pursuit of this goal, investigators depend on conditions that determine whether their model aligns with the chosen distribution. Here we obtain three characterizations of the ADP distribution based on: a) the conditional expectation of certain function of the random variable; b) the hazard rate function and c) reverse hazard rate function.

Based on conditional expectation of certain function of the random variable

Proposition 1. *$Y \sim ADP(\lambda, \beta, v)$ if and only if*

$$E\left\{v^{\log(1+\frac{\lambda}{\beta}y)} + v^{\log(1+\frac{\lambda}{\beta}(y+1))} | Y > k\right\} = v^{\log(1+\frac{\lambda}{\beta}(k+1))}. \tag{16}$$

Proof. If Y has PMF given in Equation (1), then LHS of Equation (16) will be

$$\begin{aligned}
&[1-F_Y(k)]^{-1}\sum_{y=k+1}^{\infty}\left[v^{\log\left[1+\frac{\lambda}{\beta}y\right]} + v^{\log\left[1+\frac{\lambda}{\beta}(y+1)\right]}\right]\times \\
&\left[v^{\log\left[1+\frac{\lambda}{\beta}y\right]} - v^{\log\left[1+\frac{\lambda}{\beta}(y+1)\right]}\right] \\
&= [1-F_Y(k)]^{-1}\sum_{y=k+1}^{\infty} v^{2\log\left[1+\frac{\lambda}{\beta}y\right]} - v^{2\log\left[1+\frac{\lambda}{\beta}(y+1)\right]} \\
&= \left[v^{2\log\left[1+\frac{\lambda}{\beta}(k+1)\right]}\right]\left[v^{-\log\left[1+\frac{\lambda}{\beta}(k+1)\right]}\right] \\
&= v^{\log\left[1+\frac{\lambda}{\beta}(k+1)\right]}.
\end{aligned}$$

Conversely, if Equation (16) holds, then

$$\begin{aligned}
&\sum_{y=k+1}^{\infty}\left[v^{\log\left[1+\frac{\lambda}{\beta}y\right]} + v^{\log\left[1+\frac{\lambda}{\beta}(y+1)\right]}\right] f_Y(y) \\
&= \sum_{y=k+1}^{\infty}\left[v^{\log\left[1+\frac{\lambda}{\beta}y\right]} + v^{\log\left[1+\frac{\lambda}{\beta}(y+1)\right]}\right]\times \\
&\left[v^{\log\left[1+\frac{\lambda}{\beta}y\right]} - v^{\log\left[1+\frac{\lambda}{\beta}(y+1)\right]}\right] \\
&= \sum_{y=k+1}^{\infty} v^{2\log\left[1+\frac{\lambda}{\beta}y\right]} - v^{2\log\left[1+\frac{\lambda}{\beta}(y+1)\right]} \\
&= v^{2\log\left[1+\frac{\lambda}{\beta}(k+1)\right]} \\
&= (1-F_Y(k))v^{\log\left[1+\frac{\lambda}{\beta}(k+1)\right]} \\
&= (1-F_Y(k+1)+f_Y(k+1))v^{\log\left[1+\frac{\lambda}{\beta}(k+1)\right]}.
\end{aligned} \tag{17}$$

Also we have,

$$\begin{aligned}
&\sum_{y=k+2}^{\infty}\left[v^{\log\left[1+\frac{\lambda}{\beta}y\right]} + v^{\log\left[1+\frac{\lambda}{\beta}(y+1)\right]}\right] f_Y(y) \\
&= (1-F_Y(k+1))v^{\log\left[1+\frac{\lambda}{\beta}(k+2)\right]}.
\end{aligned} \tag{18}$$

Now, substracting Equation (18) from Equation (17), we arrive at

$$\begin{aligned}
&v^{\log\left[1+\frac{\lambda}{\beta}(k+2)\right]} f_Y(k+1) = \\
&(1-F_Y(k+1))\left[v^{\log\left[1+\frac{\lambda}{\beta}(k+1)\right]} - v^{\log\left[1+\frac{\lambda}{\beta}(k+2)\right]}\right]. \\
&\text{Hence,}\quad r_Y(k+1) = \frac{f_Y(k+1)}{1-F_Y(k+1)} = \frac{v^{\log\left[1+\frac{\lambda}{\beta}(k+1)\right]} - v^{\log\left[1+\frac{\lambda}{\beta}(k+2)\right]}}{v^{\log\left[1+\frac{\lambda}{\beta}(k+2)\right]}}
\end{aligned}$$

which, in view of Equation (5), implies that Y has PMF in Equation (1). □

Based on hazard rate function

Proposition 2. *$Y \sim ADP(\lambda, \beta, v)$ if and only if its hazard rate function satisfies the difference equation*

$$r_Y(k+1) - r_Y(k) = v^{\log\left[\frac{1+\frac{\lambda}{\beta}(k+1)}{1+\frac{\lambda}{\beta}(k+2)}\right]} - v^{\log\left[\frac{1+\frac{\lambda}{\beta}k}{1+\frac{\lambda}{\beta}(k+1)}\right]}; \quad k \in \mathbf{N_0} \tag{19}$$

with boundary condition

$$r_Y(0) = v^{\log[\frac{1}{1+\lambda/\beta}]} - 1.$$

Proof. If Y has PMF in Equation (1) then clearly Equation (19) holds. Now if Equation (19) holds, then for every $y \in \mathbf{N}$, we have

$$\sum_{k=0}^{y-1} r_Y(k+1) - r_Y(k) = \sum_{k=0}^{y-1} v^{\log\left[\frac{1+\frac{\lambda}{\beta}(k+1)}{1+\frac{\lambda}{\beta}(k+2)}\right]} - v^{\log\left[\frac{1+\frac{\lambda}{\beta}k}{1+\frac{\lambda}{\beta}(k+1)}\right]}.$$

$$r_Y(y) - r_Y(0) = -v^{\log[\frac{1}{1+\lambda/\beta}]} + v^{\log\left[\frac{1+\frac{\lambda}{\beta}y}{1+\frac{\lambda}{\beta}(y+1)}\right]}.$$

In view of the fact $r_Y(0) = v^{\log[\frac{1}{1+\lambda/\beta}]} - 1$, from the last equation we have

$$r_Y(y) = v^{\log\left[\frac{1+\frac{\lambda}{\beta}y}{1+\frac{\lambda}{\beta}(y+1)}\right]} - 1$$

which in view of Equation (5), implies Y has PMF in Equation (1). □

Based on reverse hazard rate function

Proposition 3. *$Y \sim ADP(\lambda, \beta, v)$ if and only if its reverse hazard rate function satisfies the difference equation*

$$r_Y^*(k+1) - r_Y^*(k) = \frac{v^{\log(1+\frac{\lambda}{\beta}(k+1)} - v^{\log(1+\frac{\lambda}{\beta}(k+2))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(k+2))}} - \frac{v^{\log(1+\frac{\lambda}{\beta}k)} - v^{\log(1+\frac{\lambda}{\beta}(k+1))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(k+1))}}; \quad k \in \mathbf{N_0} \tag{20}$$

with boundary condition $r_Y^*(0) = 1$.

Proof. If Y has PMF in Equation (1) then clearly Equation (20) holds. Now if Equation (20) holds, then for every $y \in \mathbf{N}$, we have

$$\sum_{k=0}^{y-1} r_Y^*(k+1) - r_Y^*(k) = \sum_{k=0}^{y-1} \frac{v^{\log(1+\frac{\lambda}{\beta}(k+1)} - v^{\log(1+\frac{\lambda}{\beta}(k+2))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(k+2))}} - \frac{v^{\log(1+\frac{\lambda}{\beta}k)} - v^{\log(1+\frac{\lambda}{\beta}(k+1))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(k+1))}}.$$

This implies,

$$r_Y^*(y) - r_Y^*(0) = \frac{v^{\log(1+\frac{\lambda}{\beta}y} - v^{\log(1+\frac{\lambda}{\beta}(y+1))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(y+1))}} - 1.$$

In view of the fact that $r_Y^*(0) = 1$, from the last equation we have

$$r_Y^*(y) = \frac{v^{\log(1+\frac{\lambda}{\beta}y} - v^{\log(1+\frac{\lambda}{\beta}(y+1))}}{1 - v^{\log(1+\frac{\lambda}{\beta}(y+1))}}$$

which in view of Equation (6), implies Y has PMF in Equation (1). □

4. Different methods of estimation and testing of hypothesis

4.1. Maximum likelihood estimation

In this section, the estimation of ADP distribution parameters is carried out by using the method of the maximum likelihood (ML) based on a complete sample. Let $y_1, y_2, ...y_n$ be random sample from ADP(λ, β, v) distribution. Then log-likelihood function (L) can be formulated as

$$L = \sum_{i=1}^{n} log\left[v^{\log(1+\frac{\lambda}{\beta}y_i)} - v^{\log(1+\frac{\lambda}{\beta}(y_i+1))}\right]. \tag{21}$$

Differentiating Equation (21) with respect to λ, β and v respectively, we get the likelihood equations as follows:

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{n} \frac{\frac{\log v}{\beta}\left[y_i\left(1+\frac{\lambda}{\beta}y_i\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}y_i\right)} - (y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]}{v^{\log\left(1+\frac{\lambda}{\beta}y_i\right)} - v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}}. \tag{22}$$

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^{n} \frac{\frac{-\lambda\log v}{\beta^2}\left[y_i\left(1+\frac{\lambda}{\beta}y_i\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}y_i\right)} - (y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]}{v^{\log\left(1+\frac{\lambda}{\beta}y_i\right)} - v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}}. \tag{23}$$

$$\frac{\partial L}{\partial v} = \sum_{i=1}^{n} \frac{\log\left(1+\frac{\lambda}{\beta}y_i\right) v^{\log\left(1+\frac{\lambda}{\beta}y_i\right)-1} - \log\left(1+\frac{\lambda}{\beta}(y_i+1)\right) v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)-1}}{v^{\log\left(1+\frac{\lambda}{\beta}y_i\right)} - v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}}. \tag{24}$$

As closed-form solutions for the log-likelihood Equations (22-24) do not exist, the ML estimates can be derived through numerical methods or a direct search for the global maximum on the log-likelihood surface. The asymptotic covariance matrix of the ML estimates for parameters λ, β and v can be calculated using the inverse of the Fisher information matrix.

4.2. Least squares estimation

This method is based on the observed sample $y_1, y_2, ..., y_n$ from n ordered random sample of any distribution with CDF, where $F(.)$ denotes the CDF, we get

$$E(F(y_j)) = \frac{j}{(n+1)}.$$

The least sqaures (LS) estimators are obtained by minimizing

$$LS = \sum_{j=1}^{n}(F(y_j) - \frac{j}{n+1})^2. \tag{25}$$

Putting the CDF of ADP in Equation (25) we get

$$LS = \sum_{j=1}^{n}\left[1 - \left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right) - \frac{j}{n+1}\right]^2. \tag{26}$$

After differentiating Equation (26) with respect to the parameters λ, β and v and equating to zero, the normal equations are as follows:

$$\begin{aligned}\frac{\partial LS}{\partial \lambda} = &-2\sum_{j=1}^{n}\left[1 - \left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right) - \frac{j}{n+1}\right] \\ &\times\left[\frac{\log v}{\beta}(y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]\end{aligned} \tag{27}$$

$$\begin{aligned}\frac{\partial LS}{\partial \beta} =&2\sum_{j=1}^{n}\left[1-\left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right)-\frac{j}{n+1}\right]\\ &\times\left[\frac{\lambda\log v}{\beta^2}(y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1}v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]\end{aligned} \tag{28}$$

$$\begin{aligned}\frac{\partial LS}{\partial v} =&-2\sum_{j=1}^{n}\left[1-\left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right)-\frac{j}{n+1}\right]\\ &\times\left[\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)-1}\right]\end{aligned} \tag{29}$$

The above non-linear equations cannot be solved analytically. So the LS estimators of λ, *beta* and v can be obtained by using some iterative techniques likes Newton-Raphson method.

4.3. Weighted least squares estimation

The weighted least squares (WLS) estimators can be obtained by minimizing

$$WLS=\sum_{j=1}^{n}w_j(F(y_i)-\frac{j}{n+1})^2 \tag{30}$$

with respect to the unknown parameters, where $w_j=\frac{1}{Var(F(Y_j))}=\frac{(n+1)^2(n+2)}{j(n-j+1)}$.

Putting the CDF of ADP distribution in Equation (30), we get

$$WLS=\sum_{j=1}^{n}\left[1-\left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right)-\frac{j}{n+1}\right]^2. \tag{31}$$

After differentiating Equation (31) with respect to the parameters λ,β and v and equating to zero, the normal equations are as follows:

$$\begin{aligned}\frac{\partial WLS}{\partial \lambda} =&-2\sum_{j=1}^{n}w_j\left[1-\left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right)-\frac{j}{n+1}\right]\\ &\times\left[\frac{\log v}{\beta}(y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1}v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]\end{aligned} \tag{32}$$

$$\begin{aligned}\frac{\partial WLS}{\partial \beta} =&2\sum_{j=1}^{n}w_j\left[1-\left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right)-\frac{j}{n+1}\right]\\ &\times\left[\frac{\lambda\log v}{\beta^2}(y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1}v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]\end{aligned} \tag{33}$$

$$\begin{aligned}\frac{\partial WLS}{\partial v} =&-2\sum_{j=1}^{n}w_j\left[1-\left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right)-\frac{j}{n+1}\right]\\ &\times\left[\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)-1}\right]\end{aligned} \tag{34}$$

The above non-linear equations cannot be solved analytically. So the WLS estimators of λ, β and v can be obtained by using some iterative techniques likes Newton-Raphson method.

4.4. Cramer-von Mises estimation

The Cramer-von Mises (CVM) estimates of the parameter λ, β and v are obtained by minimizing the following expression with respect to the parameters λ, β and v respectively.

$$CVM=\frac{1}{12n}+\sum_{j=1}^{n}(F(y_j)-\frac{-1+2j}{2n})^2. \tag{35}$$

For in the case of ADP distribution, put CDF of ADP in Equation (35).

$$CVM = \frac{1}{12n} + \sum_{j=1}^{n} \left(1 - v^{\log(1+\frac{\lambda}{\beta}(y+1))} - \frac{-1+2j}{2n}\right)^2. \tag{36}$$

By differentiating Equation(36) with respect to the parameters λ, β and v and equating to zero, we get the normal equations as follows:

$$\begin{aligned}\frac{\partial CVM}{\partial \lambda} =& -2\sum_{j=1}^{n}\left[1 - \left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right) - \frac{-1+2j}{2n}\right] \\ &\times \left[\frac{\log v}{\beta}(y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]\end{aligned} \tag{37}$$

$$\begin{aligned}\frac{\partial CVM}{\partial \beta} =& 2\sum_{j=1}^{n}\left[1 - \left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right) - \frac{-1+2j}{2n}\right] \\ &\times \left[\frac{\lambda \log v}{\beta^2}(y_i+1)\left(1+\frac{\lambda}{\beta}(y_i+1)\right)^{-1} v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)}\right]\end{aligned} \tag{38}$$

$$\begin{aligned}\frac{\partial CVM}{\partial v} =& -2\sum_{j=1}^{n}\left[1 - \left(v^{\log(1+\frac{\lambda}{\beta}(y+1))}\right) - \frac{-1+2j}{2n}\right] \\ &\times \left[\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right) v^{\log\left(1+\frac{\lambda}{\beta}(y_i+1)\right)-1}\right]\end{aligned} \tag{39}$$

These Equations (37–39) cannot be solved analytically. The estimates of λ, β and v can be obtained by setting the normal equations equations equal to zero and solving simultaneously with the help of statistical packages like *optim* or *nlm* in **R** programming.

4.5. Testing of hypothesis

Here we present a procedure-the generalized likelihood ratio test for testing the significance of the additional parameter β of the ADP(λ, β, v), since when $\beta = 1$, ADP reduces to Discrete Lomax distribution. Thus, we consider the null hypothesis as $H_0 : \beta = 1$ against the alternative $H_0 : \beta \neq 1$.
The test statistic for generalized likelihood ratio test is,

$$-2log\omega = 2[L(y;\hat{\eta}) - L(y;\hat{\eta}^*)] \tag{40}$$

where $\hat{\eta}$ is the ML estimator of $\eta = (\lambda, \beta, v)$ and $\hat{\eta}^*$ is ML estimator of $\beta = 1$. The test statistic $-2log\omega$ given in Equation (40) is asymptotically distributed as chi-square with one degree of freedom.

5. Simulation study

In this section, we evaluate the performance of the estimators by employing a simulation technique across various sample sizes (n). The simulation is conducted using the optim() function in R Core Team (2023) software. We generated 2,000 samples of three different sizes (n = 50, 100, 250) from ADP(λ, β, v) distribution, considering different values of model parameters using different methods of estimation outlined in Table 2 and Table 3. The average values of estimates, bias and MSEs of ML, LS, WLS and CVM are obtained and displayed in Table 2 and Table 3. Initial values are chosen to compute the estimates in such way that the optimization function have minimum bias. The steps are:
1. Compute the estimates for the 2000 samples, say $\hat{\eta}$ for j = 1, 2, ..., 2000.

2.Compute Bias by using the formula, $Bias(\hat{\eta}) = \frac{1}{2000}\sum_{j=1}^{2000}(\hat{\eta} - \eta)$.
3. Compute MSE by using the formula, $MSE(\hat{\eta}) = \frac{1}{2000}\sum_{j=1}^{2000}(\hat{\eta} - \eta)^2$.
Also, graphs for MSE for simulation results obtained in Table 2 and Table 3 are illustrated in Figure 3 and Figure 4.
Notably, as the sample size (n) increases, there is a consistent decrease in bias and MSE for different methods of estimations. From Table 2 and Table 3, we observe that all the estimates show the property of consistency i.e., the MSEs decrease as sample size increase. Comparing the different methods of estimation, the results show that the ML produces the best results for estimating the parameters λ, β and v in terms of MSEs in all of the cases.

Table 2: Simulation results of $\lambda = 0.5$, $\beta = 0.5$ and $v = 0.5$

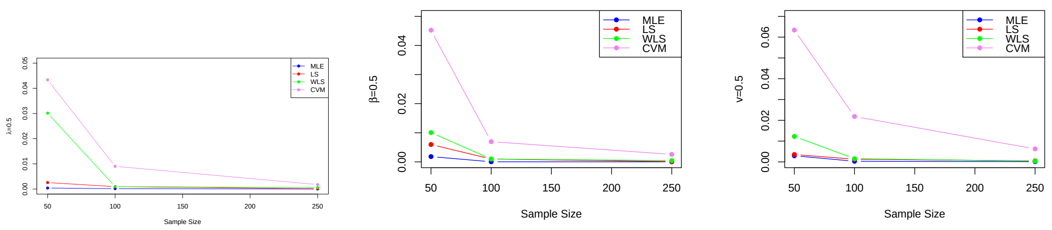
Sample size(n)	Parameter↓	ML	LS	WLS	CVM
50	$\hat{\lambda}$	0.52037	0.44922	0.32636	0.70819
	Bias	0.02037	0.05077	0.17363	-0.20818
	MSE	0.00041	0.00258	0.03014	0.04334
	$\hat{\beta}$	0.48271	0.45732	0.39961	0.71271
	Bias	0.01773	0.04268	0.10038	-0.21271
	MSE	0.00182	0.00597	0.01007	0.04524
	$\hat{v}$	0.55431	0.55953	0.38934	0.75185
	Bias	0.05430	-0.05952	0.11065	-0.25185
	MSE	0.00295	0.00354	0.01224	0.06342
100	$\hat{\lambda}$	0.48769	0.53079	0.53109	0.40496
	Bias	0.01230	-0.03079	-0.03109	0.09503
	MSE	0.00015	0.00095	0.00097	0.00903
	$\hat{\beta}$	0.50196	0.53122	0.53112	0.47658
	Bias	-0.00196	-0.03122	-0.03119	0.08341
	MSE	3.867×10^{-5}	0.00097	0.00097	0.00695
	$\hat{v}$	0.51736	0.53589	0.533938	0.64773
	Bias	-0.01736	-0.03589	-0.03986	-0.01477
	MSE	0.00030	0.00128	0.00155	0.02183
250	$\hat{\lambda}$	0.49770	0.48860	0.47600	0.45795
	Bias	0.00229	0.01139	0.02399	0.04205
	MSE	5.2754×10^{-6}	0.00013	0.00057	0.00177
	$\hat{\beta}$	0.50431	0.48958	0.47889	0.55093
	Bias	-0.00431	0.01042	0.0211	-0.05093
	MSE	1.8535×10^{-5}	0.00011	0.00044	0.00259
	$\hat{v}$	0.50926	0.47798	0.51817	0.57900
	Bias	-0.00926	0.02201	0.01817	-0.07900
	MSE	8.5833×10^{-5}	0.00048	0.00033	0.00624

6. Applications

For application, we have considered four real data sets. The parameters are estimated by using the method of maximum likelihood (using R software). We compare the fit of the ADP distribution with Discrete Pareto (DP)(see Krishna and Pundir (2009)), Discrete Lomax (DL)(see Prieto *et al.* (2014)) , Discrete Pareto type (IV) (DP IV) (see Ghosh (2020)), Discrete Burr XII (DB XII) (see Para and Jan (2016)) and Discrete exponentiated Weibull (EDW)(see Nekoukhou and Bidram (2015)) distributions. The values of estimates, the log-

Table 3: Simulation results of $\lambda = 0.2$, $\beta = 0.9$ and $\upsilon = 0.5$

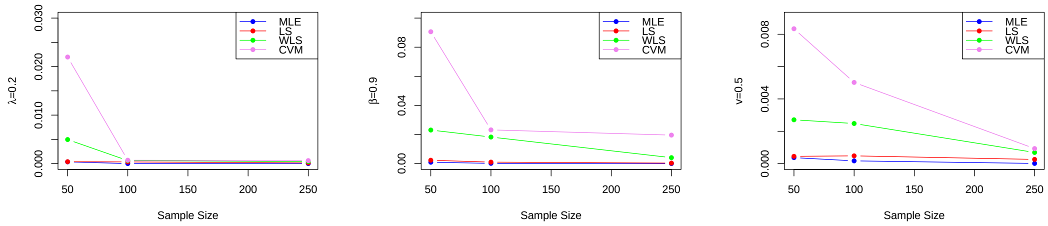
Sample size(n)	Parameter↓	ML	LS	WLS	CVM
50	$\hat{\lambda}$	0.18111	0.18001	0.12951	0.34823
	Bias	0.01889	0.01999	0.07048	-0.14823
	MSE	0.00036	0.00039	0.00496	0.02197
	$\hat{\beta}$	0.93075	0.85183	0.74811	1.20102
	Bias	-0.03075	0.04816	0.15189	-0.30106
	MSE	0.00095	0.00232	0.02307	0.09061
	$\hat{\upsilon}$	0.48072	0.52129	0.55207	0.59135
	Bias	0.01928	-0.02129	-0.05207	-0.09135
	MSE	0.00037	0.00045	0.00271	0.00834
100	$\hat{\lambda}$	0.20767	0.18115	0.22510	0.22649
	Bias	-0.00765	0.01885	-0.02510	-0.02649
	MSE	5.8760×10^{-6}	0.00035	0.00063	0.00070
	$\hat{\beta}$	0.91302	0.93195	0.76488	1.05226
	Bias	-0.01302	-0.03195	0.13512	-0.15226
	MSE	0.00017	0.00102	0.01826	0.02318
	$\hat{\upsilon}$	0.51334	0.52209	0.54984	0.42914
	Bias	-0.01334	-0.02209	-0.04984	-0.07085
	MSE	0.00017	0.00048	0.00248	0.00502
250	$\hat{\lambda}$	0.20544	0.18820	0.17961	0.22491
	Bias	-0.00544	0.01179	0.02038	-0.02491
	MSE	2.9615×10^{-5}	0.00014	0.00042	0.00062
	$\hat{\beta}$	0.90342	0.88095	0.83621	0.76002
	Bias	-0.00342	0.01904	0.06397	0.13997
	MSE	1.1758×10^{-5}	0.00036	0.00409	0.01959
	$\hat{\upsilon}$	0.50413	0.51610	0.47365	0.46949
	Bias	-0.00413	0.01610	0.02635	0.03051
	MSE	1.7076×10^{-5}	0.00026	0.00069	0.00093



(a) (b) (c)

Figure 3: MSE graphs for $\lambda = 0.5$, $\beta = 0.5$ and $\upsilon = 0.5$

likelihood function (-LogL), the Kolmogrov-Smirnov (K-S) statistic, Akaike Information Criterion (AIC), Akaike Information Criterion with correction (AICc), Bayesian Information Criterion (BIC) and Hannon-Quinn Information Criterion (HQIC) are calculated for the six distributions in order to verify which distribution fits better to the data.



(a) (b) (c)

Figure 4: MSE graphs for $\lambda = 0.2$, $\beta = 0.9$ and $\upsilon = 0.5$

Data set I: *Medical data*

A data set is taken from Ramana, Babu, and Venkateswarlu (2012), which is a critical comparative study of liver patients from USA and India. This data set contains records of 583 patients, and there are 10 variables per patient: age, gender, total bilirubin, direct bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and alkphos. Here, we have taken the data of total bilirubin content in patients. Since the data set is continuous, we have taken the integer part of observations. The summary statistics of data set I is given below.

Mean	Variance	Skeweness	Kurtosis
2.722	39.239	4.853	39.497

Data set II: *Biological data*

The data set includes 171 molecules designed for functional domains of a core clock protein, CRY1, responsible for generating circadian rhythm. 56 of the molecules are toxic and the rest are non-toxic. Here we used period lengthening dataset called "nHBin" of 10-Fold cross validation accuracies, (for more details of data see Gul, Rahim, Isin, Yilmaz, Ozturk, Turkay, and Kavakli (2021)). The summary statistics of data set II is given below.

Mean	Variance	Skeweness	Kurtosis
0.316	0.582	2.828	11.566

Data set III: *Reliability data*

The data used is failure time data of an air conditioning system on an airplane sourced from Linhart and Zucchini (1986).The mean value of the data is 53.7407, which is smaller than the variance, 4087.7380 this indicates over dispersion. Also, the skewness is 1.6632 and kurtosis is 2.1453.

Data set IV: *Epidemological data*

This data represents the daily new cases of COVID-19 in Egypt. The data is available at https://covid19.who.int/ and contains the daily new cases between 15 March and 10 June 2020. The data set along with R codes for computations are given in the Appendix.

From Table 4, Table 5, Table 6 and Table 7, it is clear that ADP distribution provides a good fit to these data sets. For the data set I ADP distribution has highest p value when compared with other five models and lowest -LogL, AIC, BIC, AICc, HQIC and K-S statistic values.

Table 4: Parameter estimates and goodness of fit for various models fitted to the dataset I.

Model	MLEs	-LogL	AIC	BIC	AICc	HQIC	K-S	p-value
DP	$\hat{p} = 0.013$	1277.736	2557.472	2561.840	2557.479	2559.175	0.761	0.00
DL	$\hat{\alpha} = 0.107$	1229.895	2463.790	2472.526	2463.811	2467.195	0.312	0.067
	$\hat{\lambda} = 0.341$							
DP IV	$\hat{\theta}$=0.765	1690.572	3387.144	3400.249	3387.185	3392.252	0.594	0.01
	$\hat{\sigma} = 1.085$							
	$\hat{\gamma} = 0.094$							
DB XII	$\hat{\beta}$ =0.365	1501.676	3009.352	3022.457	3009.393	3014.460	0.499	0.07
	$\hat{\gamma}$ =1.083							
	$\hat{c}$ =0.185							
EDW	$\hat{p}$= 0.917	1333.916	2673.832	2686.937	2673.873	2678.940	0.648	<0.00
	$\hat{\alpha}$ =0.548							
	$\hat{\gamma}$ =0.175							
ADP	$\hat{\lambda}$ =0.103	1221.147	2448.294	2461.399	2448.335	2453.402	0.292	0.431
	$\hat{\beta} = 0.158$							
	$\hat{v}$= 0.502							

Also, in case of data set II, data set III and data set IV, ADP distribution has highest p value when compared to others and lowest -LogL, AIC, BIC, AICc, HQIC and K-S statistic values.

For testing, $H_0 : \beta = 1$ against $H_0 : \beta \neq 1$ by generalized likelihood ratio test procedure we have computed the values of $L(\hat{\eta}; y)$ and $L(\hat{\eta}^*; y)$ as -1221.147 and -1229.895 respectively for data set I. The test statistic value is computed as 8.748. Since the critical value for the test at 5% level of significance is 3.84 at one degree of freedom, the null hypothesis is rejected in this case.
For data set II, the values of $L(\hat{\eta}; y)$ and $L(\hat{\eta}^*; y)$ are - 123.2164 and -125.7054 respectively. The test statistic value is 4.978. The null hypothesis is rejected since critical value 3.84 at 0.05 level of significance is less than the calculated value.
For data set III, the values of $L(\hat{\eta}; y)$ and $L(\hat{\eta}^*; y)$ are - 134.5773 and -136.5341 respectively. The test statistic value is 3.914. Thus the null hypothesis is rejected at 5% level of significance since the critical value is 3.84 at one degree of freedom.

7. Discussion with other method

Cordeiro, Ortega, Popović, and Pescim (2014) proposed a class of distributions called Lomax generator with two positive parameters to generalize any continuous baseline distributions.

Table 5: Parameter estimates and goodness of fit for various models fitted to the dataset II.

Model	MLEs	-LogL	AIC	BIC	AICc	HQIC	K-S	p-value
DP	$\hat{p}$ = 0.1757	128.9439	259.8878	263.0295	259.9115	261.1626	0.7004	0.010
DL	$\hat{\alpha}$ = 0.1154	125.7054	255.4108	261.6941	255.4822	257.9603	0.5446	0.120
	$\hat{\lambda}$ = 0.2018							
DP IV	$\hat{\theta}$=0.1324	123.9902	253.9804	263.4054	254.1241	257.8047	0.7367	0.074
	$\hat{\sigma}$ = 1.0512							
	$\hat{\gamma}$ = 0.7410							
DB XII	$\hat{\beta}$ =0.0723	123.7497	253.4994	262.9244	253.6431	257.3237	0.8312	0.02
	$\hat{\gamma}$ =1.0471							
	$\hat{c}$ =0.7104							
EDW	$\hat{p}$= 0.8577	128.8873	263.7746	273.1996	263.9183	267.5989	0.8715	<0.00
	$\hat{\alpha}$ =0.6497							
	$\hat{\gamma}$ =0.0705							
ADP	$\hat{\lambda}$ =0.1539	123.2164	252.4328	261.8578	252.5765	256.2571	0.4536	0.287
	$\hat{\beta}$ = 0.4551							
	$\hat{v}$= 0.1256							

The Probability Density Function (PDF) and its SF are

$$f(x) = \theta\mu^{\theta}\frac{g(x)}{(1-G(x))\{\mu-\log(1-G(x))\}^{\theta+1}}; x \geqslant 0 \tag{41}$$

and

$$S(x) = \left[\frac{\mu}{\mu-\log(1-G(x))}\right]^{\theta}; x > 0, \theta, \mu > 0 \tag{42}$$

respectively, where $G(x)$ is the CDF of the baseline model.
The Lomax generator family of distributions can generalize all classical continuous distributions. The density funtion of Lomax generator allows greater flexibility of tails and can be widely applied in many areas of engineering and biology. Also the moments of this family play an important role for measuring inequality, for example, income quantiles and Lorenz and Bonferroni curves, which depend upon the incomplete moments of a distribution.

Here we consider the Survival Discretization (SD) method. Let X be a continuous rv. Then the discerte analogue Y of X can be derived by using the SF as follows: Let $S(.)$ be the SF of the rv X. According to the SD approach, the PMF can be expressed as

$$P(Y = y) = S(y) - S(y+1); \quad y = 0, 1, 2, 3, \ldots. \tag{43}$$

Table 6: Parameter estimates and goodness of fit for various models fitted to the dataset III.

Model	MLEs	-LogL	AIC	BIC	AICc	HQIC	K-S	p-value
DP	$\hat{p} = 0.7444$	151.4199	304.8398	306.1356	304.9998	305.2251	0.3829	0.0007
DL	$\hat{\alpha} = 0.1655$	136.5341	277.0682	279.6599	277.5682	277.8388	0.2511	0.0665
	$\hat{\lambda} = 0.0273$							
DP IV	$\hat{\theta}$=0.0445	155.3150	316.6300	320.5175	317.6735	317.7890	0.8520	0.0000
	$\hat{\sigma} = 1.1709$							
	$\hat{\gamma} = 0.9545$							
DB XII	$\hat{\beta}$ =0.4065	167.6467	341.2934	345.1809	342.3369	342.4494	0.5671	0.0000
	$\hat{\gamma}$ =0.5732							
	$\hat{c}$ =0.3025							
EDW	$\hat{p}$= 0.3410	154.5298	315.0596	318.9471	316.1031	316.2156	0.3361	0.0045
	$\hat{\alpha}$ =0.1478							
	$\hat{\gamma}$ =3.1741							
ADP	$\hat{\lambda}$ =0.1100	134.5773	275.1466	279.0341	276.1901	276.3026	0.1197	0.8338
	$\hat{\beta}$ = 0.0016							
	$\hat{v}$= 0.1250							

Employing SD approach on Equation (41), the PMF of discrete Lomax generator family can be obtained as,

$$f_Y(y) = \left[\frac{\mu}{\mu - \log(1 - G(y))}\right]^{\theta} - \left[\frac{\mu}{\mu - \log(1 - G(y+1)))}\right]^{\theta}; \quad y \in \mathbf{N_0} \tag{44}$$

where $\mu, \theta \in (0, \infty)$ and $\mathbf{N_0} = \{0, 1, 2, \ldots\}$.
By putting CDF of geometric distribution, $G(y) = 1 - \delta^y;\ 0 < \delta < 1,\ y \in \mathbf{N_0}$, in Equation (44), we obtain

$$\begin{aligned}\left[\frac{\mu}{\mu-\log(1-(1-\delta^y))}\right]^{\theta} - \left[\frac{\mu}{\mu-\log(1-(1-\delta^{y+1}))}\right]^{\theta} &= \left[\frac{\mu}{\mu-y\log\delta}\right]^{\theta} - \left[\frac{\mu}{\mu-(y+1)\log\delta}\right]^{\theta} \\ &= \left[1 - \frac{\log\delta}{\mu}y\right]^{-\theta} - \left[1 - \frac{\log\delta(y+1)}{\mu}\right]^{-\theta}.\end{aligned}$$

By reparameterization, $\delta = e^{-\gamma}$, we get

$$f_Y(y) = \left[1 + \frac{\gamma}{\mu}y\right]^{-\theta} - \left[1 + \frac{\gamma}{\mu}(y+1)\right]^{-\theta};\ y \in \mathbf{N_0}.$$

This is same as the PMF of ADP distribution.

8. Conclusions and future works

Here we introduce a new discrete distribution which is an alternative to discrete Pareto distribution. Various properties of ADP distribution are studied. This distribution is a competing

Table 7: Parameter estimates and goodness of fit for various models fitted to the dataset IV.

Model	MLEs	-LogL	AIC	BIC	AICc	HQIC	K-S	p-value
DP	$\hat{p}$ = 0.8205	675.674	1353.347	1355.824	1353.394	1354.345	0.368	0.000
DL	$\hat{\alpha}$ = 0.9983	619.267	1242.533	1247.488	1242.675	1244.530	0.162	0.019
	$\hat{\lambda}$ = 1.3725							
DP IV	$\hat{\theta}$=0.9980	685.136	1376.272	1383.704	1376.558	1379.266	0.873	0.000
	$\hat{\sigma}$ = 0.5030							
	$\hat{\gamma}$ = 0.0113							
DB XII	$\hat{\beta}$ =0.6360	657.638	1321.277	1328.709	1321.563	1324.271	0.362	0.000
	$\hat{\gamma}$ =5.5837							
	$\hat{c}$ =0.7216							
ADP	$\hat{\lambda}$ =0.0620	617.399	1240.799	1248.231	1241.085	1243.794	0.099	0.350
	$\hat{\beta}$ = 0.0003							
	$\hat{v}$= 0.2168							

model when compared with other existing discrete Pareto models. The proposed model offers the ability to model a hazard rate function that exhibits a decreasing failure rate patterns and is suitable for modeling with overdispersed datasets. From the application of real datasets, it is clear that ADP distribution may serve as a viable alternative to discrete Pareto models available in the literature for modeling count data arising in various fields of scientific investigation such as reliability theory, hydrology, medicine, meteorology, survival analysis and engineering. This paper identifies several promising paths for future investigation, offering substantial potential to enhance comprehension and utilization of the proposed model across diverse domains. These avenues include extending the analysis to encompass bivariate scenarios, investigating alternative censored methods, embracing Bayesian techniques for parameter estimation, exploring various loss functions, and exploring the application of neutrosophic statistics. Also, considering that the hazard rate function of the ADP distribution decreased, for modeling real-world scenarios with increasing and bathtub-shaped hazard rates, we were unable to apply this distribution. This facilitates the development of better adapted versions of discrete Pareto models.

Acknowledgments

Authors wish to thank the Editor and the reviewers for their comments and suggestions which significantly helped in improving the presentation of the article.

References

Akinsete A, Famoye F, Lee C (2008). "The Beta-Pareto dDistribution." *Statistics*, **42**(6), 547–563.

Arnold BC (2008). "Pareto and Generalized Pareto Distributions." In D Chotikapanich (ed.), *Modeling Income Distributions and Lorenz Curves*, pp. 119–145. Springer.

Bhati D, Bakouch HS (2019). "A New Infinitely Divisible Discrete Distribution with Applications to Count Data Modeling." *Communications in Statistics-Theory and Methods*, **48**(6), 1401–1416.

Buddana A, Kozubowski TJ (2014). "Discrete Pareto Distributions." *Economic Quality Control*, **29**(2), 143–156.

Cifarelli DM, Gupta RP, Jayakumar K (2010). "On Generalized Semi-Pareto and Semi-Burr Distributions and Random Coefficient Minification Processes." *Statistical Papers*, **51**, 193–208.

Cordeiro GM, Ortega EM, Popović BV, Pescim RR (2014). "The Lomax Generator of Distributions: Properties, Minification Process and Regression Model." *Applied Mathematics and Computation*, **247**, 465–486.

Ghosh I (2020). "A New Discrete Pareto Type (IV) Model: Theory, Properties and Applications." *Journal of Statistical Distributions and Applications*, **7**(1), 1–17.

Gul S, Rahim F, Isin S, Yilmaz F, Ozturk N, Turkay M, Kavakli IH (2021). "Structure-based Design and Classifications of Small Molecules Regulating the Circadian Rhythm Period." *Scientific Reports*, **11**(1), 1–15.

Jayakumar K, Jose J (2022). "Discrete New Generalized Pareto Distribution." *Statistica*, **82**(4), 373–391.

Jayakumar K, Krishnan B, Hamedani GG (2020). "On a New Generalization of Pareto Distribution and Its Applications." *Communications in Statistics-Simulation and Computation*, **49**(5), 1264–1284.

Jayakumar K, Mathew T (2008). "On a Generalization to Marshall-Olkin Scheme and Its Application to Burr Type XII Distribution." *Statistical Papers*, **49**(3), 421–439.

Johnson NL, Kotz S (1982). "Developments in Discrete Distributions, 1969-1980." *International Statistical Review/Revue Internationale de Statistique*, **50**(1), 71–101.

Karlis D, Xekalaki E (2005). "Mixed Poisson Distributions." *International Statistical Review/Revue Internationale de Statistique*, **73**(1), 35–58.

Kemp AW (2004). "Classes of Discrete Lifetime Distributions." *Communications in Statistics - Theory and Methods*, **33**(12), 3069–3093.

Krishna H, Pundir PS (2009). "Discrete Burr and Discrete Pareto Distributions." *Statistical Methodology*, **6**(2), 177–188.

Linhart H, Zucchini W (1986). *Model Selection.* John Wiley & Sons.

Marshall AW, Proschan F (1970). "Mean Life of Series and Parallel Systems." *Journal of Applied Probability*, **7**(1), 165–174.

Nekoukhou V, Bidram H (2015). "The Exponentiated Discrete Weibull Distribution." *Sort*, **39**, 127–146.

Para BA, Jan TR (2016). "On Discrete Three Parameter Burr Type XII And Discrete Lomax Distributions and Their Applications to Model Count Data from Medical Science." *Biometrics & Biostatistics International Journal*, **4**(2), 1–15.

Pickands III J (1975). "Statistical Inference Using Extreme Order Statistics." *The Annals of Statistics*, **3**(1), 119–131.

Prieto F, Gómez-Déniz E, Sarabia JM (2014). "Modelling Road Accident Blackspots Data with the Discrete Generalized Pareto Distribution." *Accident Analysis & Prevention*, **71**, 38–49.

R Core Team R (2023). "R: A Language and Environment for Statistical Computing." *R Foundation for Statistical Computing, Vienna, Austria.*

Ramana BV, Babu MSP, Venkateswarlu N (2012). "A Critical Comparative Study of Liver Patients from USA and INDIA: An Exploratory Analysis." *International Journal of Computer Science Issues*, **9**(3), 506.

Rényi A (1961). "On Measures of Entropy and Information." In J Neyman (ed.), *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 547–562. Berkely.

Steutel FW, van Harn K (2003). *Infinite Divisibility of Probability Distributions on the Real Line.* CRC Press.

Stoppa G (1990). "A New Model for Income Size Distributions." In C Dagum, M Zenga (eds.), *Income and Wealth Distribution, Inequality and Poverty. Studies in Contemporary Economics*, pp. 33–41. Springer.

Tahir MH, Cordeiro GM, Mansoor M, Zubair M, Alzaatreh A (2021). "The Kumaraswamy Pareto IV Distribution." *Austrian Journal of Statistics*, **50**(5), 1–22.

Xie M, Gaudoin O, Bracquemond C (2002). "Redefining Failure Rate Function for Discrete Distributions." *International Journal of Reliability, Quality and Safety Engineering*, **9**(3), 275–285.

Appendix

R code for the data set IV.

```
#Data set
> y<-c(17, 16, 40, 30, 14, 46, 29, 9, 33, 39, 76,14, 39, 41, 40,
33, 47, 54, 69, 86, 120, 85, 103, 149, 128, 110,139, 95, 145,
126, 125, 160, 155, 168, 171, 188, 112, 189,157, 169, 433, 0,
227,463, 260, 226,269, 358, 0, 298, 272,736, 387, 0, 393, 495,
924, 346, 0, 685, 398, 0, 399, 491,510,535, 1465, 0, 774, 783,
727, 752, 702, 789,910, 1127, 1289,1367, 1536, 1399, 1152,
1079,1152, 1348, 1497, 1467, 1365,1385)
> y<-sort(y)
> n<-length(y)
> hist(y)
#ADP distribution
   f<-function(par){
  -sum(log(par[1]^(log(1+(par[2]/par[3])*y))
  -par[1]^(log(1+(par[2]/par[3])*(y+1)))))
}
   out<-nlm(f,p=c(0.12, .1, .1),hessian=T)

   pdlog<-function(y,v,lambda,beta){
  1-v^(log(1+((lambda/beta)*(y+1))))
}
   ks.test(y,"pdlog", 0.0620280034 ,0.0002732876, 0.2168630137)
   L=617.399
   k<-3
```

```
    AIC<-(2*L)+(2*k)
    BIC<-(2*L)+(k*log(n))
    AICC<-(2*L)+((2*k*n)/(n-k-1))
    HQIC<-(2*L)+(2*k*log(log(n)))
#DP distribution
f<-function(par){
  -sum(log(par[1]^(log(1+y))-par[1]^(log(y+2))))
}
out<-nlm(f,p=c(0.75),hessian=T)
pdlog<-function(y,v){
  1-v^(log(y+2))
}
ks.test(y,"pdlog", 0.820534)
L=675.6735
k<-1
AIC<-(2*L)+(2*k)
BIC<-(2*L)+(k*log(n))
AICC<-(2*L)+((2*k*n)/(n-k-1))
HQIC<-(2*L)+(2*k*log(log(n)))
#DB distribution
f<-function(p){
  -sum( log((1+((y/p[2])^p[3])))*log(p[1]) +
  log((1-(p[1]^(log((1+(((y+1)/p[2])^p[3]))/(1+((y/p[2])^p[3]))))))) )
}
out=optim(p=c(.001,0.01,0.004),f)
pdb3<-function(y,beta,gamma,c){
  1-(beta^log( (1+(((y+1)/gamma)^c))))
}
ks.test(y,"pdb3", 0.6360314, 5.5837399, 0.7216526)
L=657.6385
k<-3
AIC<-(2*L)+(2*k)
BIC<-(2*L)+(k*log(n))
AICC<-(2*L)+((2*k*n)/(n-k-1))
HQIC<-(2*L)+(2*k*log(log(n)))
f<-function(p){
  -sum(log( p[1]^(1+p[2]*y)-p[1]^(1+p[2]*(y+1))))
}
out=optim(p=c(0.1,0.2),f)
pdlog<-function(y,alpha,lambda){
  1-alpha^(1+lambda*(y+1))
}
ks.test(y,"pdlog", 0.9982677, 1.3725924)
L=619.267
k<-2
AIC<-(2*L)+(2*k)
AIC
BIC<-(2*L)+(k*log(n))
BIC
AICC<-(2*L)+((2*k*n)/(n-k-1))
AICC
HQIC<-(2*L)+(2*k*log(log(n)))
HQIC
#DP type IV distribution
```

```
f<-function(p){
  -sum(log(p[1]^(log(1+(y/p[2])^(1/p[3]))) -
  p[1]^(log(1+((y+1)/p[2])^(1/p[3])))))
}
out=optim(p=c(0.5,0.5,0.1),f)
pdpIV<-function(y,theta,sigma,gamma){
  1- theta^(log(1+((y+1)/sigma^(1/gamma))))
}

ks.test(y,"pdpIV",0.99801714, 0.50304968 ,0.01130628)
L= 685.1361
k<-3
AIC<-(2*L)+(2*k)
BIC<-(2*L)+(k*log(n))
AICC<-(2*L)+((2*k*n)/(n-k-1))
HQIC<-(2*L)+(2*k*log(log(n)))
```

Affiliation:

K. Jayakumar
Department of Statistics
University of Calicut
Malappuram, Kerala-673635, India
E-mail: jkumarguruvayoor@gmail.com

Jiji Jose
Department of Statistics
University of Calicut
Malappuram, Kerala-673635, India
E-mail: jijijose2793@gmail.com

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 54 *Submitted:* 2024-08-04
2025 *Accepted:* 2025-03-03