

RESEARCH

Open Access

A new discrete pareto type (IV) model: theory, properties and applications



Indranil Ghosh

Correspondence:
ghoshi@uncw.edu
University of North Carolina,
Wilmington, USA

Abstract

Discrete analogue of a continuous distribution (especially in the univariate domain) is not new in the literature. The work of discretizing continuous distributions begun with the paper by Nakagawa and Osaki (1975) to the best of the knowledge of the author. Since then several authors proposed discrete analogues of known continuous models. In this paper, we propose and study a discrete analogue of the continuous Pareto (type IV) distribution, namely the discrete Pareto (type IV) distribution (DP-IV, henceforth, in short) that has three parameters. Its probability mass function can be approximately symmetric, right-skewed and left-skewed shapes, and the hazard rate function possesses decreasing and upside-down bathtub shapes. Also, the proposed discrete distribution can be under-, over- or equi- dispersion. The flexibility of the new discrete model is illustrated by means of three applications to real life data sets arising out of various domains affecting our life.

Mathematics Subject Classification (2010): 60E05; 62F10; 62N05

Introduction

The discrete distributions are useful when count phenomenon occurs. The discrete models are as important as the continuous models. Nowadays, both types of models can be used in fascinating ways to explore real life data sets available in different fields of studies. One convincing way is the compounding technique, where the discrete and continuous models are mixed together for better exploration of phenomena under study. The other interesting technique began with the work of (Nakagawa and Osaki 1975), who first introduced the concept of discretizing a continuous model into discrete one. There are many situations where it is inappropriate to describe the lifetime of devices on a continuous scale. For example, a piece of equipment operates in cycles and experimenter observes that the number of cycles successfully completed prior to the failure. In such case, the time to failure is more appropriately represented by the number of times they are used before they fail, which is a discrete random variable. Salvia and Bollinger (1982), and Padgett and Spurrier (1985) discussed discrete hazard rate functions (h.r.f.) and failure rate models by giving illustrations to such situation. Xie et al. (2002) also defined another discrete h.r.f. $h(k) = \log[S(k-1)/S(k)]$ for a random variable K , where $S(\cdot)$ is the reliability function, which gives similar results to those for continuous h.r.f.s.



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

(Bracquemond and Gaudoin 2003) presented a survey on discrete lifetime distributions and suggested two ways by which discrete distributions can be derived from the continuous ones; (i) to consider a characteristic property of a continuous distribution and to build the similar property in discrete time, and (ii) to consider discrete lifetime as the integer part of continuous lifetime. Lai (2013) also appreciated the discretization of a continuous model but stressed that a continuous lifetime random variable may be characterized either by its cumulative distribution function (c.d.f.), probability density function (p.d.f.) or h.r.f. which are equivalent in the sense that one can be uniquely determined by the other.

In reliability theory, classification of lifetime models is defined in terms of their survival function (s.f.) and other reliability characteristics. For example, the increasing (decreasing) failure rate IFR (DFR) class, increasing (decreasing) failure rate average IFRA (DFRA) class, the new better (worse) than used NBU (NWU) class, new better (worse) than used in expectation NBUE (NWUE) class and increasing (decreasing) mean residual lifetime IMRL (DMRL) class etc. (See Kemp 2004) and the references cited therein). The discretization of a continuous lifetime distribution retains the same functional form of the survival function, therefore, many reliability characteristics and properties shall remain unchanged. Thus, discretization of a continuous lifetime model is an interesting and simple approach to derive a discrete lifetime model corresponding to the continuous one.

In the last two or three decades, there has been a growing interest in introducing discretized continuous distributions. For more details the reader is referred to (Chakraborty 2015) and the references cited therein. In the literature, there are several different methods of discretizing a continuous probability model. We will consider the approach which is due to (Nakagawa and Osaki 1975) and (Roy 2003). This is a rather more general approach for discretizing continuous models, adopted in this paper. The following proposition may lead to this approach.

Proposition 1 *Given a continuous random variable X with survival function $S_X(x)$, a discrete random variable Y can be defined as $Y = \lfloor x \rfloor$, where $\lfloor x \rfloor = \max\{m \in \mathbb{Z} \mid m \leq x\}$, the floor function. The probability mass function (p.m.f.) $P(Y = y)$ of Y is then given by*

$$\begin{aligned} P(Y = y) &= P(y \leq X < y + 1) \\ &= P(X \geq y) - P(X \geq y + 1) \\ &= S_X(y) - S_X(y + 1), \end{aligned} \tag{1}$$

where $y \in \mathbb{Z}$, and \mathbb{Z} is the set of integers. Consequently, a continuous failure-time model can be used to generate a discrete time model by introducing a grouping on the time axis. To put it in a simple way, if X is a continuous random variable, then the p.m.f. of its integer part, that is $T = dX = \lfloor X \rfloor$, can be viewed as a discrete concentration of the p.d.f. of X . Such discretized distributions retain the same functional form of the survival function as that of the continuous ones and the reliability characteristics also do not change.

The p.d.f. a continuous Pareto (Type IV) distribution (with the location parameter 0, inequality parameter γ , and the shape parameter α , also known the tail index, for details, see Arnold (1983)) is given by

$$f(x) = \frac{\alpha}{\gamma} \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}-1} \left[1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right]^{-\alpha-1}, x \geq 0.$$

The associated c.d.f. and the survival function are, respectively,

$$F(x) = 1 - \left(1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right)^{-\alpha},$$

$$S(x) = \left(1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}. \quad (2)$$

Using Eq. (1), the discrete Pareto (IV) (DPIV, henceforth, in short) distribution can be defined as

$$g(x) = \theta^{\log\left(1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right)} - \theta^{\log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)}; \quad (3)$$

$x \in \mathbb{N}^*$; where $\mathbb{N}^* = \mathbb{N} \cup \{0\}$, $\theta = \exp(-\alpha)$, and $0 < \theta < 1$, $\sigma > 0$. A random variable X with the p.m.f. as given in Eq. (3) will be said to have the DPIV distribution.

From (3), the c.d.f. and the survival function of a random variable that follows the DPIV distribution are given as follows

$$G(x) = 1 - \theta^{\log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \quad x = 1, 2, \dots, \quad (4)$$

$$S(x) = \theta^{\log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \quad x = 1, 2, \dots. \quad (5)$$

The next result discusses the limiting behavior of the DPIV distribution corresponding to various parameter choices at the boundary.

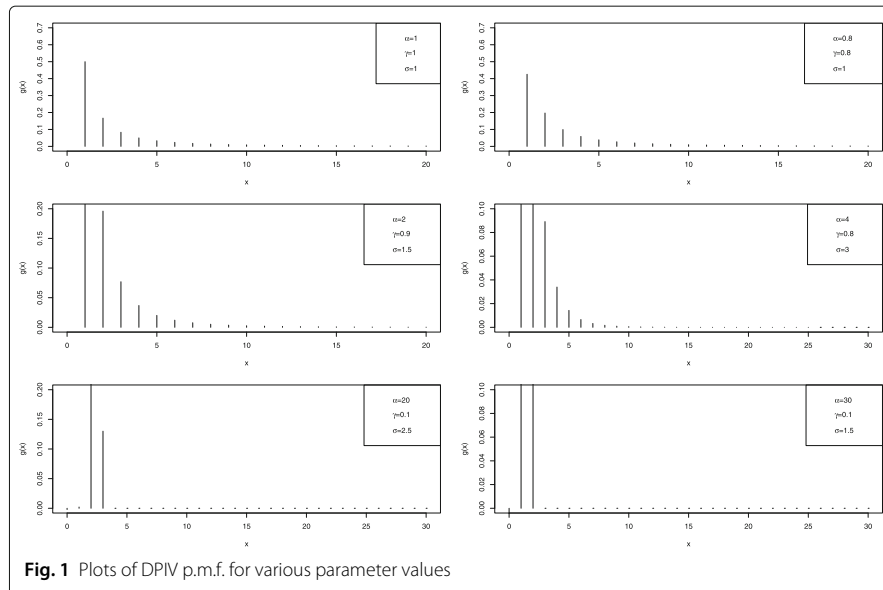
Result 1

- $\lim_{x \rightarrow \infty} g(x) = 0$.
- $\lim_{\sigma \rightarrow 0, \infty} g(x) = 0$.
- $\lim_{\gamma \rightarrow 0, \infty} g(x) = \theta^{\log\left(\frac{x}{\sigma} + 1\right) + 1} - \theta^{\log\left(\frac{x+1}{\sigma} + 1\right) + 1}$.
- $\lim_{\theta \rightarrow 0, 1} g(x) = 0$.

Some representative plots of the DPIV p.m.f. are provided in Fig. 1, for varying parameter values.

From the plots in Fig. 1, it appears that the distribution is right skewed. If we wanted to apply this to some real life application, we would desire the data to also possess this right skewed characteristic for a better model fit. It is also important to note that the Pareto (Type IV) distribution is highly sensitive to changes in the α parameter as this is the shape parameter (also known as the tail index). Furthermore, from Fig. 1, it represents the fact that for larger values of the parameter α , and γ the mode moves to the right, indicating that the proposed distribution is quite versatile in nature; while smaller values of α appear to have a significant effect on the respective probabilities, and of course, on the values of the moments. We found that mass points were more evenly distributed when $\alpha \leq 2$, $\gamma = 0.4$. As we will see in the “Estimation” section, this sensitivity will become an issue in estimating the parameters under the maximum likelihood method.

The rest of the paper will be organized in the following way. “Structural properties” introduces the discrete Pareto (IV) distribution. The maximum likelihood estimation in DPIV distribution is discussed in detail with simulation studies in “Estimation”. For illustrative purposes, three different data sets from various real life scenarios are re-analyzed



to show the applicability of the proposed DPIV distribution “[Application](#)”. Finally, we conclude this paper by providing some final remarks in “[Concluding remarks](#)” sections.

Structural properties

In this section, we discuss some important structural properties of the DPIV distribution. At first, we have the following lemma.

Lemma 1 *If a random variable Y follows the Pareto distribution with parameters α, γ, σ , then $X = \lfloor Y \rfloor$ follows the DPIV (θ, γ, σ) .*

Proof Follows immediately from (3).

Stochastic ordering is an integral tool to judge comparative behaviors of random variables. Many stochastic orders exist and have various applications. Theorem 1 and Corollary 1 (below) give some results on the stochastic orderings of the DPIV. The orders considered here are the stochastic order \leq_{st} , and the expectation order \leq_E . □

Theorem 1 *The DPIV (θ, γ, σ) has the following properties.*

- Suppose $X_1 \sim \text{DPIV}(\theta, \gamma, \sigma_1)$ and $X_2 \sim \text{DPIV}(\theta, \gamma, \sigma_2)$. If $\sigma_1 < \sigma_2$, then $X_1 \leq_{st} X_2$.
- Suppose $X_1 \sim \text{DPIV}(\theta, \gamma_1, \sigma)$ and $X_2 \sim \text{DPIV}(\theta, \gamma_2, \sigma)$. If $\gamma_1 > \gamma_2$, then $X_2 \leq_{st} X_1$.
- Suppose $X_1 \sim \text{DPIV}(\theta_1, \gamma, \sigma)$ and $X_2 \sim \text{DPIV}(\theta_2, \gamma, \sigma)$. If $\theta_2 > \theta_1$, then $X_1 \leq_{st} X_2$.

Proof Follows immediately from the c.d.f. of the DPIV (θ, γ, σ) distribution.

Next, we describe the expectation ordering in the next Corollary which follows from Theorem 1. □

Corollary 1 • Suppose $X_1 \sim \text{DPIV}(\theta, \gamma, \sigma_1)$ and $X_2 \sim \text{DPIV}(\theta, \gamma, \sigma_2)$. If $\sigma_1 < \sigma_2$, then $X_1 \leq_E X_2$.

- Suppose $X_1 \sim \text{DPIV}(\theta, \gamma_1, \sigma)$ and $X_2 \sim \text{DPIV}(\theta, \gamma_2, \sigma)$. If $\gamma_1 > \gamma_2$, then $X_2 \leq_E X_1$.
- Suppose $X_1 \sim \text{DPIV}(\theta_1, \gamma, \sigma)$ and $X_2 \sim \text{DPIV}(\theta_2, \gamma, \sigma)$. If $\theta_2 > \theta_1$, then $X_1 \leq_E X_2$.

Notice that for the DPIV (θ, γ, σ) distribution

$$\begin{aligned} & g^2(x) - g(x-1) \times g(x+1) \\ &= \theta^{2 \log\left(1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right) + \theta^{2 \log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)} - \theta^{\log\left(1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right) + \log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \\ & \quad - \theta^{\log\left(1 + \left(\frac{x-1}{\sigma}\right)^{\frac{1}{\gamma}}\right) + \log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)} + \theta^{\log\left(1 + \left(\frac{x-1}{\sigma}\right)^{\frac{1}{\gamma}}\right) + \log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)}. \end{aligned} \quad (6)$$

For $\theta > 0, \sigma > 0, \gamma > 0$, the distribution is infinitely divisible since the expression in Eq. (6) can be negative and $g(0) \neq 0, g(1) \neq 0$, see Warde and Katti (1971) for details. Then, in this case the distribution has its mode at zero. Furthermore, the infinitely divisible distribution plays an important role in many areas of statistics, for example, in stochastic processes and in actuarial statistics. When a distribution G is infinitely divisible, then for any integer $y \geq 2$, there exists a distribution G_y such that G is the y -fold convolution of G_y , namely, $G = G_y^{*y}$. Also, when a distribution is infinitely divisible an upper bound for the variance can be obtained when $\theta > 0, \sigma > 0, \gamma > 0$ (see Johnson and Kotz (1982), page 75), which is given by

$$\text{Var}(X) \leq \frac{g(1)}{g(0)}.$$

The following results hold for the p.m.f. of the DPIV distribution in Eq. (1) which are listed as follows:

- For all $k = 0, 1, \dots$, and for any $m \geq 1$,

$$\binom{k+m}{m} g(x+m)g(0) \geq g(k)g(m),$$

for further details, see Steutel and Van Harn (2004), page 51, Proposition 8.4.

- For all $x = 0, 1, \dots$, $g(x) \leq \exp(-1)$. See Steutel and Van Harn (2004), page 56, Proposition 9.2.
- The distribution is strictly log-concave and strongly unimodal, see theorem 3 in Keilson and Gerber (1971).
- The cumulants of an infinitely divisible distribution on the set of non-negative integers (as far as they exist) are non-negative, see Steutel and Van Harn (2004), page 47, Corollary 7.2. This will imply that the skewness of the new distribution is positive, since the third cumulant equals the third central moment.

Increasing and decreasing failure rate

The purpose of this section is to find a relationship between the parameters of this model in order to study the failure rates. From the beginning of this section, we have that the failure rate, $r(x)$ is given by : $r(x) = \frac{P(x)}{S(x)} = 1 - \theta^{\phi(x)}$; $\phi(x) = \log\left(\frac{1 + \left(\frac{1+x}{\sigma}\right)^{\frac{1}{\gamma}}}{1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}}\right)$. Next,

on setting $r(1) = r(2)$, we get $\sigma = \left[\frac{(4-2^{1+\frac{1}{\gamma}})}{2^{\frac{2}{\gamma}} - 3^{\frac{1}{\gamma}}}\right]^{-\gamma}$. If γ is an integer, the failure rate is indeterministic. Therefore, this parametric relationship yields information about the

failure rate if $\gamma < 2$. Note that from (4), $r(x)$ is decreasing if $\theta^{\phi(x)}$ is increasing. Since $0 \leq \theta \leq 1$, $\theta^{\phi(x)}$ will be increasing if $\phi(x)$ is increasing. Furthermore, we observe the following:

- If $\frac{1}{\gamma} < 1$, then $\phi(x)$ is increasing, and with θ increasing (equivalently α decreasing), $r(x)$ is decreasing. This proves the fact that it has a DFR in such a scenario.
- If $\frac{1}{\gamma} > 1$, then $\phi(x)$ is decreasing, and with θ decreasing (equivalently α increasing), $r(x)$ is increasing. This proves the fact that it has a IFR in such a scenario.

According to (Kemp 2004), page 3074 one can say the following relationships for discrete distributions which are applicable to the DPIV distribution Eq. (3) given below.
 $IFR/DFR \implies IFRA/DFRA \implies NBU/NWU \implies NBUE/NWUE \implies DMRL/IMRL$.

Moments and generating functions

The r^{th} moment of a random variable X with the p.m.f. in Eq. (1) will be for any $r \in \mathbb{Z}^+$:

$$\begin{aligned} E(X^r) &= \sum_{x=0}^{\infty} x^r p(x) \\ &= \sum_{x=1}^{\infty} (x^r - (x-1)^r) S(x) \\ &\leq r \sum_{x=1}^{\infty} x^{r-1} \theta^{\log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \\ &\leq r \sum_{x=1}^{\infty} x^{r-1} \left(\frac{x}{\sigma}\right)^{-\frac{\alpha}{\gamma}}, \quad \text{using } \theta = \exp(-\alpha) \\ &\leq r \sum_{x=1}^{\infty} \sigma^{\frac{\alpha}{\gamma}} \left(\frac{1}{x^{1-r+\frac{\alpha}{\gamma}}}\right) \\ &= r \sigma^{\frac{\alpha}{\gamma}} \sum_{x=1}^{\infty} \left(\frac{1}{x^{1-r+\frac{\alpha}{\gamma}}}\right). \end{aligned}$$

Then, $E(X^r)$ will be convergent if $\frac{\alpha}{\gamma} > r$, i.e., $\exp(-\gamma \times r) > \theta$, where $\theta = \exp(-\alpha)$. Consequently, we can write the following theorem:

Theorem 2 $E(X^r)$ exists if and only if $\exp(-\gamma \times r) > \theta$.

Proof Immediately follows from the previous discussion. □

Probability generating function and factorial moments

The relationship between the probabilities and the associated factorial moments (see Johnson et al. (2005), page 59)

$$P(X = x) = \sum_{r \geq 0} (-1)^r \frac{\mu'_{[x+r]}}{x! r!},$$

which is due to (Frechet 1940; 1943). Again, one may write

$$\sum_{i \geq x} P(X = i) = \sum_{j \geq x} (-1)^{x+j} \binom{j-1}{x-1} \frac{\mu'_{[j]}}{j!},$$

which is due to (Laurent 1965).

Theorem 3 *Characterization via minimization*

Let $X'_i, i = 1, 2, 3, \dots$ be non-negative independent and identically distributed (i.i.d.) integer valued random variables with $X_{(1)} = \min_{1 \leq i \leq n} X_i$. Then, $X_{(1)} \sim \text{DPIV}(\theta^n, \sigma, \gamma)$ iff $X_i \sim \text{DPIV}(\theta, \sigma, \gamma)$.

Proof Sufficiency part:

Let $X_i \sim \text{DPIV}(\theta, \sigma, \gamma)$. Then, $S(X) = \theta^{\log\left(1 + \left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}}\right)}, x = 1, 2, 3, \dots$. Consider $\forall x_{(1)} = 1, 2, \dots$:

$$\begin{aligned} S(X_{(1)}) &= P(X_{(1)} \geq x_{(1)}) \\ &= \{P(X_{(1)} \geq x_{(1)})\}^n \\ &= \theta^{n \log\left\{1 + \frac{1+x_{(1)}}{\sigma}\right\}}. \end{aligned}$$

Hence $X_{(1)} \sim \text{DPIV}(\theta^n, \sigma, \gamma)$.

Necessary part:

Let $S(X_{(1)}) = \theta^{n \log\left\{1 + \frac{1+x_{(1)}}{\sigma}\right\}}; x_1 = 1, \dots$

We know that

$$\begin{aligned} S(X) &= P(X_1 \geq x) \\ &= (P(X_{(1)} \geq x))^n \\ &= \theta \log\left(1 + \left(\frac{1+x_{(1)}}{\sigma}\right)^{\frac{1}{\gamma}}\right). \end{aligned}$$

Hence the proof. □

Theorem 4 Let $X'_i, i = 1, 2, 3, \dots$ be non-negative independent integer valued random variables with $X_{(1)} = \min_{1 \leq i \leq n} X_i$. Then, $X_{(1)} \sim \text{DPIV}(\delta, \sigma, \gamma)$ if and only if $X_i \sim \text{DPIV}(\theta_i, \sigma, \gamma)$, where $\delta = \prod_{i=1}^n \theta_i$.

Proof It is similar to that of Theorem 3 and hence omitted. □

Theorem 5 If $Y \sim \text{DPIV}(\theta, \sigma, \gamma)$, then

$$\frac{P(Y > x)}{\left(1 + \left(\frac{1+x}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \rightarrow 1,$$

as $x \rightarrow \infty$.

Proof Let $x \rightarrow \infty$, and also let $t = t(x)$ be the unique integer such that $t(x) \leq x \leq t(x) + 1$. As a consequence, we have $S(t(x)) \geq P(Y > x) \geq S(t(x) + 1)$. Therefore,

$$\left[\frac{\left(1 + \left(\frac{1+x}{\sigma}\right)^{\frac{1}{\gamma}}\right)}{\left(1 + \left(\frac{1+t(x)}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \right]^{\alpha} \geq \frac{P(Y > x)}{\left[\left(1 + \left(\frac{1+x}{\sigma}\right)^{\frac{1}{\gamma}}\right)\right]^{-\alpha}} \geq \left[\frac{\left(1 + \left(\frac{1+x}{\sigma}\right)^{\frac{1}{\gamma}}\right)}{\left(1 + \left(\frac{t(x)+2}{\sigma}\right)^{\frac{1}{\gamma}}\right)} \right]^{\alpha}.$$

Next, note that since $\frac{x}{t(x)} \rightarrow 1$, the sequence in the middle is bounded by two sequences which converges to 1 as $x \rightarrow \infty$. Hence, the proof. \square

Estimation

For a random sample of size n drawn from the p.m.f. in Eq. (3), the log-likelihood function is given by

$$\ell = \sum_{i=1}^n \log \left[\theta \log \left(1 + \left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} \right) - \theta \log \left(1 + \left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} \right) \right]. \quad (7)$$

The corresponding maximum likelihood equations are (by taking partial derivatives of ℓ w.r.t. θ , σ and γ respectively)

$$\begin{aligned} \frac{\partial \ell}{\partial \theta} &= \left[\theta \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) - \theta \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \right] \\ &\times \left[\log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \theta \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right)^{-1} - \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \theta \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right)^{-1} \right]. \end{aligned} \quad (8)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \sigma} &= \left[\gamma \sigma \left(\theta \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) - \theta \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \right) \right] \\ &\times \left[\log(\theta) \left(\frac{\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} \theta \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right)}{\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1} - \frac{\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} \theta \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right)}{\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1} \right) \right]. \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \gamma} &= \left[\gamma^2 \left(\theta \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) - \theta \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \right) \right] \\ &\times \left[\log(\theta) \left(\frac{\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \theta \log \left(\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right)}{\left(\frac{x_i+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1} - \frac{\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right) \theta \log \left(\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1 \right)}{\left(\frac{x_i}{\sigma}\right)^{\frac{1}{\gamma}} + 1} \right) \right]. \end{aligned} \quad (10)$$

The maximum likelihood estimates of θ , σ and γ can be obtained by setting Eqs. (8)-(10) equal to zero and solving simultaneously using bivariate Newton-Raphson method. The asymptotic variance-covariance matrix of the MLEs of parameters θ , σ and γ are obtained by inverting the Fisher's information matrix with elements which are negative expected

values of second order derivatives of the log-likelihood function (ℓ). Using the general theory of MLEs, (see [Appendix A](#)) the asymptotic distribution of $(\hat{\theta}, \hat{\sigma}, \hat{\gamma})$ is a trivariate normal with mean (θ, σ, γ) and variance-covariance matrix is given by

$$\begin{bmatrix} E\left(-\frac{\partial^2 \ell}{\partial \theta^2}\right) & E\left(-\frac{\partial^2 \ell}{\partial \theta \partial \sigma}\right) & E\left(-\frac{\partial^2 \ell}{\partial \theta \partial \gamma}\right) \\ E\left(-\frac{\partial^2 \ell}{\partial \sigma \partial \theta}\right) & E\left(-\frac{\partial^2 \ell}{\partial \sigma^2}\right) & E\left(-\frac{\partial^2 \ell}{\partial \sigma \partial \gamma}\right) \\ E\left(-\frac{\partial^2 \ell}{\partial \gamma \partial \theta}\right) & E\left(-\frac{\partial^2 \ell}{\partial \gamma \partial \sigma}\right) & E\left(-\frac{\partial^2 \ell}{\partial \gamma^2}\right) \end{bmatrix}^{-1}.$$

The exact expressions for various expectations above are cumbersome. However, in practice we would estimate above matrix by the inverse of observed Fisher’s information matrix using the following approximations

$$\begin{aligned} E\left(-\frac{\partial^2 \ell}{\partial \theta^2}\right) &\approx -\frac{\partial^2 \ell}{\partial \theta^2} \Big|_{\theta=\hat{\theta}, \sigma=\hat{\sigma}, \gamma=\hat{\gamma}} \\ E\left(-\frac{\partial^2 \ell}{\partial \sigma^2}\right) &\approx -\frac{\partial^2 \ell}{\partial \sigma^2} \Big|_{\theta=\hat{\theta}, \sigma=\hat{\sigma}, \gamma=\hat{\gamma}} \\ E\left(-\frac{\partial^2 \ell}{\partial \gamma^2}\right) &\approx -\frac{\partial^2 \ell}{\partial \gamma^2} \Big|_{\theta=\hat{\theta}, \sigma=\hat{\sigma}, \gamma=\hat{\gamma}} \\ E\left(-\frac{\partial^2 \ell}{\partial \sigma \partial \theta}\right) &\approx -\frac{\partial^2 \ell}{\partial \sigma \partial \theta} \Big|_{\theta=\hat{\theta}, \sigma=\hat{\sigma}, \gamma=\hat{\gamma}} \\ E\left(-\frac{\partial^2 \ell}{\partial \gamma \partial \theta}\right) &\approx -\frac{\partial^2 \ell}{\partial \gamma \partial \theta} \Big|_{\theta=\hat{\theta}, \sigma=\hat{\sigma}, \gamma=\hat{\gamma}} \\ E\left(-\frac{\partial^2 \ell}{\partial \sigma \partial \gamma}\right) &\approx -\frac{\partial^2 \ell}{\partial \sigma \partial \gamma} \Big|_{\theta=\hat{\theta}, \sigma=\hat{\sigma}, \gamma=\hat{\gamma}}. \end{aligned} \tag{11}$$

The expressions are provided in [Appendix B](#). For simulation study, we consider the following choices of the parameters (for a random sample of sizes $n = 50, 100, 200$) respectively. A thorough simulation study was performed by generating 2000 samples of various sizes from $DPIV(\theta, \gamma, \sigma)$.

- **Choice 1:** $\theta = 0.2, \gamma = 0.2, \sigma = 2$.
- **Choice 2:** $\theta = 0.4, \gamma = 0.7, \sigma = 2.5$.
- **Choice 3:** $\theta = 0.7, \gamma = 1.5, \sigma = 3$.
- **Choice 4:** $\theta = 0.9, \gamma = 1, \sigma = 1.7$.
- **Choice 5:** $\theta = 0.5, \gamma = 0.5, \sigma = 1$.
- **Choice 6:** $\theta = 0.8, \gamma = 0.7, \sigma = 1.2$.
- **Choice 7:** $\theta = 0.1, \gamma = 2, \sigma = 2.5$.

The parameter estimates with their associated 95% confidence intervals are provided in [Tables 1, 2, 3, 4, 5, 6 and 7](#).

Comment on the simulation study: From [Tables 1, 2, 3, 4, 5, 6 and 7](#), we observe that the convergence of N-R method is slightly strong for smaller values of θ, γ , and σ as θ, γ , and σ are close to the corresponding actual values in expectation. The associated

Table 1 Maximum likelihood estimates of the parameters (Choice 1)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.158 (0.1014, 1.2823)	0.146 (0.1027, 0.8933)	1.894 (1.3722, 3.2426)
100	0.179 (0.1321, 1.0874)	0.168 (0.145, 0.7826)	1.932 (0.7865, 3.0105)
200	0.192 (0.1526, 0.9586)	0.182 (0.1628, 0.4532)	1.943 (0.1235, 2.2703)

Table 2 Maximum likelihood estimates of the parameters (Choice 2)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.285 (0.178, 1.6263)	0.5834 (0.2366, 1.5376)	1.958 (1.2315, 3.2428)
100	0.347 (0.2239, 1.0873)	0.632 (0.3285, 1.1745)	2.223 (1.7856, 3.0127)
200	0.381 (0.2783, 0.9734)	0.671 (0.4029, 0.9822)	2.430 (1.9852, 2.9733)

Table 3 Maximum likelihood estimates of the parameters (Choice 3)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.622 (0.3041, 1.2895)	1.3768 (0.2563, 2.7617)	2.689 (1.2542, 3.5624)
100	0.638 (0.4134, 1.0872)	1.4261 (0.7286, 2.0157)	2.734 (1.5804, 3.2381)
200	0.657 (0.5029, 0.9326)	1.4792 (0.8252, 1.7556)	2.8631 (1.6953, 3.0317)

Table 4 Maximum likelihood estimates of the parameters (Choice 4)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.802 (0.3089, 1.3702)	0.0793 (0.0215, 0.9851)	1.527 (0.9821, 2.2362)
100	0.837 (0.4219, 1.1857)	0.089 (0.0426, 0.6329)	1.594 (1.1026, 2.0139)
200	0.876 (0.5392, 1.0722)	0.0927 (0.0658, 0.5843)	1.6488 (1.2301, 1.9842)

Table 5 Maximum likelihood estimates of the parameters (Choice 5)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.396 (0.1172, 0.8324)	0.412 (0.1236, 0.8907)	0.864 (0.3215, 1.3849)
100	0.432 (0.2396, 0.7842)	0.428 (0.2169, 0.8136)	0.893 (0.4107, 1.3262)
200	0.465 (0.3176, 0.7021)	0.448 (0.2895, 0.6536)	0.923 (0.4868, 1.1709)

Table 6 Maximum likelihood estimates of the parameters (Choice 6)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.721 (0.3162, 1.1859)	0.613 (0.2814, 1.3753)	0.985 (0.7832, 2.2465)
100	0.745 (0.4431, 1.2807)	0.627 (0.3413, 1.2527)	1.013 (0.8556, 2.0167)
200	0.763 (0.5192, 0.8623)	0.663 (0.4836, 0.9162)	1.032 (0.9387, 1.8946)

Table 7 Maximum likelihood estimates of the parameters (Choice 7)

n	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$
50	0.0289 (0.0036, 0.2968)	1.428 (1.1023, 3.0681)	2.136 (1.3418, 3.3969)
100	0.043 (0.0096, 0.2215)	1.573 (1.1859, 2.8766)	2.268 (1.4322, 3.1851)
200	0.068 (0.0172, 0.1817)	1.874 (1.2368, 2.4579)	2.418 (1.5732, 3.0156)

confidence intervals (in the parenthesis) cover the actual values of θ, γ , and σ quite satisfactorily. From the estimated value of θ , i.e., $\hat{\theta}$ one can obtain an estimate of α by using invariance property of the MLE. Also, notice that there exists an appreciable amount of error when the sample size is very small (values of the estimates for $n = 50$ in Tables 1, 2, 3, 4, 5, 6 and 7. Since, α appears in powers of X_i 's (from the original model, before reparametrization) it is expected to have more fluctuations in N-R method, resulting in weak convergence of the method which results in an unstable estimated values in some sense. Furthermore, as mentioned in the "Introduction" section, for a choice of $\alpha \leq 2$ (equivalently $\theta = \exp(-\alpha) \leq 0.1353$, the estimated value of θ is not good. A full scale simulation study with all possible such combination of values of each of the parameters of DPIV distribution under the Bayesian paradigm is required which will be a subject matter of a separate article. One may consider the method of moments estimation to examine if such an anomaly can be removed or at least, in principle be reduced to a desired accuracy level.

Application

Data set 1

In this section we re-analyze the data which is used by Krishna and Pundir (2009). The data set comprises the recordings of (Phyo 1973) of the total number of carious teeth among the four deciduous molars in a sample of 100 children 10 and 11 years old. Symmetry between right and left molars is presumed and only the right molars are considered with a time unit of two years. The data are given in Table 8. The p -values of χ^2 -statistic are 0.000024, 0.2614, 0.5438, 0.4169, 0.1339 and 0.6238 for Poisson, geometric, DBD, DPareto, GBPareto(discrete) and DPIV distributions, respectively. This reveals that Poisson and Geometric distributions are not good fit at all, whereas GBPareto (discrete), DBD and DPareto are good fit with DPIV being the best one.

We compute the expected frequencies (E_i) for fitting Poisson, Geometric, DBD, DPareto, GBPareto, DPIV distributions and pool the frequencies for 3 or more in order to apply χ^2 -test for goodness of fit. For the calculation of expected frequencies we use ML estimates in each case. The estimated value of the parameter is given in parenthesis in column one of Table 9.

Data set 2

A second application of the distributions is for modeling discrete data in which the frequencies at successive values increase. In general, most real discrete data are unimodal, multimodal or with decreasing frequencies. However, for this data set a reverse pattern can be observed. As an illustration, we consider data on duckweed fronds for plants growing in pure water observed weekly (for details, see Hand et al. (1993)) presented in Table 10. We fit this data to DPIV(θ, γ, σ) (equivalently, DPIV), Poisson, Geometric, and

to the generalized Poisson distribution (GPD) with p.m.f. (see, for details, Consul (1989)) given by

Table 8 Data set on total number of carious teeth among the four deciduous molars

Total number of carious teeth (x)	0	1	2	3	4	Total
Frequency	64	17	10	6	3	100

Table 9 Table for goodness of fit

Total number of carious teeth (x)	0	1	2	≥ 3	Total	p-value
Observed Frequency (O_i)	64	17	10	9	100	
Poisson(0.67)						
E_i	51.2	34.3	11.5	3	100	
$\frac{(O_i - E_i)^2}{E_i}$	3.2	8.7	0.2	12	24.1	0.000024
Geometric(0.5988)						
E_i	59.9	24	9.6	6.5	100	
$\frac{(O_i - E_i)^2}{E_i}$	0.28	2.74	0.02	0.96	4.0	0.2614
DBD (1.292,0.2108)						
E_i	66	19.4	6.7	7.9	100	
$\frac{(O_i - E_i)^2}{E_i}$	0.05	0.3	1.61	0.15	2.14	0.5438
DPareto (0.1935)						
E_i	68	15.6	6.2	10.2	100	
$\frac{(O_i - E_i)^2}{E_i}$	0.24	0.13	2.33	0.14	2.84	0.4169
GBPareto (1.2742, 1.5783, 2.0132)						
E_i	61	18	6	15	100	
$\frac{(O_i - E_i)^2}{E_i}$	0.1525	0.0609	0.9756	4.3902	5.5793	0.1339
DPIV(0.735, 1.282, 0.679)						
E_i	59	16.4	16.4	8.2	100	
$\frac{(O_i - E_i)^2}{E_i}$	0.4237	0.0219	2.4975	0.0780	3.02129	0.3884

Table 10 The number of duckweed fronds for plants growing in pure water observed weekly

x-value (Week)	Observed	GPD (E)	DPIV(E)	Poisson(E)	Geometric(E)
0	20	30.27	18.20	2.95	317.56
1	30	35.73	32.12	19.81	276.45
2	52	56.47	51.24	66.62	240.66
3	77	73.48	87.62	149.34	209.51
4	135	130.49	138.67	251.04	182.38
5	211	198.56	197.15	337.62	158.77

Table 11 The estimated parameters and goodness of fit for the duckweed fronds for plants growing data

Model	Parameters	χ^2	χ^2 p-value
DPIV	$\hat{\theta} = 0.6539$ $\hat{\gamma} = 1.1822$ $\hat{\sigma} = 2.0175$	2.6865	0.611579
GPD	$\hat{\lambda} = 1.2847$ $\hat{\theta} = 2.8652$	5.86103	0.2097
Poisson	$\hat{\lambda} = 6.7242$	9.9348	< 0.00001
Geometric	$\hat{p} = 0.12946$	52.54	< 0.00001

Table 12 The number of outbreaks of strike in the coal-mining industry in UK

x-value (Week)	Observed	GPD(E)	DPIV(E)	Poisson(E)	Geometric(E)
0	46	50.01	44.22	57.74	78.25
1	76	65.77	79.12	57.38	38.99
2	24	32.23	22.10	28.51	19.44
3	9	7.23	8.24	9.44	9.68
≥ 4	1	0.76	2.32	2.34	4.83

$$P(\theta, \lambda) = \begin{cases} \theta(\theta + \lambda x)^{x-1} e^{-(\theta+\lambda x)} / x! & x = 0, 1, 2, \dots \\ 0 & \text{for } x > m \text{ if } \lambda < 0, \end{cases}$$

where $\theta > 0$, $\max(-1, -\theta/m) \leq \lambda \leq 1$ and $m(\geq 4)$ is the largest positive integer for which $\theta + m\lambda \geq 0$ when $\lambda < 0$.

From the above table (Table 11) it appears that clearly, the DPIV distribution provides the best fit.

Data set 3

In this Section, the discrete gamma-Lomax distribution is applied to a data set that is taken from (Consul 1989). The data set represents the observed frequencies of the number of outbreaks of strike in the coal-mining industry in the U.K. during 1948 – 1959. The data are depicted in Table 12 along with the expected frequencies corresponding to DPIV, Poisson, Geometric, and the Generalized Poisson distribution utilized by Consul (1989) as given in earlier “Data set 2” section. Consul (1989) applied the Generalized Poisson distribution (GPD) to this data set to examine the efficacy of the GPD model.

From the above table (Table 13) it appears that clearly, the DPIV distribution provides the best fit.

Concluding remarks

In this paper, we have proposed a new discrete analogue of the continuous Pareto (IV) distribution (DPIV distribution in short with three parameters) and derived some of its interesting distributional properties. The DPIV distribution offers good flexibilities in terms of shapes for the probability mass functions and hazard rate functions. The “Application” shows that the DPIV can be useful in fitting data which are positively skewed as well as to other data sets with slightly different shapes. The estimation of the model parameters are discussed in the classical set up under the method of maximum likelihood. From the “Application” sections, it appears that the DPIV distribution provides a better alternative to the existing discrete Pareto probability models.

Table 13 The estimated parameters and goodness of fit for the outbreaks of strike in the coal-mining industry in UK data

Model	Parameters	χ^2	χ^2 p-value
DPIV	$\hat{\theta} = 0.7892$ $\hat{\varphi} = 1.4623$ $\hat{\sigma} = 2.3245$	1.179	0.7580
GPD	$\hat{\lambda} = -0.1450$ $\hat{\theta} = 1.1377$	4.5234	0.0334
Poisson	$\hat{\lambda} = 0.9935$	9.9348	0.0191
Geometric	$\hat{p} = 0.5016$	52.54	< 0.00001

Appendix

Appendix A

In this section we provide some justification on the consistency of the maximum likelihood estimators which are given as follows: Consistency of MLEs: Under certain regularity conditions, Lehman and Casella (1998) in Theorem 3.10 page 449 has proved that MLE $\hat{\delta}$ is a consistent estimator of the parameter δ . Several conditions can be checked accordingly in our case. First condition states that the parameter space $\Omega = (\theta \in (0, 1), \sigma \in (0, \infty), \gamma \in (0, \infty))$ is a subset on the real line. The support of the random variable X is independent of the parameters. Next, the second condition $E\left(\frac{\partial \log f}{\partial \delta}\right) = 0$ can be verified easily. The third condition that the Fisher’s information $I(\delta) = E\left(-\frac{\partial^2 \log f}{\partial \delta^2}\right)$ is positive definite which can be verified easily. The final condition to verify is that $\left|\frac{\partial^3 \log f}{\partial \delta^3}\right| \leq M(x)$ with $E(M(X)) < \infty$.

In this case, we consider $M(x) = \left[\frac{1}{f(x;\delta)\left(1+\frac{\delta}{\sigma}\right)^k} \times \frac{\partial^3 \log f}{\partial \delta^3}\right]^2$. For a carefully selected large k , we verified numerically (on using Mathematica software) that for all $\Omega = (\theta \in (0, 1), \sigma \in (0, \infty), \gamma \in (0, \infty))$ and $x = 1, 2, \dots$ $\left|\frac{\partial^3 \log f}{\partial \delta^3}\right| \leq M(x)$. Furthermore, observe that $E(M(X)) = \left[\left(1 + \frac{\delta}{\sigma}\right)^k\right]^{-1} \text{Var}\left(\frac{\partial^3 \log f}{\partial \delta^3}\right)$, because $\sum_{x \in \mathbb{N}^*} \frac{\partial^3 \log f}{\partial \delta^3} = 0$. Then, $\text{Var}\left(\frac{\partial^3 \log f}{\partial \delta^3}\right)$ is finite whenever $E(X)$ and $\text{Var}(X)$ are finite. Therefore, all the regularity conditions are satisfied, one may say the MLE of δ given by $\hat{\delta}$ is a consistent estimator of δ and in our case $\hat{\delta} \sim N_3(\delta, [I(\delta)]^{-1})$.

Appendix B

In this section, we provide the elements of the observed Fisher Information matrix which are as follows:

$$\begin{aligned} & \frac{\partial^2 \ell}{\partial \theta^2} \\ &= \left[\theta^2 \left(\theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} - \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} \right)^2 \right] \\ & \times \left[-\log^2\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right) \theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right) + \log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} + \log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right) \right. \\ & \left. - 2\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right) \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} - \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} \right) \theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} + \log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right) \right. \\ & \left. \times \left(\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right) \theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} - \theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} + \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} \right) \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} \right]^{-1}. \end{aligned}$$

$$\begin{aligned} & \frac{\partial^2 \ell}{\partial \sigma^2} \\ &= \left[\log \theta \left(-\log \theta \left(\left(\left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}} + 1 \right) \left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}} \theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} - \left(\left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}} + 1 \right) \left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}} \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} \right) \right)^2 \right. \\ & \left. \times \gamma \left(\left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}} + 1 \right) \left(\left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}} + 1 \right) \left(\theta^{\log\left(\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} - \theta^{\log\left(\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} + 1\right)} \right) \right) \right]^{-1}. \end{aligned}$$

$$\begin{aligned} & \frac{\partial^2 \ell}{\partial \sigma \partial \gamma} \\ &= \left[\log \theta \left(\theta^{\log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} - \theta^{\log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} \right) \left(\frac{\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} \theta^{\log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}}}{\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} - \frac{\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} \theta^{\log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}}}{\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} \right) \right. \\ & \quad - \gamma^{-1} \log \theta \left(\frac{\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} \log\left(\frac{x+1}{\sigma}\right) \theta^{\log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}}}{\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} - \frac{\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} \log\left(\frac{x}{\sigma}\right) \theta^{\log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}}}{\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} \right) \\ & \quad \times \left(\frac{\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}} \log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}}{\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} - \frac{\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} \log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}}{\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} \right) \\ & \quad + \left[\left(\theta^{\log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} - \theta^{\log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} \right) \left(\left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}+1} \right)^2 \left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}} \log\left(\frac{x}{\sigma}\right) \theta^{\log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} \left(\log(\theta) \left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}+1} \right) \right. \\ & \quad \left. - \left(\left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}+1} \right)^2 \left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}} \log\left(\frac{x+1}{\sigma}\right) \theta^{\log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} \left(\log(\theta) \left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}+1} \right) \right] \\ & \quad \times \gamma \left(\left(\frac{x}{\sigma} \right)^{\frac{1}{\gamma}+1} \right)^2 \left(\left(\frac{x+1}{\sigma} \right)^{\frac{1}{\gamma}+1} \right)^2 \\ & \quad \times \left[\gamma^2 \sigma \left(\theta^{\log\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}+1}} - \theta^{\log\left(\frac{x+1}{\sigma}\right)^{\frac{1}{\gamma}+1}} \right)^2 \right]^{-1}. \end{aligned}$$

Abbreviations

DPIV: discrete Pareto (type IV) distribution; IFR (DFR): increasing (decreasing) failure rate; IFRA (DFRA): increasing (decreasing) failure rate average; NBU (NWU): new better (worse) than used; NBUE (NWUE): new better (worse) than used in expectation; IMRL (DMRL): Increasing (decreasing) mean residual lifetime; s.f.: survival function; p.m.f.: probability mass function; h.r.f.: hazard rate function

Acknowledgements

The authors acknowledges several of the references from which some useful ideas have generated in preparing this manuscript.

Authors' contributions

Dr. Indranil Ghosh, being the solo author contributed in its entirety in preparation for the manuscript.

Funding

The author did not receive any funding in preparing this manuscript.

Availability of data and materials

All the data sets utilized in this manuscript are available on the specific references made freely.

Competing interests

The authors have no conflict of interest.

Received: 18 March 2020 Accepted: 21 July 2020

Published online: 01 August 2020

References

Arnold, B. C.: Pareto Distributions. International Co-operative Publishing House, Maryland (1983)
 Bracquemond, C., Gaudoin, O.: A survey on discrete lifetime distributions. *Int. J. Reliab. Qual. Saf. Eng.* **10**, 69–98 (2003)
 Chakraborty, S.: Generating discrete analogues of continuous probability distributions—A survey of methods and constructions. *J. Stat. Distrib. Appl.* **2**(Article 6), 1–30 (2015)
 Consul, P. C.: Generalized Poisson Distributions: Properties and Applications. Marcel Dekker, Inc., New York (1989)
 Fréchet, M.: Les Probabilités Associaees a un Systeme d'Evenements Compatibles et Dependants, 1, Evenements en Nombre Fini Fixe, Actualites scientifiques et industrielles, No. 859. Hermann, Paris (1940)
 Fréchet, M.: Les Probabilités Associaees a un Systeme d'Evenements Compatibles et Dependants, 2, Evenements en Nombre Fini Fixe, Actualites scientifiques et industrielles, No. 942. Hermann, Paris (1943)
 Hand, D. J., Daly, F., McConway, K., Lunn, D., Ostrowski, E.: A handbook of small data sets. CRC Press, Boca Raton, Florida (1993)

- Johnson, N. L., Kotz, S.: Developments in discrete distributions, 1969–1980, correspondent paper. *Int. Stat. Rev./Rev. Int. Stat.* **50**(1), 71–101 (1982)
- Johnson, N. L., Kotz, S., Kemp, A. W.: *Univariate Discrete Distributions*. Second Edition. John Wiley and Sons, Hoboken, New Jersey (2005)
- Kemp, A. W.: Classes of discrete lifetime distributions. *Comput. Stat. Theory Methods*. **33**, 3069–3093 (2004)
- Keilson, J., Gerber, H.: Some results for discrete unimodality. *J. Am. Stat. Assoc.* **66**, 386–389 (1971)
- Krishna, H., Pundir, P. S.: Discrete Burr and discrete Pareto distributions. *Stat. Methodol.* **6**, 177–188 (2009)
- Lai, C. D.: Issues concerning constructions of discrete lifetime models. *Qual. Tech. Quant. Manag.* **10**, 251–262 (2013)
- Laurent, A. G.: *Probability distributions, factorial moments, empty cell test, Classical and Contagious Discrete Distributions*(Patil, G. P., ed.) Calcutta Statistical Publishing Society, Oxford, Pergamon (1965)
- Lehman, E. L., Casella, G.: *Theory of Point Estimation*. Springer, New York (1998)
- Nakagawa, T., Osaki, S.: The discrete Weibull distribution. *IEEE Trans. Reliab.* **R-24**, 300–301 (1975)
- Phyo, I.: Use of a Chain Binomial in the Epidemiology of Caries. *J. Dent. Res.* **52**, 750–752 (1973)
- Padgett, W. J., Spurrer, J. D.: On discrete failure models. *IEEE Trans. Reliab.* **R-34**, 253–256 (1985)
- Roy, D.: The discrete normal distribution. *Commun. Stat. Theory Methods*. **32**, 1871–1883 (2003)
- Salvia, A. A., Bollinger, R. C.: On discrete hazard functions. *IEEE Trans. Reliab.* **R-31**, 458–459 (1982)
- Steutel, F. W., Van Harn, K.: *Infinite Divisibility of Probability Distributions on the Real Line*. Marcel Dekker, New York (2004)
- Warde, W. D., Katti, S. K.: Infinite divisibility of discrete distributions II. *Ann. Math. Stat.* **42**, 1088–1090 (1971)
- Xie, M., Gaudoin, O., Bracquemond, C.: Redefining failure rate function for discrete distributions. *Int. J. Reliab. Qual. Saf. Eng.* **9**, 275–285 (2002)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
