Pakistan Journal of Statistics and Operation Research

A New Skewed Discrete Model: Properties, Inference, and Applications

Ahmed Z. Afify[1], M. El-Morshedy[2,3]*, M. S. Eliwa[3]

*Corresponding Author

1. Department of Statistics, Mathematics and Insurance, Benha University, Egypt ahmed.afify@fcom.bu.edu.eg
2. Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Saudi Arabia m.elmorshedy@psau.edu.sa
3. Mathematics Department, Faculty of Science, Mansoura University, Egypt mseliwa@mans.edu.eg

Abstract

In this paper, a new probability discrete distribution for analyzing over-dispersed count data encountered in biological sciences was proposed. The new discrete distribution, with one parameter, has a log-concave probability mass function and an increasing hazard rate function, for all choices of its parameter. Several properties of the proposed distribution including the mode, moments and index of dispersion, mean residual life, mean past life, order statistics and L-moment statistics have been established. Two actuarial or risk measures were derived. The numerical computations for these measures are conducted for several parametric values of the model parameter. The parameter of the introduced distribution is estimated using eight frequentist estimation methods. Detailed Monte Carlo simulations are conducted to explore the performance of the studied estimators. The performance of the proposed distribution has been examined by three over-dispersed real data sets from biological sciences.

Key Words: Mean residual life; COVID-19 data; Reliability; Maximum likelihood; Over-dispersed data; Simulation; Risk measures.

Mathematical Subject Classification: 60E05, 62E15.

1. Introduction

The discrete probability distributions have their great importance in modeling real count data in many applied sciences such as public health, medicine, agriculture, epidemiology, and sociology, among others. Several discrete distributions have been introduced for modeling count data. However, some traditional discrete models such as Poisson geometric distributions have limited applications in reliability, failure times, and counts. This is so, because some real count data show either under-dispersion or over-dispersion. This has motivated several statisticians to explore new discrete models based on classical continuous distributions for modeling discrete failure times and reliability data.

In the last two decades, several authors have introduced discrete models by the discretization of continuous distributions. For example, Krishna and Pundir (2009) proposed discrete analogues of the Pareto and Burr, Jazi et al. (2010) introduced the discrete inverse Weibull, and Gómez-Déniz (2010) introduced the discrete generalized exponential distribution. However, there is still a clear need to construct more flexible discrete distributions to serve several applied areas such as social sciences, economics, and reliability studies to properly suit different types of count data. Furthermore, Al-Babtain et al. (2020) proposed the natural discrete Lindley distribution. Eliwa et al. (2020a, 2020c) proposed discrete Gompertz-G family and the three-parameter discrete Lindley distribution. El-Morshedy et al. (2020a, 2020b) introduced the discrete Burr-Hatke and exponentiated discrete Lindley distributions, respectively. Almazah et al. (2021b) proposed the transmuted record type geometric distribution. Aljohani et al. (2021) introduced the uniform Poisson–Ailamujia model.

Recently, Ramos and Louzada (2019) proposed a new one parameter distribution for instantaneous failures as an application. The Ramos and Louzada distribution can be specified by the following reliability or survival function (SF) (for $x \geq 0$)

$$R(x;\theta) = \frac{\theta^2 - \theta + x}{\theta(\theta-1)} e^{-\frac{x}{\theta}}, \tag{1}$$

where $\theta \geq 2$ is the shape parameter. The probability density function (PDF) correspond to Equation (1) reduces to

$$g(x;\theta) = \frac{\theta^2 - 2\theta + x}{\theta^2(\theta-1)} e^{-\frac{x}{\theta}}. \tag{2}$$

In this paper, a discrete version of Equation (1) is proposed, to model over-dispersed count data, using the most commonly used technique to construct discrete analogies from continuous ones. Consider the underlying continuous non-negative failure time X with SF, $R(x) = P(X \geq x)$, and failure times were grouped into unit intervals. The associated probability mass function (PMF) is specified (for $x = 0, 1, 2, ...$) by

$$p(x) = R(x) - R(x+1),$$

The introduced discrete model is referred to as one parameter discrete (OPD) distribution and has closed form expressions for its PMF, moments, and other properties. We have studied some of its statistical and reliability properties. Further, we derived two important risk measures of the OPD model namely, value at risk and tail value at risk. We also focus on the estimation of its parameter from frequentist point of view. We briefly study several estimators called, maximum likelihood estimator (MLE), maximum product of spacings estimator (MPSE), least-squares estimator (LSE), percentile estimator (PCE), Anderson-Darling estimator (ADE), Cramér-von-Mises estimator (CVME), weighted least-squares estimator (WLSE), and right-tail Anderson-Darling estimator (RADE). The simulation results were introduced to compare these estimators and assess their performance. Some authors have adopted different estimators to estimate the parameters of discrete models such as Al-Babtain et al. (2021) and Almazah et al. (2021a). Finally, the flexibility of the introduced OPD distribution was illustrated by modeling three real count data from the medicine field.

We are motivated to propose the OPD distribution due to its desirable properties such as its simple closed form expressions for the PMF and cumulative distribution function (CDF), moments, and other characteristics. It also can de adopted to model several real count data in different applied fields.

The rest of the article is unfolded as follows. In Section 2, we defined the proposed OPD distribution. In Section 3, some reliability and mathematical properties of the OPD distribution were derived in explicit forms. We two risk measures for the OPD distribution and present some numerical computations for them in Section 4. In Section 5, eight estimation approaches are presented to estimate the model parameter. In Section 6, a simulation study was performed to assess and explore the performance of the aforementioned eight estimation methods. The importance and flexibility of the OPD distribution was addressed using three real count data sets in Section 7. Finally, we present some conclusions in Section 8.

2. The OPD Distribution

The random variable X is said to have the OPD distribution with a parameter $\theta \geq 2$, if its SF can be expressed as

$$S(x;\theta) = \frac{\theta^2 - \theta + x + 1}{\theta(\theta-1)} e^{-\frac{x+1}{\theta}};\ x \in \mathbb{N}_0, \tag{3}$$

where $\mathbb{N}_0 = \{0, 1, 2, 3, ..., v\}$ for $0 < v < \infty$. The behavior of the SF is given by

$$S(x;\alpha)=\begin{cases}\frac{\theta^2-\theta+1}{\theta(\theta-1)}e^{-\frac{1}{\theta}}; & x\longrightarrow 0\\ \frac{x+3}{2}e^{-\frac{x+1}{2}}; & \theta\longrightarrow 2\\ 1; & \theta\longrightarrow\infty.\end{cases}$$

The corresponding CDF and PMF to Equation (3) are

$$F(x;\theta)=1-\frac{\theta^2-\theta+x+1}{\theta(\theta-1)}e^{-\frac{x+1}{\theta}};\ x\in\mathbb{N}_0 \tag{4}$$

and

$$P_x(x;\theta)=\left\{\frac{\theta^2-\theta+x-(\theta^2-\theta+x+1)e^{-\frac{1}{\theta}}}{\theta(\theta-1)}\right\}e^{-\frac{x}{\theta}};\ x\in\mathbb{N}_0, \tag{5}$$

respectively. The PMF in Equation (5) is log-concave for all values of θ, where $\frac{P_x(x+1;\theta)}{P_x(x;\theta)}$ is a decreasing function in x for all value of θ. The behaviors of the CDF and PMF are given by

$$F(x;\alpha)=\begin{cases}1-\frac{\theta^2-\theta+1}{\theta(\theta-1)}e^{-\frac{1}{\theta}}; & x\longrightarrow 0\\ 1-\frac{x+3}{2}e^{-\frac{x+1}{2}} & \theta\longrightarrow 2\\ 0; & \theta\longrightarrow\infty\end{cases} \tag{6}$$

and

$$P_x(x;\theta)=\begin{cases}1-\frac{\theta^2-\theta+1}{\theta(\theta-1)}e^{-\frac{1}{\theta}}; & x\longrightarrow 0\\ \frac{x+2-(x+3)e^{-\frac{1}{2}}}{2}e^{-\frac{x}{2}} & \theta\longrightarrow 2\\ 0; & \theta\longrightarrow\infty,\end{cases} \tag{7}$$

respectively. Figure 1 shows the PMF plots for various values of the parameter θ. This figure shows that the OPD distribution has a unimodal PMF which is skewed to the right.

The hazard rate function (HRF) is

$$h(x;\theta)=1-\left\{\frac{\theta^2-\theta+x+1}{\theta^2-\theta+x}\right\}e^{-\frac{1}{\theta}};\ \ x\in\mathbb{N}_0. \tag{8}$$

Based on Gupta et al. (1997) log-concavity concept, it is found that the HRF of the OPD model is increasing where $\Delta\eta(w)=\eta(w+1)-\eta(w)>0;\eta(w)=1-\Pr(X=w+1)/\Pr(X=w)$. The behavior of the HRH is

$$h(x;\theta)=\begin{cases}1-\frac{\theta^2-\theta+1}{\theta^2-\theta}e^{-\frac{1}{\theta}}; & x\longrightarrow 0,\\ 1-e^{-\frac{1}{\theta}} & x\longrightarrow\infty\\ 1-\frac{x+3}{x+2}e^{-\frac{1}{2}} & \theta\longrightarrow 2\\ 0; & \theta\longrightarrow\infty.\end{cases} \tag{9}$$

Figure 2 shows the HRF plots for various values of θ. This figure illustrates that the HRF of the OPD distribution has only an increasing shape.

Suppose Z_1 and Z_2 are two independent OPD random variables with parameters θ_1 and θ_2, respectively. Then the HRF of $Y=\min(Z_1,Z_2)$ can be written as

$$\begin{aligned}h_Y(x;\theta_1,\theta_2) &= \frac{\Pr(X_1\geq x)\ \Pr(X_2\geq x)-\ \Pr(X_1\geq x+1)\ \Pr(X_2\geq x+1)}{\Pr(X_1\geq x)\Pr(X_2\geq x)}\\ &= \frac{\Pr(X_1\geq x)\ \Pr(X_2=x)+\ \Pr(X_1=x)\ \Pr(X_2\geq x)-\ \Pr(X_1=x)\ \Pr(X_2=x)}{\Pr(X_1\geq x)\ \Pr(X_2\geq x)}\\ &= 1-\frac{\theta_1^2-\theta_1+x+1}{\theta_1^2-\theta_1+x}\left[2-\frac{\theta_2^2-\theta_2+x+1}{\theta_2^2-\theta_2+x}e^{-\frac{1}{\theta_2}}\right]e^{-\frac{1}{\theta_1}}.\end{aligned}$$

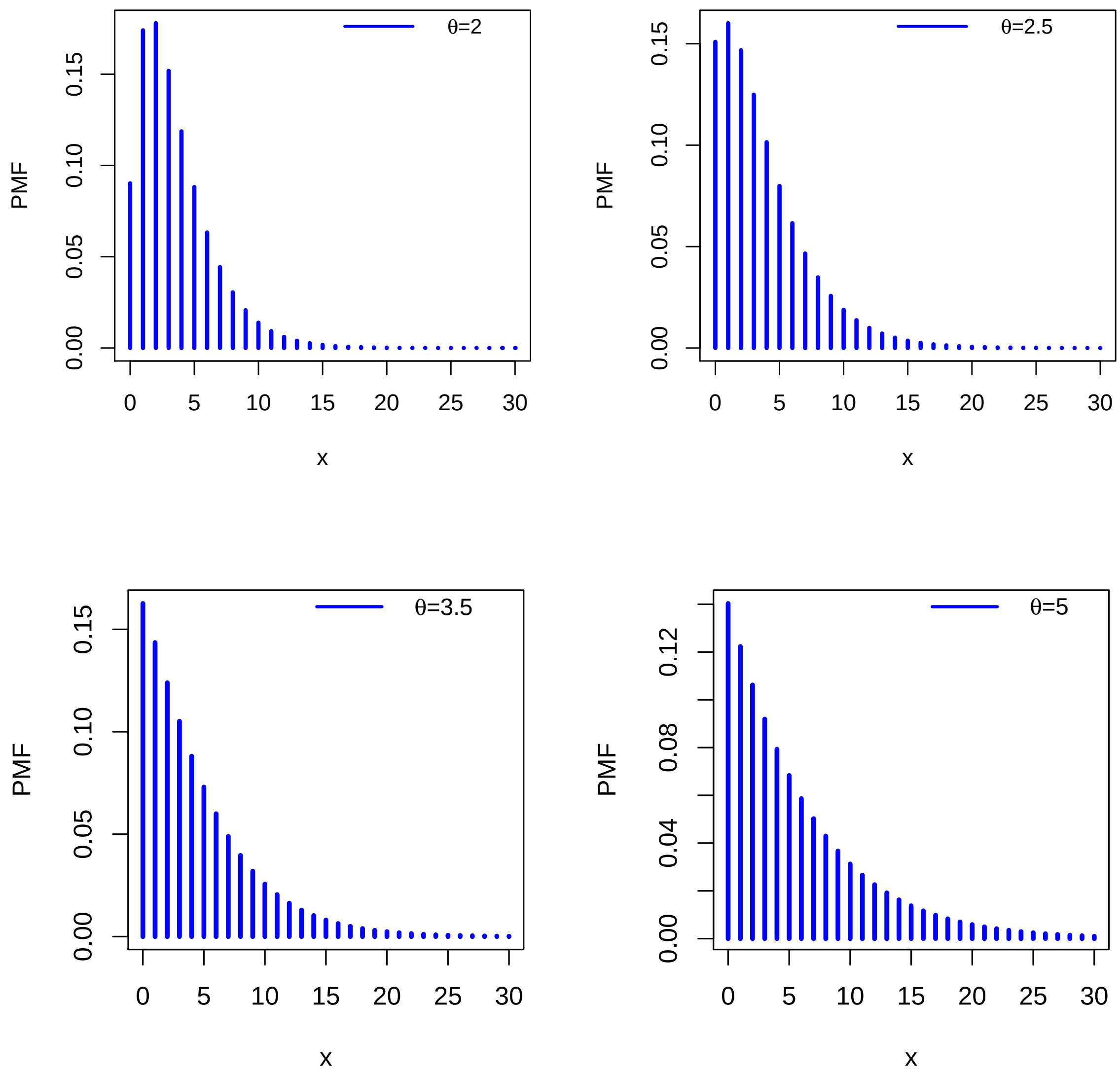

Figure 1: Plots of the PMF of the OPD distribution.

3. Some Properties

3.1. Mode

If X have a OPD distribution, then the mode of X is derived by solving the non-linear equation

$$\left(\frac{1}{\theta(\theta-1)}-\frac{\theta^2-\theta+x}{\theta^2(\theta-1)}\right)e^{-\frac{x}{\theta}}+\left(\frac{1}{\theta(\theta-1)}-\frac{\theta^2-\theta+x+1}{\theta^2(\theta-1)}\right)e^{-\frac{x}{\theta}+1}=0. \tag{10}$$

Based on Equation (10), the mode of the OPD distribution is

$$\mathrm{M}(X)=\left(e^{\frac{1}{\theta}}-1\right)^{-1}-\theta(\theta-2). \tag{11}$$

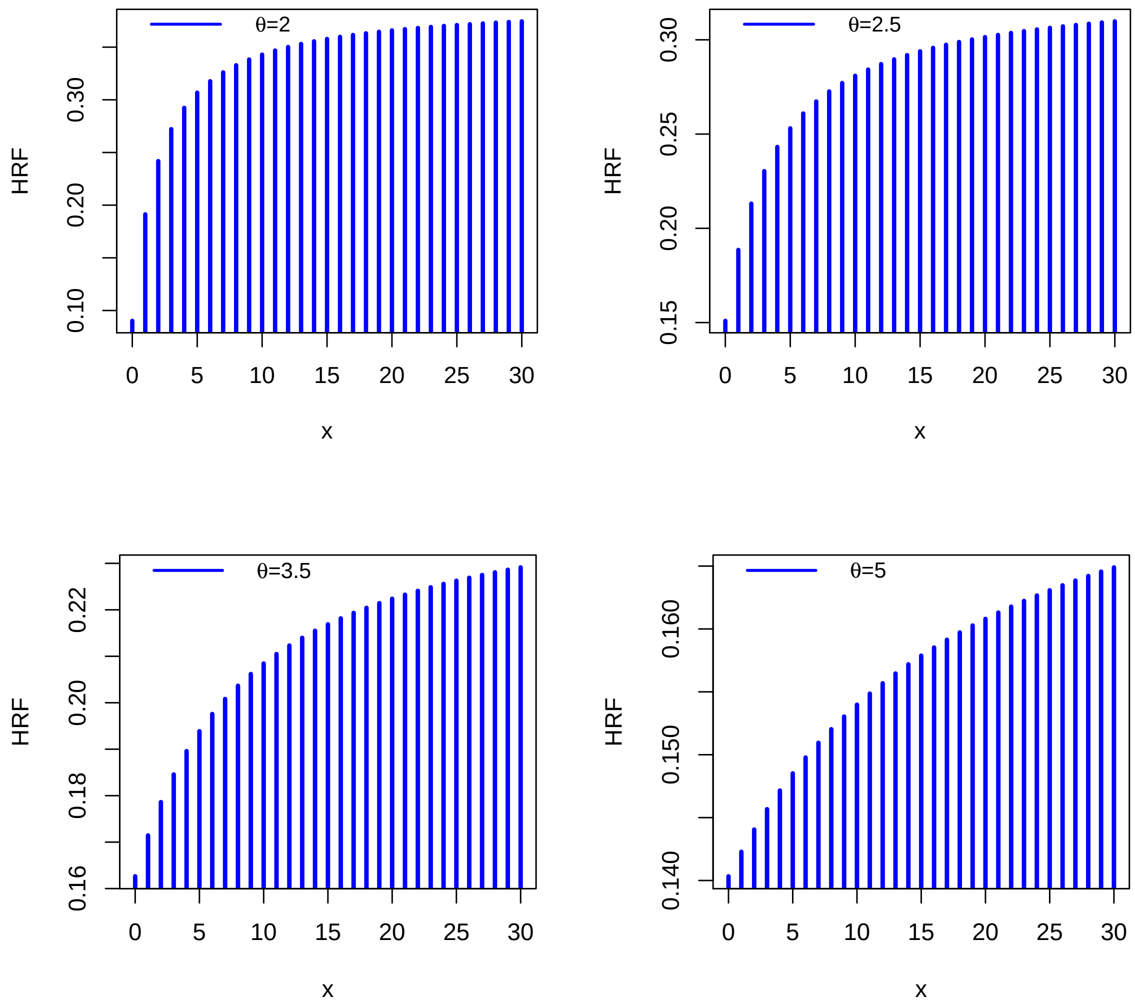

Figure 2: Plots of the HRF of the OPD distribution.

3.2. Moments and Index of Dispersion

The probability generating function (PGF) of X has the form

$$G_X(v) = \frac{\theta(\theta-1)e^{\frac{2}{\theta}} + \left[(-v-1)\ \theta^2 + (v+1)\theta + v - 1\right] e^{\frac{1}{\theta}} + v\theta(\theta-1)}{\theta(\theta-1)\left(v - e^{\frac{1}{\theta}}\right)^2}, \tag{12}$$

where $G_X(v) = \sum_{x=0}^{\infty} v^x P_x(x;\theta)$. The first two moments of the OPD distribution are

$$E(X) = \frac{\left(\theta^2-\theta+1\right) e^{\frac{1}{\theta}} - \theta(\theta-1)}{\theta\ (\theta-1)\left(e^{\frac{1}{\theta}}-1\right)^2} \tag{13}$$

and

$$E(X^2) = \frac{\left(\theta^2-\theta+1\right) e^{\frac{2}{\theta}} + 3\ e^{\frac{1}{\theta}} - \theta\ (\theta-1)}{\theta\ (\theta-1)\ \left(e^{\frac{1}{\theta}}-1\right)\left(e^{\frac{2}{\theta}} - 2\ e^{\frac{1}{\theta}} + 1\right)},$$

respectively. Based on the first two moments of the OPD distribution, the variance follows as

$$Var(X) = \left\{ \frac{\left(\theta^4 - 2\theta^3 + 2\theta^2 - \theta\right) e^{\frac{2}{\theta}} + (-2\theta^4 + 4\theta^3 - 2\theta^2 - 1)e^{\frac{1}{\theta}} + \theta(\theta^3 - 2\theta^2 + 1)}{\left(e^{\frac{2}{\theta}} - 2e^{\frac{1}{\theta}} + 1\right)\left[\theta(\theta-1)\left(e^{\frac{1}{\theta}} - 1\right)\right]^2} \right\} e^{\frac{1}{\theta}}. \quad (14)$$

Based on the first four moments, about origin, of the OPD distribution, the skewness and kurtosis can be derived in explicit forms by utilizing well-known relationships (see Eliwa et al., 2020c). Another important statistical tool, called index of dispersion (ID) which is defined by ID $= Var(X)/|E(X)|$. The ID indicates whether a certain model is appropriate to over or (under)-dispersed data. If ID $> (<) 1$, the model is over- (under-dispersed). Table 1 lists some descriptive statistics of the OPD model for various values of θ.

Table 1: Some values of descriptive measures of the OPD model.

Parameter ⟶	θ							
Measure ↓	2.0	2.5	3.0	3.5	5.5	6.0	6.5	7.0
Mean	3.50034	3.67787	4.01391	4.41429	6.23400	6.71111	7.19231	7.67659
Variance	8.07857	11.79500	15.70791	20.03320	42.12557	48.88338	56.13917	63.89356
ID	2.30794	3.20702	3.91337	4.53826	6.75739	7.28395	7.80545	8.32318
Skewness	1.39503	1.51904	1.62940	1.70797	1.86181	1.88109	1.89658	1.90923
Kurtosis	5.94046	6.35298	6.82090	7.19973	8.05644	8.17617	8.27457	8.35624

From Table 1, one can note that:

1. The mean, variance and ID increase with θ grows.
2. The OPD distribution is convenient only for modeling over-dispersed data. Moreover, it is capable of modeling leptokurtic and positively skewed data.

3.3. Mean Residual Life and Mean Past Life

The mean residual life (MRL) is a beneficial tool for analyzing the burn-in and maintenance policies. The discrete MRL is defined as

$$\Lambda(i;\theta) = \frac{1}{1 - F(i-1;\theta)} \sum_{z=i+1}^{\infty} \left[1 - F(z-1;\theta)\right]; \quad i \in \mathbb{N}_0.$$

If X have an OPD random variable, then the MRL takes the form

$$\begin{aligned} \Lambda(i;\theta) &= \frac{1}{(\theta^2 - \theta + i)\, e^{-\frac{i}{\theta}}} \sum_{z=i+1}^{\infty} \left(\theta^2 - \theta + z\right) e^{-\frac{z}{\theta}} \\ &= \left\{ -\frac{\left(-\theta^2 + \theta - i - 1\right) e^{\frac{i+2}{\theta}} + \left(\theta^2 - \theta + i\right) e^{\frac{i+1}{\theta}}}{(\theta^2 - \theta + i)\left(e^{\frac{1}{\theta}} - 1\right)^2 e^{\frac{i+3}{\theta}}} \right\} e^{\frac{2}{\theta}}. \end{aligned}$$

The variance residual life (VRL) function can be defined as

$$\begin{aligned} \Omega_{\text{VRL}}(i;\theta) &= E(X^2|X \geq i) - \left[E(X|X \geq i)\right]^2 \\ &= \frac{2e^{\frac{i}{\theta}}}{\theta^2 - \theta + i} \sum_{z=i}^{\infty} \left[z\left(\theta^2 - \theta + z + 1\right) e^{-\frac{z+1}{\theta}} \right] - (2i-1)\Lambda(i) - \left[\Lambda(i)\right]^2 \\ &= \frac{2\left\{\Phi(i,\theta) e^{\frac{i+1}{\theta}} + i\left(\theta^2 - \theta + i + 1\right) e^{\frac{2i+5}{\theta}}\right\} e^{-2(i+1)}}{(\theta^2 - \theta + i)\left(e^{\frac{1}{\theta}} - 1\right)\left[\left(e^{\frac{1}{\theta}} - 2\right) e^{\frac{2}{\theta}} + e^{\frac{1}{\theta}}\right]} - (2i-1)\Lambda(i) - \left[\Lambda(i)\right]^2, \end{aligned}$$

where

$$\Phi(i,\theta) = (i-1)(\theta^2 - \theta + i)e^{\frac{i+2}{\theta}} - 2e^{\frac{i+3}{\theta}} \left\{ (i - \frac{1}{2})\theta^2 + (-i - \frac{1}{2})\theta + i^2 - 1 \right\}.$$

Thus, X is increasing (decreasing) VRL if

$$\Omega_{\text{VRL}}(i+1) \geq (\leq)\Lambda(i)\left[1+\Lambda(i+1)\right].$$

The residual coefficient of variation (RCV) is derived in a closed form as $\text{RCV}(i)=\sqrt{\Omega_{\text{VRL}}(i)}/|\Lambda(i)|$.

The mean past life (MPL) is an important reliability concept and it used to measures the time elapsed since the failure of X given that the device is failed before i. The discrete MPL is specified by

$$\delta(i;\theta)=\frac{1}{F(i-1;\theta)}\sum_{z=1}^{i}F(z-1;\theta);\;\; i\in\mathbb{N}_0-\{0\},$$

where $\delta(0)=0$.
If X have an OPD random variable, then the MPL reduces to

$$\begin{aligned}\delta(i;\theta) &= \frac{1}{\theta(\theta-1)-(\theta^2-\theta+i)e^{-\frac{i}{\theta}}}\sum_{z=1}^{i}\left[\theta(\theta-1)-(\theta^2-\theta+z)e^{-\frac{z}{\theta}}\right]\\ &= \frac{i\theta(\theta-1)\left(e^{\frac{1}{\theta}}-1\right)^2+\left\{(\theta^2-\theta+i+1)e^{\frac{1}{\theta}}-\theta^2+\theta-i\right\}e^{-\frac{i}{\theta}}-\left\{(\theta^2-\theta+1)e^{\frac{1}{\theta}}-\theta^2+\theta\right\}}{\left\{\theta(\theta-1)-(\theta^2-\theta+i)e^{-\frac{i}{\theta}}\right\}\left(e^{\frac{1}{\theta}}-1\right)^2}.\end{aligned}$$

The mean of the OPD model can be defined by

$$\text{Mean}=i-\delta(i;\theta)F(i-1;\theta)+\Lambda(i;\theta)\left[1-F(i-1;\theta)\right];\; i\in\mathbb{N}_0-\{0\}.$$

The reversed HRF (RHRF) and the MPL are related by the following formula

$$r(i;\theta)=\frac{1-\delta(i+1;\theta)+\delta(i;\theta)}{\delta(i;\theta)};\; i\in\mathbb{N}_0-\{0\}.$$

The CDF of the OPD model follows from the MPL as

$$F(k;\theta)=F(0)_{i=1}^{k}\left[\frac{\delta(i;\theta)}{\delta(i+1;\theta)-1}\right];\;\; k\in\mathbb{N}_0-\{0\},$$

where $F(0)=\left(_{i=1}^{\infty}\left[\frac{\delta(i;\theta)}{\delta(i+1;\theta)-1}\right]\right)^{-1}$.
Table 2 lists some numerical computations of reliability concepts for different values of the parameter θ at time $i=10$, whereas Table 3 reports the same concepts with $\theta=3.5$ and different values of i.

Table 2: Some numerical computations of reliability concepts for various values of θ.

Parameter $\longrightarrow$	θ							
Measure $\downarrow$	2.0	2.5	3.0	3.5	5.5	6.0	6.5	7.0
MRL	1.86797	2.48178	3.08505	3.67268	5.88325	6.41180	6.93449	7.45261
VRL	5.14411	8.23863	11.98135	16.31913	38.98868	45.91049	53.32271	61.22410
RCV	1.21419	1.15655	1.12100	1.09993	1.06133	1.056759	1.053032	1.04991
MPL	6.85219	6.95595	6.93975	6.88505	6.61423	6.55416	6.49875	6.44782

Table 3: Some numerical computations of reliability concepts for various values i.

Parameter $\longrightarrow$	Time (i hour)$\|_{\theta=3.5}$							
Measure $\downarrow$	1.0	1.5	2.0	3.0	6.0	9.0	12.0	15.0
MRL	4.27168	4.21080	4.15559	4.05927	3.84866	3.70924	3.61014	3.53607
VRL	19.40438	19.12360	18.86252	18.39243	17.29997	16.52799	15.95560	15.51498
RCV	1.03122	1.03853	1.04512	1.05651	1.08072	1.09604	1.10645	1.11392
MPL	0.999100	1.26234	1.53115	2.09001	3.95468	6.10769	8.51856	11.13375

From Tables 2 and 3, it is noted that:

1. The MRL and VRL increase whereas the RCV and MPL decrease for increased values of θ and fixed time i.
2. The MRL and VRL decrease whereas the RCV and MPL increase for increased values of i and fixed θ.

3.4. Order Statistics and L-moment Statistics

Let $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be the order statistics of a random sample from the OPD model. Hence, the CDF of ith order statistic is

$$\begin{aligned} F_{i:n}(x;\theta) &= \sum_{k=i}^{n} \binom{n}{k} \left[F_i(x;\theta)\right]^k \left[1-F_i(x;\theta)\right]^{n-k} \\ &= \sum_{k=i}^{n} \sum_{j=0(m)}^{n-k} \Upsilon_{(m)}^{(n,k)} F_i(x;\theta,k+j), \end{aligned} \tag{15}$$

where $\Upsilon_{(m)}^{(n,k)} = (-1)^j \binom{n}{k} \binom{n-k}{j}$ and $F_i(x;\theta,k+j)$ represents the CDF of the exponentiated-OPD distribution with a power parameter $k+j$.
The corresponding PMF to Equation (15) is

$$f_{i:n}(x;\theta) = \sum_{k=i}^{n} \sum_{j=0}^{n-k} \Upsilon_{(m)}^{(n,k)} f_i(x;\theta,k+j), \tag{16}$$

where $f_{i:n}(x;\theta) = F_{i:n}(x;\theta) - F_{i:n}(x-1;\theta)$. The cth moments of $X_{i:n}$ can be expressed as

$$E(X_{i:n}^c) = \sum_{x=0}^{\infty} \sum_{k=i}^{n} \sum_{j=0}^{n-k} \Upsilon_{(m)}^{(n,k)} x^c f_i(x;\theta,k+j). \tag{17}$$

Based on Equation (17), the L-moments (L-Ms) can be derived from the following relation

$$\Theta_q = \frac{1}{q} \sum_{j=0}^{q-1} \varpi(j,q)\, rE\left(X_{q-j:q}\right), \tag{18}$$

where $\varpi(j,q) = (-1)^j \left(\; q-1j. \; \right)$
Utilizing Equation (18), we can introduce some statistical measures like L-M of mean $= \Theta_1$, L-M coefficient of skewness $= \frac{\Theta_3}{\Theta_2}$ and L-M coefficient of kurtosis $= \frac{\Theta_4}{\Theta_2}$.

4. Actuarial Measures

Probability distributions can be used to describe risk exposure. Actuaries are often interested in key risk indicators which are important in determining the degree to which their companies are subject to risks due to the changes in interest rates, prices of equity or exchange rates.
In this section, we determine the value at risk (VaR) and tail value at risk (TVaR) of the OPD distribution, which play a crucial role in portfolio optimization.

4.1. VaR Measure

The VaR of any random variable X is $Var_p = F_X^{-1}(p)$, for a probability level $p \in (0,1)$. It is the p^{th} quantile of its CDF (Artzner, 1999). The VaR of the OPD distribution is defined by

$$Var_p = \theta - 1 - \theta^2 - \theta W\left((\theta-1)(p-1)\exp(1-\theta)\right),$$

where $W(\cdot)$ is the Lambert W function (see, Lambert, 1758).

4.2. TVaR Measure

The TVaR is also known as conditional-value at-risk or conditional tail expectation. The TVaR is defined by (Klugman et al., 2012)

$$\begin{aligned} TVaR_p &= E(X|X>x_p) = x_p + \frac{\sum_{x=x_p}^{\infty}(x-x_p)}{p(x)} \\ &= x_p + \frac{E(X) - x_p + \sum_{x=0}^{x_p-1}(x_p-x)p(x)}{1-F(x_p)}. \end{aligned}$$

Using the QF and mean of the OPD distribution, the TVaR of the OPD distribution can be derived as

$$TVaR_p = x_p + 1 + \frac{\theta^2-\theta+x_p+2}{\left[\exp(\frac{1}{\theta})-1\right][\theta^2-\theta+x_p+1]} + \frac{1}{\left[\exp(\frac{1}{\theta})-1\right]^2[\theta^2-\theta+x_p+1]}.$$

4.3. Simulations of VaR and TVaR

In this sub-section, some numerical results for the VaR and TVaR measures of the OPD distribution are obtained for some values of θ. These results are obtained based on the following two steps.

1. We generated random sample of size $n = 100$ from the OPD distribution and estimated θ using the maximum likelihood method.

2. The two measures are calculated from 2,000 repetitions.

Tables 4 and 5 report the results of our simulation. These results are also summarize graphically in Figures 3 and 4. We conclude that the two measures are increasing functions in the parameter θ and the significance level (SL).

Table 4: Numerical results of the VaR and TVaR for the OPD distribution.

SL	VaR	TVaR	VaR	TVaR
	$\theta = 2.5$		$\theta = 3.5$	
0.70	4.58777	8.88141	5.10367	9.99363
0.75	5.31731	9.56545	5.94021	10.78697
0.80	6.19179	10.39204	6.94738	11.74781
0.85	7.29484	11.44304	8.22317	12.9723
0.90	8.81275	12.90121	9.98613	14.67517
0.95	11.33398	15.34509	12.92726	17.53666
0.99	16.96094	20.85813	19.52380	24.01228

Table 5: Numerical results of the VaR and TVaR for the OPD distribution.

SL	VaR	TVaR	VaR	TVaR
	$\theta = 5.0$		$\theta = 10.0$	
0.70	6.27630	12.34423	11.83046	22.79134
0.75	7.32231	13.35345	13.75414	24.69535
0.80	8.58940	14.57994	16.10262	27.02089
0.85	10.20437	16.14853	19.12130	30.01171
0.90	12.45037	18.33829	23.35990	34.21394
0.95	16.22425	22.03441	30.56695	41.36560
0.99	24.76200	30.44703	47.14509	57.84021

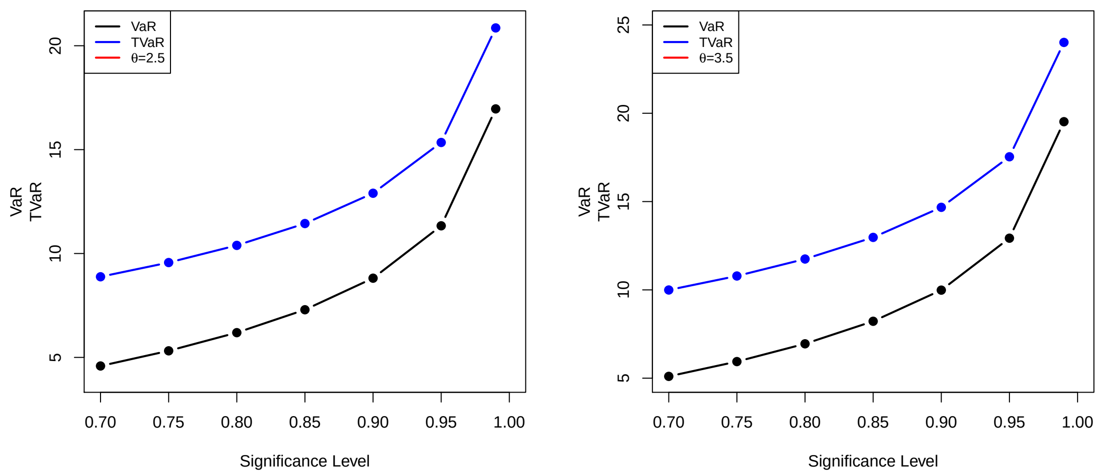

Figure 3: Plots of the VaR and TVaR of the OPD distribution.

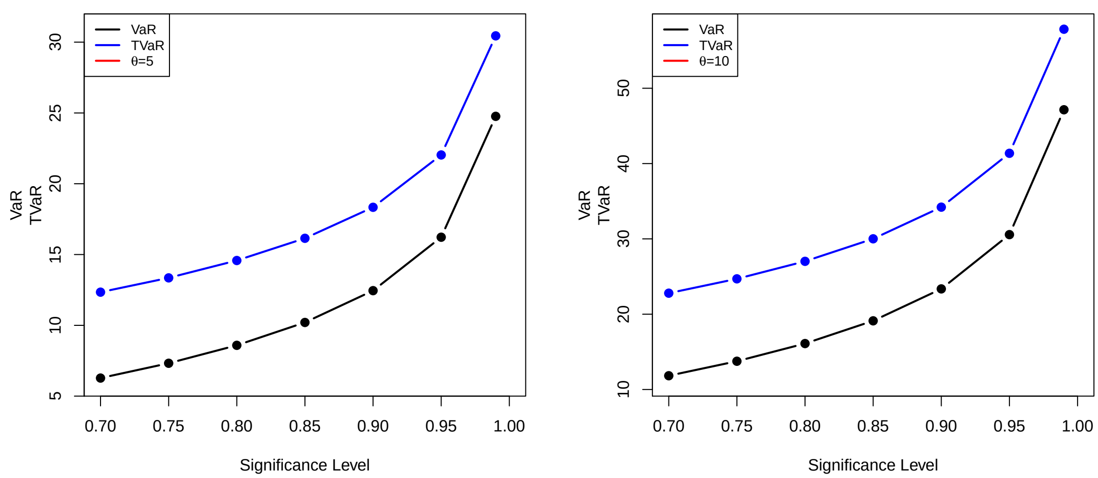

Figure 4: Plots of the VaR and TVaR of the OPD distribution.

5. Parameter Estimation

In this section, we estimate the unknown parameter θ of the OPD distribution using eight classical estimators.

5.1. Maximum Likelihood Estimator

In this section, we present the MLE of the parameter of the OPD distribution. Let $x_1, x_2, ..., x_n$ be n independent random variables with OPD(θ) distribution, then the log-likelihood function of the OPD distribution is given by

$$L(\theta; x) = \sum_{i=0}^{n} \log \left[\theta^2 - \theta + x_i - (\theta^2 - \theta + x_i + 1)e^{-\frac{1}{\theta}}\right] - \frac{1}{\theta} \sum_{i=0}^{n} x_i - n \log \left[\theta(\theta - 1)\right]. \tag{19}$$

To estimate the unknown parameter θ, we take the partial derivative of the $L(\theta; x)$ function with respect to θ and equate the result equation to zero. The result equation $\frac{\partial}{\partial\theta} L(\theta; x) = 0$ cannot be solved in closed form. Thus, we obtain the solution of the MLE for θ using the Newton-Raphson procedure.

5.2. Least Square Estimator

Let $x_{(1)}, x_{(2)}, \cdots, x_{(n)}$ be the order statistics of a random sample from the OPD distribution. The LSE of the OPD parameter, say $\widehat{\theta}_{OLS}$, can be derived by solving the non-linear equation defined by

$$\sum_{i=1}^{n}\left[1-\frac{\theta^2-\theta+x_i+1}{\theta(\theta-1)}e^{-\frac{x_i+1}{\theta}}-\frac{i}{n+1}\right]\Delta_\theta\left(x_{(i)}|\theta\right)=0, \tag{20}$$

where

$$\Delta_\theta\left(x_{(i)}|\theta\right)=\frac{\partial}{\partial\theta}\left[1-\frac{\theta^2-\theta+x_i+1}{\theta(\theta-1)}e^{-\frac{x_i+1}{\theta}}\right]. \tag{21}$$

Note that the solution of $\Delta_\theta\left(x_{(i)}|\theta\right)$ can be obtained numerically.

5.3. Weighted Least Square Estimator

The WLSE of the OPD parameter, $\widehat{\theta}_{WLS}$, can be derived by solving the non-linear equation defined by

$$\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-i+1)}\left[1-\frac{\theta^2-\theta+x_i+1}{\theta(\theta-1)}e^{-\frac{x_i+1}{\theta}}-\frac{i}{n+1}\right]\Delta_\theta\left(x_{(i)}|\theta\right)=0, \tag{22}$$

where $\Delta_\theta\left(x_{(i)}|\theta\right)$ is provided in Equation (21).

5.4. Cramer-von Mises Estimator

The CVME has less bias than other minimum distance estimators. The CVME follows as the difference between the estimate of the CDF and the empirical CDF. The CVME of the OPD parameter can be derived by solving the non-linear equation defined by

$$\sum_{i=1}^{n}\left[1-\frac{\theta^2-\theta+x_i+1}{\theta(\theta-1)}e^{-\frac{x_i+1}{\theta}}-\frac{2i-1}{2n}\right]\Delta_\theta\left(x_{(i)}|\theta\right)=0, \tag{23}$$

where $\Delta_\theta\left(x_{(i)}|\theta\right)$ is defined in Equation (21).

5.5. Maximum Product of Spacings Estimator

For $i=1,2,\ldots,n+1$, let

$$D_i(\theta)=F\left(x_{(i)}|\theta\right)-F\left(x_{(i-1)}|\theta\right),$$

be the uniform spacings of a random sample from the OPD distribution, where $F\left(x_{(0)}|\theta\right)=0$, $F\left(x_{(n+1)}|\theta\right)=1$ and $\sum_{i=1}^{n+1}D_i(\theta)=1$. The MPSE of θ, $\widehat{\theta}_{MPS}$, can be obtained by maximizing the geometric mean of the spacings

$$G(\theta)=\left[\prod_{i=1}^{n+1}D_i(\theta)\right]^{\frac{1}{n+1}},$$

with respect to θ. Or by maximizing the logarithm of the geometric mean of sample spacings

$$H(\theta)=\frac{1}{n+1}\sum_{i=1}^{n+1}\log D_i(\theta).$$

5.6. Percentile Estimator

Let $u_i = i/(n+1)$ be an unbiased estimator of $F\left(x_{(i)}|\theta\right)$. Hence, the PCE of θ, denoted by $\widehat{\theta}_{PT}$, can be obtained by minimizing

$$P(\theta) = \sum_{i=1}^{n} \left(x_{(i)} - Q\left(u_i\right)\right)^2,$$

with respect to θ, where $Q\left(u_i\right) = F^{-1}\left(x_{(i)}|\theta\right)$ is the quantile function of the OPD distribution.

5.7. Anderson-Darling and Right-Tail Anderson-Darling Estimators

The ADE is another kind of minimum distance estimators. The ADE of the OPD parameter, $\widehat{\theta}_{AD}$, is obtained by minimizing

$$AD(\theta) = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left[\log F\left(x_{(i)}|\theta\right) + \log\left(1 - F\left(x_{(i)}|\theta\right)\right)\right],$$

with respect to θ, whereas the RADE of θ, $\widehat{\theta}_{RTAD}$, can be obtained by minimizing

$$RTAD(\theta) = \frac{n}{2} - 2\sum_{i=1}^{n} F\left(x_{(i:n)}|\theta\right) - \frac{1}{n}\sum_{i=1}^{n} (2i-1)\left[\log\left(1 - F\left(x_{(n+1-i:n)}|\theta\right)\right)\right],$$

with respect to θ.

6. Simulation Results

In this section, we provide a simulation study to explore the behavior of the proposed estimators for some values of the parameter θ and several samples sizes. The performance of these estimators is explored using some measures called, average values of the estimate (AVE), average of mean squared error (MSE), average absolute bias (ABB), and mean relative error (MRE). The MSE, ABB, and MRE are calculated as follows

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (\widehat{\theta} - \theta)^2, \; |ABB| = \frac{1}{N}\sum_{i=1}^{N} |\widehat{\theta} - \theta|, \; MRE = \frac{1}{N}\sum_{i=1}^{N} \frac{|\widehat{\theta} - \theta|}{\theta}.$$

We generate $N = 5,000$ random samples from the OPD distribution using its QF of sizes $n = \{20, 50, 100, 250\}$ for some values of the parameter θ, where $\theta = (2.5, 3.5, 5, 10)$. The QF of the OPD distribution takes the form

$$Q(p) = \theta - 1 - \theta^2 - \theta W\left((\theta-1)(p-1)\exp(1-\theta)\right), \; 0 < p < 1, \tag{24}$$

where $W(\cdot)$ is the Lambert W function. The simulations results are obtained by the R software©. The performance of the proposed estimators are assessed in terms of MSE, ABB, and MRE of the estimates.

The AVE, MSE, ABB, and MRE for the parameter θ are reported in Tables 6-9. One can note that, from Tables 6-9, that the estimates of the parameter θ of the OPD distribution are obtained using eight estimation methods are quite good, where the estimates are very close to the true parameter values of θ, showing small bias, MSE and MRE for all considered values of θ. The eight estimators show the consistency property, i.e., the MSE, ABB and MRE decrease as sample size increases, for all values of θ. We conclude that the MLE, LSE, WLSE, CME, MPSE, ADE, RTADE, and PCE perform very well in estimating the parameter θ of the OPD distribution.

Table 6: Numerical results of the AVE, MSE, ABB, and MRE for $\theta = 2.5$.

Measures	n	MLE	LSE	WLSE	CVME	MPSE	ADE	RADE	PCE
AVE	20	2.98204	3.09203	2.95475	2.97104	2.61870	2.62104	2.85873	2.92243
	50	2.99163	2.85804	2.66745	2.77306	2.50674	2.50609	2.68817	2.73691
	100	2.99673	2.73274	2.58621	2.67722	2.50000	2.50085	2.60857	2.64418
	250	2.99323	2.62212	2.50983	2.60472	2.50000	2.50000	2.56046	2.58134
MSE	20	0.45134	1.16927	1.03838	1.06737	0.20829	0.35717	0.76454	0.82853
	50	0.30710	0.53061	0.36116	0.47177	0.00689	0.02055	0.31741	0.32828
	100	0.27677	0.27351	0.19272	0.24482	0.00000	0.00098	0.15118	0.15528
	250	0.25382	0.10218	0.06894	0.10142	0.00000	0.00000	0.06400	0.06396
ABB	20	0.55475	0.74053	0.70357	0.69665	0.15961	0.20412	0.63087	0.68541
	50	0.49855	0.50828	0.43477	0.47947	0.01078	0.01757	0.42136	0.44700
	100	0.49677	0.36594	0.32117	0.34554	0.00000	0.00085	0.29829	0.31006
	250	0.49323	0.22444	0.19881	0.22744	0.00000	0.00000	0.19538	0.20054
MRE	20	0.22190	0.29621	0.28143	0.27866	0.06384	0.08165	0.25235	0.27416
	50	0.19942	0.20331	0.17391	0.19179	0.00431	0.00703	0.16854	0.17880
	100	0.19871	0.14638	0.12847	0.13822	0.00000	0.00034	0.11932	0.12402
	250	0.19729	0.08978	0.07952	0.09097	0.00000	0.00000	0.07815	0.08022

Table 7: Numerical results of the AVE, MSE, ABB, and MRE for $\theta = 3.5$.

Measures	n	MLE	LSE	WLSE	CVME	MPSE	ADE	RADE	PCE
AVE	20	3.56567	3.87908	3.78959	3.75153	3.60695	3.63561	3.71572	3.95809
	50	3.54583	3.68157	3.64641	3.63438	3.50373	3.50873	3.59399	3.75927
	100	3.53142	3.58719	3.55792	3.56949	3.50016	3.50000	3.54830	3.66344
	250	3.52304	3.56417	3.54242	3.54392	3.50000	3.50000	3.53108	3.58360
MSE	20	0.52856	1.65105	1.51221	1.60044	0.28887	0.44941	1.29720	1.58121
	50	0.19163	0.75533	0.65317	0.73510	0.00847	0.01803	0.56323	0.61812
	100	0.09029	0.37058	0.33509	0.38321	0.00013	0.00000	0.27631	0.28591
	250	0.03541	0.17048	0.13746	0.16833	0.00000	0.00000	0.11648	0.10799
ABB	20	0.54448	0.99444	0.97249	1.00291	0.17436	0.23214	0.91209	0.98813
	50	0.33278	0.70078	0.64560	0.69425	0.00868	0.01257	0.60425	0.62277
	100	0.23266	0.49173	0.46383	0.50078	0.00016	0.00000	0.42211	0.42601
	250	0.14840	0.33289	0.29923	0.32902	0.00000	0.00000	0.27145	0.26224
MRE	20	0.15557	0.28413	0.27785	0.28654	0.04982	0.06633	0.26060	0.28232
	50	0.09508	0.20022	0.18446	0.19836	0.00248	0.00359	0.17264	0.17793
	100	0.06648	0.14049	0.13252	0.14308	0.00005	0.00000	0.12060	0.12172
	250	0.04240	0.09511	0.08549	0.09401	0.00000	0.00000	0.07756	0.07493

Table 8: Numerical results of the AVE, MSE, ABB, and MRE for $\theta = 5$.

Measures	n	MLE	LSE	WLSE	CVME	MPSE	ADE	RADE	PCE
AVE	20	4.68899	5.25656	5.29537	5.15667	5.15979	5.17956	5.19811	5.60445
	50	4.67416	5.11881	5.12834	5.05247	5.01190	5.02033	5.09926	5.35302
	100	4.66358	5.04919	5.09114	5.04139	4.99995	5.00019	5.05142	5.21303
	250	4.65870	5.05304	5.07396	5.02636	5.00000	5.00000	5.03099	5.10970
MSE	20	1.60743	2.94715	2.67999	2.91031	0.64442	0.71902	2.32940	2.85430
	50	0.70834	1.27397	1.06567	1.17503	0.02933	0.03815	0.93027	1.03983
	100	0.41075	0.58926	0.53409	0.58652	0.00001	0.00010	0.47041	0.50944
	250	0.23797	0.24494	0.20902	0.23560	0.00000	0.00000	0.18003	0.19306
ABB	20	1.04696	1.36973	1.29660	1.37691	0.29271	0.30636	1.21460	1.31245
	50	0.69163	0.90528	0.82520	0.86600	0.02011	0.02352	0.76466	0.80177
	100	0.52902	0.60941	0.58237	0.60699	0.00005	0.00019	0.54734	0.56579
	250	0.40796	0.39499	0.36476	0.38842	0.00000	0.00000	0.33637	0.34921
MRE	20	0.20939	0.27395	0.25932	0.27538	0.05854	0.06127	0.24292	0.26249
	50	0.13833	0.18106	0.16504	0.17320	0.00402	0.00470	0.15293	0.16035
	100	0.10580	0.12188	0.11647	0.12140	0.00001	0.00004	0.10947	0.11316
	250	0.08159	0.07900	0.07295	0.07768	0.00000	0.00000	0.06727	0.06984

Table 9: Numerical results of the AVE, MSE, ABB, and MRE for $\theta = 10$.

Measures	n	MLE	LSE	WLSE	CVME	MPSE	ADE	RADE	PCE
AVE	20	9.41989	10.33912	10.42736	10.29962	10.50932	10.50630	10.35437	11.10173
	50	9.51970	10.17165	10.17896	10.12099	10.09009	10.07717	10.19161	10.64467
	100	9.47265	10.07099	10.12363	10.09440	10.00789	10.00981	10.06349	10.41150
	250	9.49873	10.04040	10.08735	10.06305	10.00000	10.00000	10.03086	10.22152
MSE	20	6.37866	8.94238	8.22124	8.71524	3.57644	3.57743	6.97808	9.37872
	50	2.61956	3.46709	3.11047	3.43227	0.37851	0.35012	2.64606	3.41852
	100	1.50702	1.77445	1.52994	1.70201	0.01679	0.02239	1.32356	1.69434
	250	0.71957	0.67454	0.59375	0.66860	0.00000	0.00000	0.53625	0.66498
ABB	20	2.04410	2.34491	2.24850	2.31773	0.95606	0.95478	2.07088	2.36470
	50	1.29706	1.47632	1.40372	1.46990	0.16628	0.15500	1.29243	1.44570
	100	0.99170	1.06220	0.98404	1.03939	0.01125	0.01339	0.91688	1.03698
	250	0.69434	0.65342	0.61376	0.65053	0.00000	0.00000	0.58654	0.64279
MRE	20	0.20441	0.23449	0.22485	0.23177	0.09561	0.09548	0.20709	0.23647
	50	0.12971	0.14763	0.14037	0.14699	0.01663	0.01550	0.12924	0.14457
	100	0.09917	0.10622	0.09840	0.10394	0.00112	0.00134	0.09169	0.10370
	250	0.06943	0.06534	0.06138	0.06505	0.00000	0.00000	0.05865	0.06428

7. Applications to Biological Real Data

This section is devoted to illustrating the importance of the OPD distribution in modeling count data from the medicine field, using three real count data sets. The first data set contains 64 observations, and it refers to numbers of daily cases in Egypt due to COVID-19 infections from 8 March to 10 May, 2020. The data are: 12, 33, 7, 4, 8, 13, 13, 17, 16, 40, 30, 14, 46, 29, 9, 33, 39, 36, 54, 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, 232, 201, 227, 215, 248, 260, 226, 269, 358, 298, 272, 348, 388, 387, 393, 495, 488. The data are available on worldometer website at https://www.worldometers.info/coronavirus/country/egypt/. The second data represents survival times (in weeks) for 33 patients. The patients are suffering from acute myelogenous leukaemia (Feigl and Zelen, 1965). The data are: 3, 3, 30, 3, 8, 4, 2, 4, 4, 65, 100, 108, 121, 4, 134, 16, 39, 26, 22, 1, 143, 56, 1, 5, 65, 17, 7, 16, 56, 65, 22, 43, 156. The third data set refers to survival times of 44 patients suffering from head and neck cancer who retreated using a combination of radiotherapy (). The data are: 12, 32, 37, 24, 24, 74, 81, 26, 41, 58, 63, 68, 78, 47, 55, 84, 155, 159, 92, 94, 110, 127, 130, 133, 140, 112, 119, 146, 173, 179, 194, 195, 339, 432, 209, 249, 281, 319, 469, 725, 817, 519, 633, 1776.

Some descriptive statistics for the three data sets are reported in Table 10. It is clear that the three data sets are over-dispersed data.

Table 10: Descriptive measures for the three real data sets.

Data	Min.	1st Qu.	Mean	3rd Qu.	Variance	ID	Max.
Data Set I	4	37.5	140.0156	208	15642.8728	111.7223	495
Data Set II	1	4	40.8788	65	2181.1724	53.3571	156
Data Set II	12	65.5	223.4091	229	93307.6427	417.6537	1776

The estimates of the parameter θ and some goodness-of-fit measures including Cramér–von Mises (CM), Anderson-Darling (AD), Kolmogorov-Smirnov (K-S) with its p-value are reported in Table 11 for the three over-dispersed data sets, respectively. These estimates and analytical measures are obtained for the eight estimation methods using the R software. It is shown, based on p-values in Table 11, that the LS method is recommended to estimate the parameter θ of the OPD distribution for COVID-19 data, whereas the AD method is recommended to estimate θ for myelogenous leukaemia data. Further, the p-values in Table 11 reveal that the WLS method is recommended to estimate the parameter θ of the OPD distribution for survival times of head and neck cancer data.

Probability–probability (P-P) plots of all estimation methods for the three data sets are depicted in Figures 5, 6 and 7, respectively. These plots support the results in Table 11, that the OPD distribution can be used to model over-dispersed data encountered in biomedical science.

Table 11: Analytical measures and the estimates of θ using various methods of estimation.

Method	CM	AD	K-S	p-value	$\widehat{\theta}$
Daily cases in Egypt due to COVID-19 infections data					
LSE	0.09137	0.53766	0.06888	0.92829	167.4075
CVME	0.09128	0.53725	0.07024	0.91789	166.7384
ADE	0.09065	0.53415	0.08048	0.81764	161.8059
RADE	0.09037	0.53280	0.08501	0.76397	159.6820
MPSE	0.08974	0.52975	0.09540	0.63192	154.9529
MLE	0.08912	0.52677	0.10576	0.50198	150.4155
PCE	0.08876	0.52502	0.11194	0.43022	147.7858
WLSE	0.08864	0.52441	0.11409	0.40668	146.8857
Myelogenous leukaemia data					
LSE	0.05126	0.40694	0.15310	0.62650	46.5452
CVME	0.05134	0.40740	0.15607	0.60290	45.9891
ADE	0.05097	0.40519	0.15287	0.62839	48.7649
RADE	0.05035	0.40149	0.16740	0.51505	54.0284
MPSE	0.05039	0.40173	0.16647	0.52203	53.6636
MLE	0.05084	0.40445	0.15580	0.60506	49.7513
PCE	0.05005	0.39975	0.17409	0.46570	56.8095
WLSE	0.05295	0.41691	0.20579	0.27003	36.3343
Survival times of head and neck cancer data					
LSE	0.17077	0.97535	0.09639	0.78360	195.1669
CVME	0.17082	0.97567	0.09654	0.78208	194.9109
ADE	0.16858	0.96299	0.10790	0.658816	205.4386
RADE	0.16905	0.96563	0.10390	0.70319	203.1994
MPSE	0.16578	0.94712	0.13206	0.40611	219.5054
MLE	0.16426	0.93855	0.14511	0.29614	227.5379
PCE	0.15643	0.89414	0.21237	0.03516	274.8875
WLSE	0.17008	0.97150	0.09499	0.797100	198.3105

 DOI: http://dx.doi.org/10.18187/pjsor.v17i4.3781

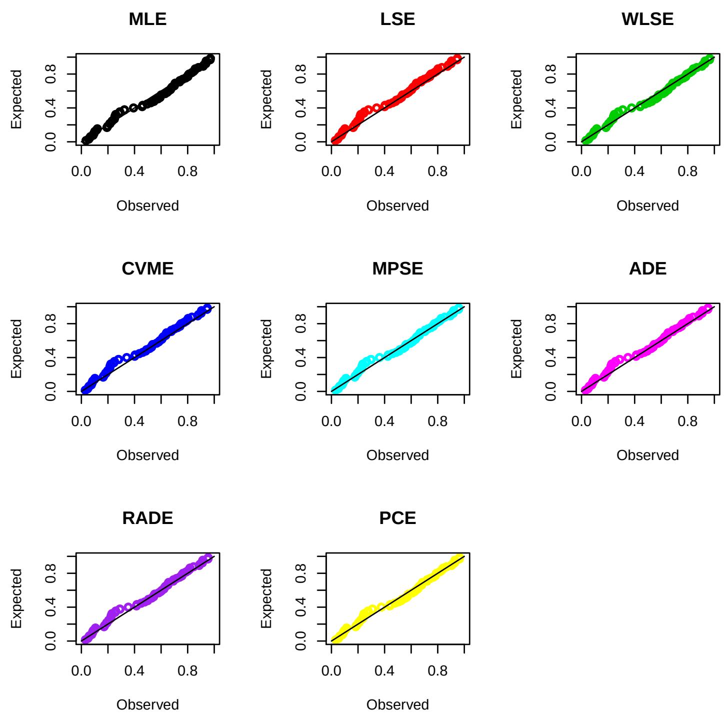

Figure 5: The PP plots of various estimation methods for COVID-19 data.

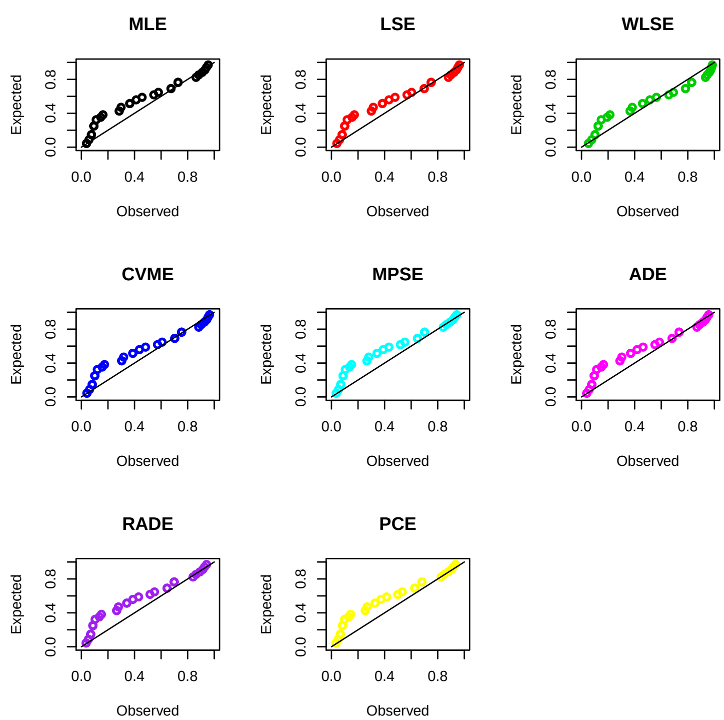

Figure 6: The PP plots of various estimation methods for myelogenous leukaemia data.

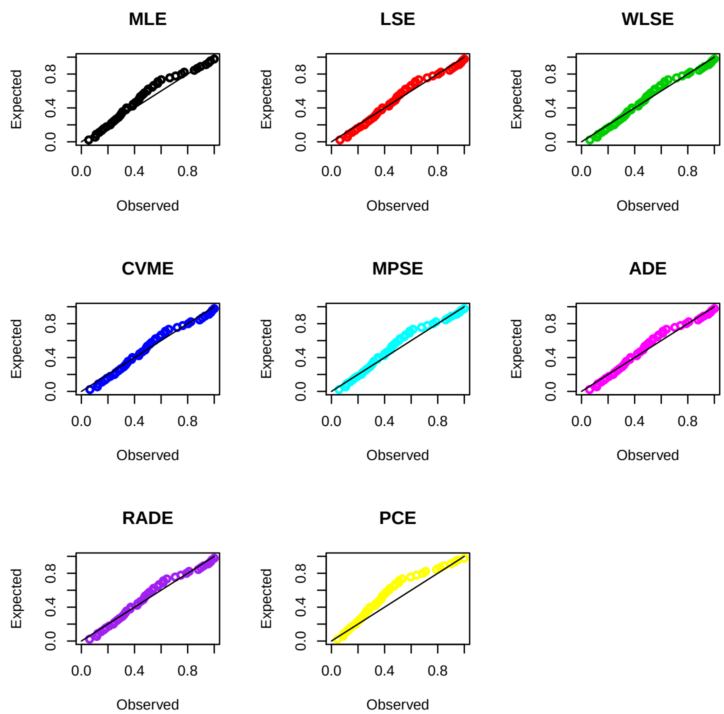

Figure 7: The PP plots of various estimation methods for survival times of head and neck cancer data.

8. Conclusions

In this paper, we propose and study a new discrete distribution which has a log-concave probability mass function and an increasing discrete hazard rate function, for all choices of its parameter. The new distribution is called one parameter discrete (OPD) distribution and it can be used effectively in analyzing over-dispersed count data. Several statistical and reliability properties are derived in closed forms, including the mode, moments, index of dispersion, mean residual life, mean past life, order statistics and L-moment statistics. Further, we derive closed form expressions for the two risk measures of the OPD distribution. The numerical computations for the VaR and TVaR measures for different parametric values of θ showed that these measures are increasing functions of θ. Eight classical estimators the parameter θ are proposed. Simulation results to explore the behavior of these estimators are conducted. Based on our study, the eight classical methods of estimation can be used effectively to estimate the OPD parameter θ. Three over-dispersed real data sets from the medicine field are used to validate the use of OPD distribution in fitting lifetime count data.

Acknowledgments: This publication was supported by the Deanship of Scientific Research at Prince Sattam bin Abdulaziz University, Alkharj, Saudi Arabia.

References

1. Al-Babtain, A. A., Ahmed, A. H. N. and Afify, A. Z. (2020). A new discrete analog of the continuous Lindley distribution, with reliability applications. Entropy, 22, 603.
2. Al-Babtain, A. A., Gemeay, A. M. and Afify, A. Z. (2021). Estimation methods for the discrete Poisson-Lindley and discrete Lindley distributions with actuarial measures and applications in medicine. Journal of King Saud University – Science, 33, 101224.
3. Aljohani, H. M., Akdoğan, Y., Cordeiro, G. M. and Afify, A. Z. (2021). The uniform Poisson–Ailamujia distribution: actuarial measures and applications in biological science. Symmetry, 13, p.1258.
4. Almazah, M. M. A., Alnssyan, B., Ahmed, A. H. N. and Afify, A.Z. (2021a). Reliability properties of the ndl family of discrete distributions with its inference. Mathematics, 9, p.1139. 5. Almazah, M. M. A., Erbayram,

T., Akdoğan, Y., AL Sobhi, M. M. and Afify, A.Z. (2021b). A new extended geometric distribution: properties, regression model, and actuarial applications. Mathematics, 9, p.1336. 6. Artzner, P. (1999). Application of coherent risk measures to capital requirements in insurance. N. Am. Actuar. J., 3, 11-25.

7. Eliwa, M. S., Alhussain, Z. A., El-Morshedy, M. (2020a). Discrete Gompertz-G family of distributions for over- and under-dispersed data with properties, estimation and applications. Mathematics, 8, p.358.

8. Eliwa, M. S., Altun, E., El-Dawoody, M. and El-Morshedy, M. (2020b). A new three-parameter discrete distribution with associated INAR (1) process and applications. IEEE access, 8, 91150-91162.

9. Eliwa, M. S., El-Morshedy, M. and Altun, E. (2020c). A study on discrete Bilal distribution with properties and applications on integer-valued autoregressive process. Revstat statistical journal. Forthcoming.

10. El-Morshedy, M., Eliwa, M. S. and Altun, E. (2020a). Discrete Burr-Hatke distribution with properties, estimation methods and regression model. IEEE access, 8, 74359-74370.

11. El-Morshedy, M., Eliwa, M. S. and Nagy, H. (2020b). A new two-parameter exponentiated discrete Lindley distribution: properties, estimation and applications. Journal of applied statistics, 47, 354-375.

12. Feigl, P. and Zelen, M. (1965). Estimation of exponential probabilities with concomitant information. Biometrics, 21, 826–838.

13. Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. Test, 19, 399–415.

14. Gupta, P. L., Gupta, R. C., and Tripathi, R. C. (1997). On the monotonic properties of discrete failure rates. Journal of statistical planning and inference, 65, 255-268.

15. Jazi, M. A., Lai, C. D., Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. Stat. Methodol., 7, 121-132.

16. Klugman, S. A., Panjer, H. H. and Willmot, G. E. (2012). Loss Models: From Data to Decisions. John Wiley & Sons: Hoboken, NJ, USA, Volume 715.

17. Krishna, H. and Pundir, P. S. (2009) Discrete Burr and discrete Pareto distributions. Stat. Methodol., 6, 177-188.

18. Ramos, P. L. and Louzada, F. (2019). A distribution for instantaneous failures. Stats, 2, 247-258.