Filomat 37:18 (2023), 6177–6203
https://doi.org/10.2298/FIL2318177E

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Modelling dispersed count data under various shapes of failure rates: A discrete probability analogue of odd Lomax generator

Mohamed. S. Eliwa[a,b,d], Mahmoud El-Morshedy[c,d], Hend S. Shahen[d]

[a]*Department of Statistics and Operation Research, College of Science, Qassim University, Buraydah 51482, Saudi Arabia*
[b]*Department of Mathematics, International Telematic University Uninettuno, I-00186 Rome, Italy*
[c]*Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia*
[d]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt*

Abstract. In this article, we introduce a discrete analogue of odd Lomax generator of distributions. The new discrete class can be utilized as a probabilistic tool to generalize any discrete baseline model. After proposing the new class, two special discrete models are investigated and discussed in detail. Some mathematical and statistical properties including, probability mass function, hazard rate function, quantile, crude moments, index of dispersion, entropies, order statistics, and L-moment statistics, are derived. It is found that the presented discrete class can be used to model symmetric and asymmetric data under different types of kurtosis shapes. It can be utilized to explain and analyze overdispersion data with extreme, zero-inflated or outliers' observations. Furthermore, it can be applied to discuss various shapes of hazard rates including monotone increasing, monotone decreasing, unimodal, bathtub, unimodal-bathtub, among others. We discuss the estimation of the class parameters by the maximum likelihood approach. The performance of the estimation method is tested via Markov chain Monte Carlo (MCMC) simulation technique. Finally, to demonstrate the proposed methodology in a real-life scenario, three real data sets are considered to show the applicability of the proposed generator.

1. Introduction

Data analysis is crucial, especially, in many applied sectors, and it has become very complex due to the diversity of data that is being generated day in and day out. Thus, many statisticians strive to create and propose different probability distributions for modeling such data. The Lomax (Lo) "shifted Pareto Type II" distribution is one of the most utilized models for this purpose, and it has many applications in a variety of fields like income and inequality of wealth, engineering, actuarial science, and biological sciences. Due to its flexibility in data modeling, many statisticians developed and studied in detail some generalizations of the Lo distribution. For instance, Marshall-Olkin Lo distribution by [1], transmuted Lo distribution by [2], weighted Lo distribution by [3], power Lo distribution by [4], Gompertz Lo distribution by [5], Weibull generalized Lo distribution by [6], inverse power Lo distribution by [7], among others.

2020 *Mathematics Subject Classification*. Primary 62E99; Secondary 62E15
Keywords. Discrete probability G-family; Hazard rate function; Dispersion phenomena; Extreme count data; Simulation; Consistent estimators.
Received: 13 September 2022; Revised: 14 November 2022; Accepted: 27 December 2022
Communicated by Biljana Popović
Email addresses: mseliwa@mans.edu.eg (Mohamed. S. Eliwa), m.elmorshedy@psau.edu.sa; mah_elmorshedy@mans.edu.eg (Mahmoud El-Morshedy), hendsalah@mans.edu.eg (Hend S. Shahen)

Recently, [8] proposed the odd Lo generator (OLoG) for positive continuous random variable (RV) based on the $T-X$ family pioneered by [9]. The survival function (SF) of the OLoG and its corresponding probability density function (PDF) can be formulated as

$$V(x;\varepsilon,\vartheta,\Lambda)=\vartheta^{\varepsilon}\left[\vartheta+\frac{Y(x;\Lambda)}{1-Y(x;\Lambda)}\right]^{-\varepsilon};\ x>0 \tag{1}$$

and

$$v(x;\varepsilon,\vartheta,\Lambda)=\frac{\varepsilon\vartheta^{\varepsilon}y(x;\Lambda)}{[1-Y(x;\Lambda)]^{2}}\left[\vartheta+\frac{Y(x;\Lambda)}{1-Y(x;\Lambda)}\right]^{-(\varepsilon+1)};\ x>0, \tag{2}$$

respectively, where $\varepsilon>0$ is shape parameter, $\vartheta>0$ is as a scale parameter, $Y(x;\Lambda)$ is the cumulative distribution function (CDF) of the baseline model with vector parameters Λ, and $y(x;\Lambda)=\frac{d}{dx}Y(x;\Lambda)$. In several cases, data need to be recorded on a discrete scale rather than on a continuous analogue such as the number of failures for a specific engineering system per week, number of daily COVID-19, number of fires in a year for any country, number of cysts of kidneys using steroids, studying the multi-system inflammatory syndrome in children (MISC) associated with COVID-19, where MISC is a condition where various body parts can become inflamed, including the brain, heart, eyes, lungs, or skin, among others. Thus, discretization of existing continuous distributions has received wide attention. Discretization approach can be applied via many techniques such as SF, Poisson-PDF, and binomial-PDF "as an example" based on specific mathematical statistical rules. Herein, we focus on the survival discretization (SD) method. According to the SD approach, the probability mass function (PMF) can be expressed as

$$\Pr(X=x)=V(x)-V(x+1);\ x=0,1,2,3,.... \tag{3}$$

Due to the flexibility of this method in creating discrete probability models, many statisticians have used this approach to present multiple PMFs for different purposes based on various continuous models. For instance, discrete Burr and discrete Pareto distribution by [10], discrete inverse Rayleigh model by [11], exponentiated discrete Weibull distribution by [12], three-parameter discrete Lindley model by [13], discrete Gompertz-G family by [14], discrete odd Weibull-G family by [15], discrete Ramos-Louzada distribution by [16], discrete Marshall-Olkin inverse Toppe-Leone model by [17], among others. Although there are number of discrete probability distributions in the statistical literature, there is still a lot of space left to derive new discretized distribution that is suitable under various conditions. Thus, in this paper, we introduce a discrete analogue of the OLoG based on SD method under the abbreviation, "DOLo-Y". Our motivations to introduce the DOLo-Y class can be summarized as follows:

- It can be used to discuss symmetric and asymmetric count data as well as zero-inflated count observations.
- It can be applied to model extreme and outlier count observations.
- It can be utilized as a suitable probabilistic tool to explain and study scattering "over (under) dispersion" phenomena.
- It can be used to study different types of kurtosis shapes including leptokurtic- and platykurtic-shaped.
- It can be applied to generate new flexible discrete models based on the baseline distribution.
- It can be utilized to analyze different forms of failure rates including monotone increasing, monotone decreasing, unimodal, bathtub, unimodal-bathtub, among others.
- It can be used for heavy tail real data modeling.
- The special sub-models of this class can be utilized in analyzing and modelling censored data because of the simplicity of its functions.

The rest of the paper is organized as follows, in Section 2, we introduce the DOLo-Y class. Various distributional statistics are derived in Section 3. In Section 4, two sub-models are discussed in detail. The DOLo-Y class parameters are estimated by utilizing the maximum likelihood method in Section 5. A simulation study is reported in Section 6. Three real data sets are analyzed to show the ability and notability of the DOLo-Y class in Section 7. Finally, Section 8 provides some conclusions.

2. Structure of the DOLo-Y class

Based on the SF of the OLoG and SD approach, the RV X is said to have the DOLo-Y class if its CDF can be expressed as

$$F(x;\varepsilon,\vartheta,\Lambda)=1-\vartheta^{\varepsilon}\left[\vartheta+\frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)}\right]^{-\varepsilon};\ x\in\mathbb{N}_0, \tag{4}$$

where $\vartheta>0$ is a scale parameter, $\varepsilon>0$ is a shape parameter, and $\mathbb{N}_0=0,1,2,3,....$ The corresponding PMF of Equation (4) can be proposed as

$$f(x;\varepsilon,\vartheta,\Lambda)=\vartheta^{\varepsilon}\left(\left[\vartheta+\frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)}\right]^{-\varepsilon}-\left[\vartheta+\frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)}\right]^{-\varepsilon}\right);\ x\in\mathbb{N}_0. \tag{5}$$

The hazard rate function (HRF) and its reversed can be formulated as

$$h(x;\varepsilon,\vartheta,\Lambda)=1-\left[\frac{\vartheta+\frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)}}{\vartheta+\frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)}}\right]^{-\varepsilon};\ x\in\mathbb{N}_0 \tag{6}$$

and

$$r(x;\varepsilon,\vartheta,\Lambda)=\frac{\left(\vartheta+\frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)}\right)^{-\varepsilon}-\left(\vartheta+\frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)}\right)^{-\varepsilon}}{\vartheta^{-\varepsilon}-\left(\vartheta+\frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)}\right)^{-\varepsilon}};\ x\in\mathbb{N}_0,$$

respectively, where $h(x;\varepsilon,\vartheta,\Lambda)=\frac{f(x;\varepsilon,\vartheta,\Lambda)}{1-F(x-1;\varepsilon,\vartheta,\Lambda)}$ and $r(x;\varepsilon,\vartheta,\Lambda)=\frac{f(x;\varepsilon,\vartheta,\Lambda)}{F(x;\varepsilon,\vartheta,\Lambda)}$.

3. Distributional properties of the DOLo-Y class

3.1. Quantile function (QF)

In this segment, the QF of the DOLo-Y class is derived. The QF can be utilized in various fields, in particular, the hydrology field "determining the level of lakes and oceans". Furthermore, the QF can be used in generating random samples for simulation objective, among others. The qth QF, say x_q, of the DOLo-Y class is the solution of $F(x_q;\varepsilon,\vartheta,\Lambda)-q=0;\ x_q>0$, then

$$x_q=Y^{-1}\left[\frac{1}{\left[\vartheta\left((1-q)^{\frac{-1}{\varepsilon}}-1\right)\right]^{-1}+1}\right]-1, \tag{7}$$

where $q\in(0,1)$ and Y^{-1} indicates to the baseline QF. Putting $q=0.5$, we get the median of the DOLo-Y class.

3.2. Crude moments and related statistical concepts

In this part, the crude or raw moments (RMS) of the DOLo-Y class, and some linked statistical measures, are derived. RMS are commonly utilized to describe the property of the probability model. The RMS can tell us much about data based on some descriptive measures including mean "$E(X)$", variance "$Var(X)$", skewness "$Sk(X)$", kurtosis "$Ku(X)$, and index of dispersion "$IoD(X)$". Thus, the statisticians are ready to compare various datasets using the primary four statistical RMS. Moreover, RMS can be used as a tool for statistical estimation and testing of hypotheses. Let the RV X have the DOLo-Y class, then the RMS can be expressed as

$$\begin{aligned}\mu_r^{'} &= E\left(X^r\right) = \sum_{x=1}^{\infty} x^r f\left(x;\varepsilon,\vartheta,\Lambda\right);\ r=1,2,3,\ ... \\ &= \sum_{x=1}^{\infty}\left[x^r-(x-1)^r\right]\vartheta^{\varepsilon}\left[\vartheta+\frac{Y\left(x;\Lambda\right)}{\overline{Y}\left(x;\Lambda\right)}\right]^{-\varepsilon}. \end{aligned} \tag{8}$$

The moment generating function (MGF) can be reported as

$$M_X\left(s\right) = E\left(e^{Xs}\right) = \sum_{x=0}^{\infty}\sum_{j=0}^{\infty}\frac{(xs)^j}{j!}\vartheta^{\varepsilon}\left(\left[\vartheta+\frac{Y\left(x;\Lambda\right)}{\overline{Y}\left(x;\Lambda\right)}\right]^{-\varepsilon}-\left[\vartheta+\frac{Y\left(x+1;\Lambda\right)}{\overline{Y}\left(x+1;\Lambda\right)}\right]^{-\varepsilon}\right);\ x\in\mathbb{N}_0. \tag{9}$$

The RMS can be derived from MGF where $E\left(X^r\right) = \frac{d^r}{ds^r}M_X\left(s\right)|_{s=0}$. Using Equation (8), the $E(X)$, $Var(X)$, $Sk(X)$, and $Ku(X)$ can be respectively given by

$$E(X) = \sum_{x=1}^{\infty}\vartheta^{\varepsilon}\left[\vartheta+\frac{Y\left(x;\Lambda\right)}{\overline{Y}\left(x;\Lambda\right)}\right]^{-\varepsilon}, \tag{10}$$

$$Var(X) = \sum_{x=1}^{\infty}\left(2x-1\right)\vartheta^{\varepsilon}\left[\vartheta+\frac{Y\left(x;\Lambda\right)}{\overline{Y}\left(x;\Lambda\right)}\right]^{-\varepsilon}-\mu_1^{'2}, \tag{11}$$

$$Sk(X) = \frac{\mu_3^{'}-3\mu_2^{'}\mu_1^{'}+2\mu_1^{'3}}{\left[Var(X)\right]^{3/2}},\ \text{and}\ Ku(X) = \frac{\mu_4^{'}-4\mu_1^{'}\mu_3^{'}+6\mu_2^{'}\mu_1^{'2}-3\mu_1^{'4}}{\left[Var(X)\right]^2}. \tag{12}$$

An important additional descriptive statistical tool called the IoD or relative variance or $Var(X)$-to-$E(X)$ ratio can be used effectively to discuss the actuarial data. The IoD of the DOLo-Y family can be formulated as

$$IoD(X) = \frac{\sum_{x=1}^{\infty}\left(2x-1\right)\vartheta^{\varepsilon}\left[\vartheta+\frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)}\right]^{-\varepsilon}}{\sum_{x=1}^{\infty}\vartheta^{\varepsilon}\left[\vartheta+\frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)}\right]^{-\varepsilon}}-\sum_{x=1}^{\infty}\vartheta^{\varepsilon}\left[\vartheta+\frac{Y\left(x;\Lambda\right)}{\overline{Y}\left(x;\Lambda\right)}\right]^{-\varepsilon};\ x\in\mathbb{N}_0. \tag{13}$$

The IoD can be used to select the data type including over dispersion "$IoD(X) > 1$", under dispersion "$IoD(X) < 1$", and equal dispersion "$IoD(X) = 1$".

3.3. Entropies

In this segment, various types of entropy (Ey) are discussed like Rényi Ey "REy", Shannon Ey "SnEy", Collision Ey "CoEy", min-Ey "MEy", and max-Ey "XEy". All of these types of Ey can be utilized in information theory because it is a measure of uncertain variability, for more detail (see, [18]). Assume the RV X have the DOLo-Y class, then the REy can be formulated as

$$I_\rho\left(X\right) = \frac{1}{1-\rho}\left[\rho\varepsilon\log\vartheta+\log\sum_{x=0}^{\infty}\left(\left[\vartheta+\frac{Y\left(x;\Lambda\right)}{\overline{Y}\left(x;\Lambda\right)}\right]^{-\varepsilon}-\left[\vartheta+\frac{Y\left(x+1;\Lambda\right)}{\overline{Y}\left(x+1;\Lambda\right)}\right]^{-\varepsilon}\right)^{\rho}\right];\ x\in\mathbb{N}_0, \tag{14}$$

where $\rho \in (0,1)$ and $\rho \neq 1$. The REy generalizes the SnEy, CoEy, MEy, and XEy where the SnEy, CoEy, MEy, and XEy can be calculated as a special case of the REy when $\rho \longrightarrow 1$, $\rho \longrightarrow 2$, $\rho \longrightarrow \infty$, and $\rho \longrightarrow 0$, respectively. As an example, the SnEy can be expressed as

$$I(X) = -\sum_{x=0}^{\infty} \vartheta^{\varepsilon} \left(\left[\vartheta + \frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)} \right]^{-\varepsilon} - \left[\vartheta + \frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)} \right]^{-\varepsilon} \right) \times \log \left(\vartheta^{\varepsilon} \left(\left[\vartheta + \frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)} \right]^{-\varepsilon} - \left[\vartheta + \frac{Y(x+1;\Lambda)}{\overline{Y}(x+1;\Lambda)} \right]^{-\varepsilon} \right) \right); \; x \in \mathbb{N}_0. \tag{15}$$

3.4. L-moment statistic

In statistics, the nth order statistic (OC) of a statistical sample is equal to its nth-smallest value. The OC, along with ranking statistics, are among the most basic tools in nonparametric statistics and inference. Given any random variables (RVs) $X_1, X_2, ..., X_n$, the OC $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ are also RVs. Let the RV X have the DOLo-Y family, then the CDF of the ith OC can be formulated as

$$\begin{aligned} F_{i:n}(x;\varepsilon,\vartheta,\Lambda) &= \sum_{k=i}^{n} \binom{n}{k} [F_i(x;\varepsilon,\vartheta,\Lambda)]^k [1 - F_i(x;\varepsilon,\vartheta,\Lambda)]^{n-k} \\ &= \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{m=0}^{k+j} \Psi_{(n,k)}^{(m,j)} [1 - F_i(x; m\varepsilon, \vartheta, \Lambda)], \end{aligned} \tag{16}$$

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

$$\begin{aligned} f_{i:n}(x;\vartheta) &= F_{i:n}(x;\varepsilon,\vartheta,\Lambda) - F_{i:n}(x-1;\varepsilon,\vartheta,\Lambda) \\ &= \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{m=0}^{k+j} \Psi_{(n,k)}^{(m,j)} \left(F_i(x-1; \varepsilon m, \vartheta, \Lambda) - F_i(x; m\varepsilon, \vartheta, \Lambda) \right). \end{aligned} \tag{17}$$

Thus, the rth RMS of $X_{i:n}$ can be reported as

$$E\left(X_{i:n}^r\right) = \sum_{x=0}^{\infty} \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{m=0}^{k+j} \Psi_{(n,k)}^{(m,j)} x^r \left(F_i(x-1; \varepsilon m, \vartheta, \Lambda) - F_i(x; m\varepsilon, \vartheta, \Lambda) \right). \tag{18}$$

Based on Equation (18), a significant descriptive statistical instrument called L-moments (L-MS) is derived. L-MS can be used to give summary statistics for probability distributions, for more detail (see, [19]). For the RV X, the L-MS statistics is given by

$$\Upsilon_{\varpi} = \frac{1}{\varpi} \sum_{i=0}^{\varpi-1} (-1)^i \binom{\varpi - 1}{i} E\left(X_{\varpi - i:\varpi}\right); \; \varpi = 1, 2, 3, ... \tag{19}$$

According to Equation (19), some statistical measures can be derived like mean, Sk, and Ku where mean $= \Upsilon_1$, Sk $= \frac{\Upsilon_3}{\Upsilon_2}$, and Ku $= \frac{\Upsilon_4}{\Upsilon_2}$.

4. Special sub-models of the DOLo-Y class

4.1. The DOLo-Geometric (DOLoGeo) distribution

Consider the RV X that follows the geometric (Geo) model with CDF $\Omega(x;\theta) = 1 - \theta^x$; $x > 0$ is $Y(x;\Lambda)$ function. Then, the PMF of the DOLoGeo model can be proposed as

$$f(x;\varepsilon,\vartheta,\theta)=\vartheta^{\varepsilon}\left[(\vartheta+\theta^{-x}-1)^{-\varepsilon}-\left(\vartheta+\theta^{-(x+1)}-1\right)^{-\varepsilon}\right];\ x\in\mathbb{N}_{0,} \tag{20}$$

where $\varepsilon, \vartheta > 0$ and $0 < \theta < 1$. For selected parameters values of the DOLoGeo model, some possible shapes for its PMF are sketched in Figure 1, the shape of the DOLoGeo distribution can be unimodal, symmetry, right skewed or zero-inflated. The behavior of the HRF of the DOLoGeo distribution is displayed graphically in Figure 2, and it can be either monotone increasing or decreasing shapes.

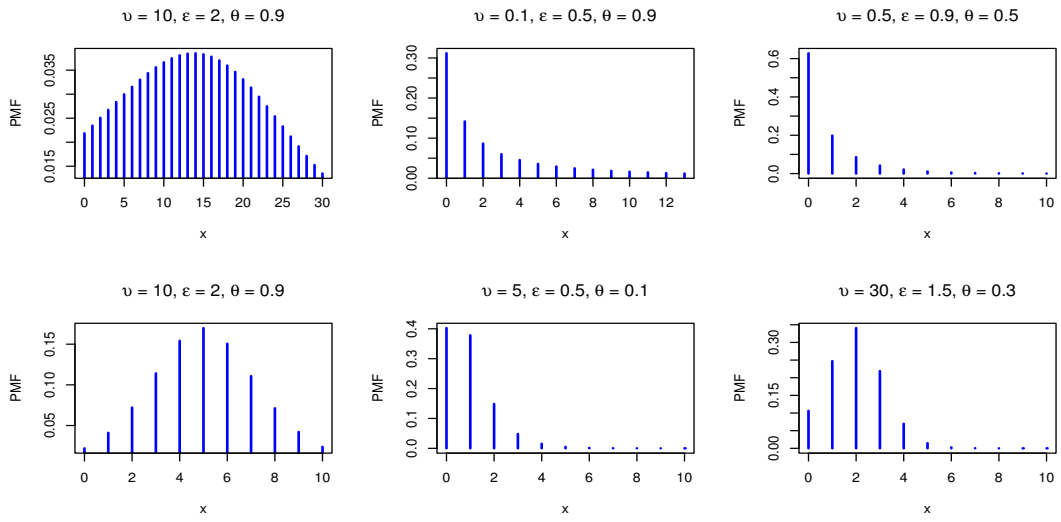

Figure 1. Various shapes for the PMF of the DOLoGeo model.

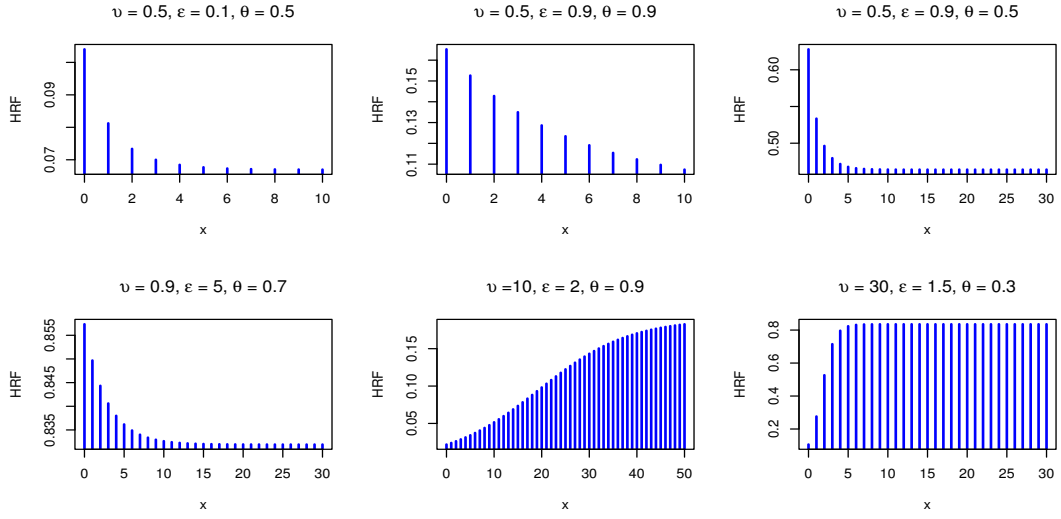

Figure 2. Various shapes for the HRF of the DOLoGeo model.

For chosen parameters values of the DOLoGeo distribution, some descriptive measures including $E(X)$, $Var(X)$, $IoD(X)$, $Sk(X)$ and $Ku(X)$ are listed in Table 1. Furthermore, some numerical computations of the Ey are reported in Tables 2 and 3.

Table 1. Numerical descriptive statistics of the DOLoGeo distribution.

Measure	$\varepsilon \downarrow \theta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	0.3	0.6282	1.0172	1.4638	2.0288	2.7982	3.9345	5.8116	9.505	18.12
$E(X)$	0.7	0.0848	0.1699	0.2783	0.4247	0.6338	0.9544	1.4998	2.611	5.977
	2	0.0005	0.0023	0.0066	0.0154	0.0327	0.0675	0.1425	0.330	1.002
	0.3	1.4802	3.1920	5.8631	10.301	18.216	33.817	69.701	173.9	514.2
$Var(X)$	0.7	0.118	0.2919	0.5797	1.0744	1.9757	3.7769	7.9605	20.72	92.63
	2	0.0005	0.0025	0.0075	0.0188	0.0440	0.1031	0.2581	0.772	3.858
	0.3	2.3563	3.1380	4.0055	5.0772	6.5098	8.5949	11.993	18.30	28.38
$IoD(X)$	0.7	1.3923	1.7181	2.0833	2.5299	3.1174	3.9576	5.3076	7.935	15.50
	2	1.0169	1.0608	1.1277	1.2201	1.3470	1.5282	1.8115	2.341	3.849
	0.3	2.8242	2.6896	2.6195	2.5747	2.5441	2.5224	2.4998	2.354	1.695
$Sk(X)$	0.7	5.0712	4.3347	3.972	3.7412	3.5789	3.4603	3.3740	3.315	3.178
	2	46.952	22.513	14.516	10.560	8.1976	6.6301	5.5279	4.748	4.252
	0.3	13.938	13.044	12.593	12.314	12.126	11.993	11.773	10.13	5.475
$Ku(X)$	0.7	36.177	28.288	24.724	22.584	21.148	20.139	19.431	18.96	16.99
	2	2277.9	561.28	252.61	144.79	94.358	66.654	49.999	39.71	33.87

Table 2. Numerical Ey of the DOLoGeo distribution at $\vartheta = 0.3$ and ε grows.

$\varepsilon \downarrow \theta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.3	0.7837	1.0413	1.2657	1.4885	1.7272	1.9997	2.3331	2.7834	3.5227
0.7	0.1828	0.3190	0.4593	0.6152	0.7975	1.0217	1.3145	1.7329	2.4516
1.5	0.0116	0.0369	0.0772	0.1373	0.2257	0.3576	0.5615	0.8999	1.5592
2	0.0021	0.0097	0.0260	0.0557	0.1071	0.1953	0.3495	0.6354	1.2491

Table 3. Numerical Ey of the DOLoGeo distribution at $\varepsilon = 0.3$ and ϑ grows.

$\vartheta \downarrow \theta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	0.9121	1.2087	1.4636	1.7127	1.9748	2.2679	2.6187	3.0823	3.8303
0.9	1.0725	1.4062	1.6855	1.9525	2.2279	2.5308	2.8886	3.3568	4.1076
2	1.2850	1.6371	1.9240	2.1950	2.4729	2.7774	3.1361	3.6049	4.3568
3.5	1.4010	1.7462	2.0306	2.3006	2.5779	2.8820	3.2405	3.7091	4.4619

According to Tables 1, 2 and 3, it is noted that the Geo model under DOLo-Y class can be used as a probability tool for modeling positively skewed data with leptokurtic shape. Moreover, it can be utilized to discuss over dispersion phenomena. For Ey, it is found that it can only take positive values.

4.2. The DOLo-Weibull (DOLoW) distribution

Consider the RV X that follows the Weibull (W) distribution with CDF $\Omega(x; a, b) = 1 - e^{-ax^b}$; $x > 0$ is $Y(x; \Lambda)$ function. Then, the PMF of the DOLoW distribution can be formulated as

$$f_X(x; \varepsilon, \vartheta, \lambda, b) = \vartheta^{\varepsilon} \left[\left(\vartheta + \lambda^{-x^b} - 1 \right)^{-\varepsilon} - \left(\vartheta + \lambda^{-(x+1)^b} - 1 \right)^{-\varepsilon} \right]; \; x \in \mathbb{N}_{0,} \tag{21}$$

where ε, ϑ, $b > 0$ and $0 < \lambda = e^{-a} < 1$. For specific parameters values of the DOLoW distribution, some possible sketches for its PMF are plotted in Figure 3, the shape of the DOLoW model can be unimodal, left skewed, right skewed, or zero-inflated. The attitude of the HRF of the DOLoW distribution is displayed in Figure 4, and it can take various shapes including monotone increasing, monotone decreasing, unimodal, bathtub and bathtub-unimodal shapes.

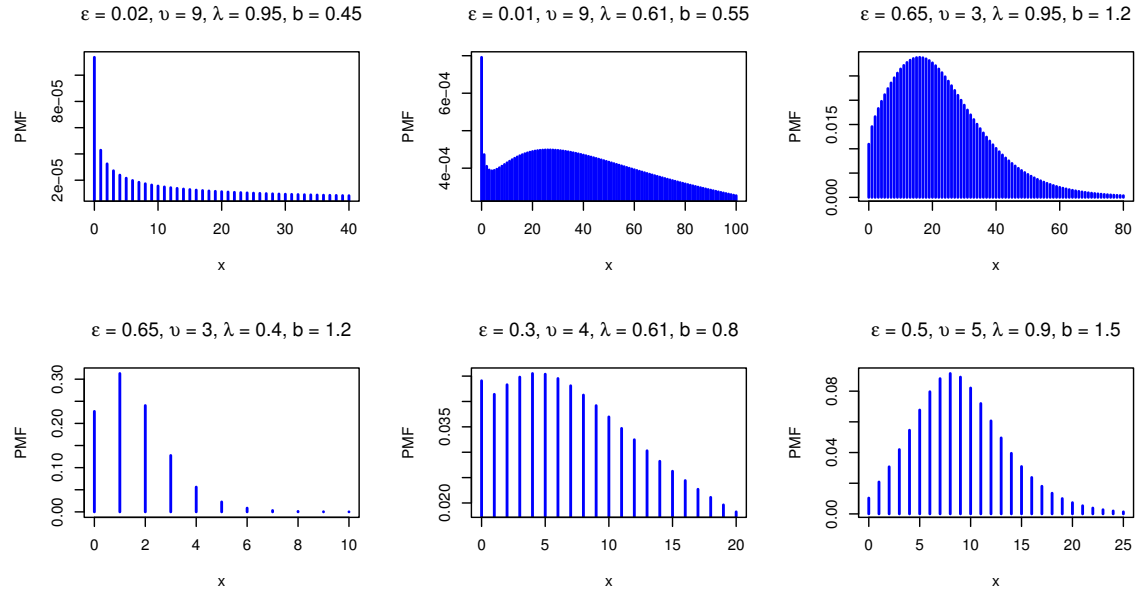

Figure 3. Various shapes for the PMF of the DOLoW model.

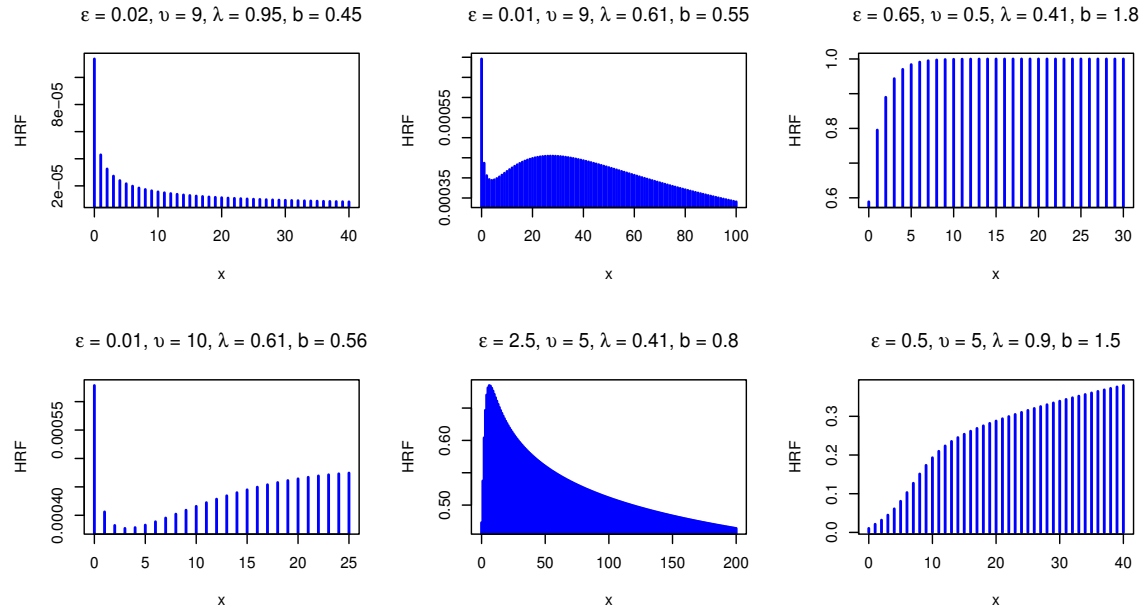

Figure 4. Various shapes for the HRF of the DOLoW model.

For selected parameters values of the DOLoW model, some numerical descriptive statistics and Ey results are reported in Tables 4, 5 and 6, respectively.

Table 4. Numerical descriptive statistics of the DOLoW distribution at $\vartheta = 0.3$ and $b = 0.2$.

Measure	$\lambda \rightarrow$ $\varepsilon \downarrow$	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
$E(X)$	**0.1**	0.52229	0.59719	0.67140	0.75151	0.84077	0.94481	1.07703	1.27031	1.63541
	0.2	0.25673	0.30387	0.34421	0.38782	0.44064	0.50820	0.59792	0.72796	0.97721
	0.5	0.03331	0.04994	0.06533	0.08144	0.09989	0.12325	0.15678	0.21344	0.33639
	0.9	0.00219	0.00454	0.00736	0.01094	0.01571	0.02254	0.03337	0.05393	0.11014
$Var(X)$	**0.1**	0.28120	0.33665	0.40460	0.47954	0.55935	0.64973	0.77124	0.96874	1.3860
	0.2	0.19132	0.21615	0.24263	0.27984	0.33314	0.40563	0.49981	0.63300	0.91391
	0.5	0.03220	0.04745	0.06108	0.07494	0.09070	0.11142	0.14375	0.20179	0.32412
	0.9	0.00219	0.00452	0.00731	0.01082	0.01546	0.02205	0.03245	0.05233	0.10638
$IoD(X)$	**0.1**	0.53840	0.56372	0.60262	0.63813	0.66529	0.68768	0.71608	0.76260	0.84749
	0.2	0.74522	0.71133	0.70488	0.72157	0.75603	0.79817	0.83591	0.86956	0.93522
	0.5	0.96669	0.95007	0.93487	0.92017	0.90793	0.90401	0.91689	0.94675	0.96520
	0.9	0.99781	0.99546	0.99264	0.98907	0.98439	0.97836	0.97223	0.97032	0.96588
$Sk(X)$	**0.1**	0.23006	0.35517	0.41170	0.38125	0.30072	0.21576	0.16266	0.12968	0.08962
	0.2	1.12279	0.92166	0.86666	0.88615	0.90866	0.87568	0.77319	0.65341	0.57879
	0.5	5.20107	4.13256	3.51930	3.07013	2.71145	2.42697	2.20115	1.94971	1.5299
	0.9	21.2909	14.7381	11.5236	9.40427	7.79064	6.44344	5.24405	4.11263	2.84794
$Ku(X)$	**0.1**	1.74328	2.26442	2.31214	2.13157	1.93909	1.8637	1.91831	1.96653	1.9889
	0.2	2.28799	2.04055	2.26551	2.63235	2.82671	2.71427	2.42068	2.26458	2.31446
	0.5	28.0511	18.0785	13.3925	10.4722	8.53891	7.40184	6.80394	6.00133	4.5727
	0.9	454.304	218.211	133.792	89.4413	61.7064	42.6011	28.8513	18.8232	10.3907

Table 5. Numerical Ey of DOLoW distribution at $\vartheta = 0.2$ and $b = 0.3$.

$\varepsilon \downarrow \lambda \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.2	1.2322	1.5485	1.8175	2.0831	2.3698	2.7042	3.1296	3.7419	4.8562
0.5	0.3161	0.4873	0.6525	0.8304	1.0363	1.2918	1.6358	2.1597	3.1707
2	0.0009	0.0045	0.0125	0.0278	0.0562	0.1089	0.2118	0.4347	1.0448

Table 6. Numerical Ey of DOLoW distribution at $\varepsilon = 0.3$ and $b = 0.2$.

$\vartheta \downarrow \lambda \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.3	0.8765	1.1866	1.4625	1.7428	2.0516	2.4170	2.8857	3.5603	4.7726
0.7	1.1979	1.6445	2.0425	2.4428	2.8754	3.3727	3.9860	4.8243	6.2286
3	2.0967	2.9061	3.5907	4.2378	4.8914	5.5910	6.3923	7.4070	8.9787

Regarding Tables 4, 5 and 6, it is found that DOLoW model is capable of modeling positively skewed and symmetric data under leptokurtic and platykurtic shapes. Further, it is appropriate for modelling under-dispersed phenomenon. For Ey, it is noted that it can only take positive values.

5. Classical estimation approach

In this segment, the estimation of the DOLo-Y class parameters is derived by using the method of the maximum likelihood (MxLk) based on a complete sample. Let $X_1, X_2, \ldots, X_n$ be a random sample from the DOLo-Y class. Then, the log-likelihood function (L) can be formulated as

$$L = n\varepsilon \ln(\vartheta) + \sum_{i=1}^{n} \ln\left([\vartheta + Q(x_i; \Lambda)]^{-\varepsilon} - [\vartheta + Q(x_i + 1; \Lambda)]^{-\varepsilon} \right). \tag{22}$$

Differentiating Equation (22) with respect to ϑ, ε and Λ, respectively, we get

$$\frac{\partial L}{\partial \vartheta} = \frac{n\varepsilon}{\vartheta} + \varepsilon \sum_{i=1}^{n} \frac{(\vartheta + Q(x_i+1;\Lambda))^{-\varepsilon-1} - (\vartheta + Q(x_i;\Lambda))^{-\varepsilon-1}}{(\vartheta + Q(x_i;\Lambda))^{-\varepsilon} - (\vartheta + Q(x_i+1;\Lambda))^{-\varepsilon}}, \tag{23}$$

$$\frac{\partial L}{\partial \varepsilon} = n \ln \vartheta + \sum_{i=1}^{n} \frac{(\vartheta + Q(x_i+1;\Lambda))^{-\varepsilon} \ln(\vartheta + Q(x_i+1;\Lambda)) - (\vartheta + Q(x_i;\Lambda))^{-\varepsilon} \ln(\vartheta + Q(x_i;\Lambda))}{(\vartheta + Q(x_i;\Lambda))^{-\varepsilon} - (\vartheta + Q(x_i+1;\Lambda))^{-\varepsilon}}, \tag{24}$$

$$\frac{\partial L}{\partial \Lambda_j} = \varepsilon \sum_{i=1}^{n} \frac{(\vartheta + Q(x_i+1;\Lambda))^{-\varepsilon-1} [Q(x_i+1;\Lambda)]_{\Lambda_j} - (\vartheta + Q(x_i;\Lambda))^{-\varepsilon-1} [Q(x_i;\Lambda)]_{\Lambda_j}}{(\vartheta + Q(x_i;\Lambda))^{-\varepsilon} - (\vartheta + Q(x_i+1;\Lambda))^{-\varepsilon}}, \tag{25}$$

where $Q(x_i;\Lambda) = \frac{Y(x;\Lambda)}{\overline{Y}(x;\Lambda)}$ and $[Q(x_i;\Lambda)]_{\Lambda_j} = \frac{\partial Q(x_i;\Lambda)}{\partial \Lambda_j}$; $j = 1, 2, ..., m$. Setting Equations (23)-(25) equal to zero, we can derive the normal equations (NE) of the MxLk technique. The NE cannot be solved analytically, therefore, an iterative procedure such as Newton-Raphson is required to solve it numerically.

6. The performance of the MxLk estimates

In this section, the performance of the MxLk estimators is tested based on various values of sample size n using a simulation technique "MCMC". The simulation results were carried out for the DOLoGeo and DOLoW distributions by the function **optim()** with the argument method = "**L-BFGS-B**" available by **R** software. We generated $N = 10000$ samples of five sizes, where $n = 20, 50, 100, 200, 500$ from the DOLoGeo and DOLoW distributions for various values of the model parameters as mentioned in Tables 7 and 8. The bias, mean squared error (MSE), and mean relative error (MRE) are the criteria of this simulation. The important results can be shown numerically in Tables 7-8, and graphically in Figures 5-8. As we can see, the bias, MSE, and MRE decrease when sample size n grows. Thus, the MxLk shows the property of consistency for all parameter combinations, and it can be used effectively in estimating the DOLoGeo and DOLoW parameters.

Table 7. Simulation results for the DOLoGeo parameters.

		$\varepsilon = 0.8, \vartheta = 0.5, \theta = 0.3$			$\varepsilon = 1.5, \vartheta = 1.3, \theta = 0.8$		
n	**Parameter**	**Bias**	**MSE**	**MRE**	**Bias**	**MSE**	**MRE**
20	ε	0.27563464	0.07498552	0.55452306	1.25312025	0.97455692	0.41812033
	ϑ	0.08432694	0.00729840	0.17102369	0.92385699	0.92389696	0.29012236
	θ	0.09896301	0.00974896	0.19798563	0.09230225	0.09032698	0.31002369
50	ε	0.15542369	0.02387420	0.31412369	0.85525631	0.73022363	0.28512336
	ϑ	0.05397103	0.00281039	0.10912369	0.59411520	0.59411290	0.19801233
	θ	0.05892013	0.00341891	0.11901236	0.05603236	0.05630337	0.11211475
100	ε	0.09586320	0.00912691	0.19912369	0.50390122	0.25553036	0.16803774
	ϑ	0.03412368	0.00123690	0.07123694	0.32636698	0.32611236	0.10901128
	θ	0.03423697	0.00136012	0.07036954	0.03332537	0.03302258	0.06512236
200	ε	0.06403147	0.00412369	0.12741236	0.38303398	0.15112369	0.12800123
	ϑ	0.02012397	0.00043965	0.04198305	0.23122547	0.23102334	0.07610257
	θ	0.02501396	0.00068012	0.05123690	0.02411452	0.02290336	0.04712336
300	ε	0.03120369	0.00095236	0.08630125	0.25033696	0.09974585	0.08366961
	ϑ	0.01850123	0.00011023	0.03741260	0.16365844	0.13695255	0.01236632
	θ	0.02298752	0.00051033	0.03941269	0.01225695	0.00745866	0.01239522
500	ε	0.02512369	0.00061203	0.06612698	0.11323632	0.00362455	0.00636633
	ϑ	0.01133694	0.00008498	0.02096841	0.08236995	0.00385962	0.00559662
	θ	0.01910369	0.00011976	0.02612361	0.00341258	0.00112581	0.00669211

Table 8. Simulation results for the DOLoW parameters.

		$\varepsilon = 0.8, \vartheta = 0.5, \lambda = 0.3, b = 0.7$			$\varepsilon = 1.5, \vartheta = 1.3, \lambda = 0.8, b = 1.5$		
n	**Parameter**	**Bias**	**MSE**	**MRE**	**Bias**	**MSE**	**MRE**
20	ε	0.37703625	0.14102368	0.25085631	0.42036991	0.17503269	0.20411233
	ϑ	0.34063265	0.11745963	0.68141415	0.32332688	0.10412596	0.28113022
	λ	0.13723032	0.01963017	0.27378952	0.13922303	0.02013691	0.28120223
	b	0.87213691	0.76103690	0.29108552	0.78033699	0.31202235	0.26122236
50	ε	0.23102367	0.05298621	0.15412397	0.25710326	0.06612236	0.17230365
	ϑ	0.20820369	0.04342103	0.41676903	0.20302363	0.04022236	0.13512366
	λ	0.08102369	0.00663691	0.16203695	0.09302236	0.00812236	0.18102366
	b	0.53541269	0.29203644	0.17901257	0.47710036	0.22712269	0.15911236
100	ε	0.12823603	0.01603697	0.08632024	0.14620373	0.02112533	0.09800369
	ϑ	0.12382669	0.01543036	0.25036941	0.11301478	0.01303362	0.07523013
	λ	0.05012369	0.00233696	0.09930369	0.05332361	0.00279128	0.10612669
	b	0.31123607	0.09741231	0.10410288	0.29803369	0.08811263	0.09911265
200	ε	0.09302395	0.00874109	0.06525663	0.10710306	0.01203005	0.07203303
	ϑ	0.08201746	0.00670394	0.16445632	0.08741033	0.00743415	0.05855255
	λ	0.03401395	0.00113126	0.06723955	0.03513954	0.00081256	0.07113961
	b	0.22803974	0.05201773	0.07650312	0.20745563	0.04303364	0.06911236
300	ε	0.08374019	0.00709641	0.04236102	0.07626888	0.00652887	0.01136955
	ϑ	0.04563036	0.00213103	0.11236996	0.02221305	0.00463025	0.02833632
	λ	0.02635441	0.00098512	0.04412320	0.01585552	0.00008545	0.00221478
	b	0.13695303	0.00369603	0.02364741	0.10141225	0.01120369	0.02366333
500	ε	0.04589120	0.00123120	0.01485230	0.02236944	0.00096366	0.00547899
	ϑ	0.01299749	0.00085636	0.07563214	0.00966369	0.00132002	0.00415569
	λ	0.01856309	0.00056361	0.02896301	0.00663303	0.00003237	0.00085996
	b	0.06362369	0.00072361	0.00725036	0.03614786	0.00133695	0.00700147

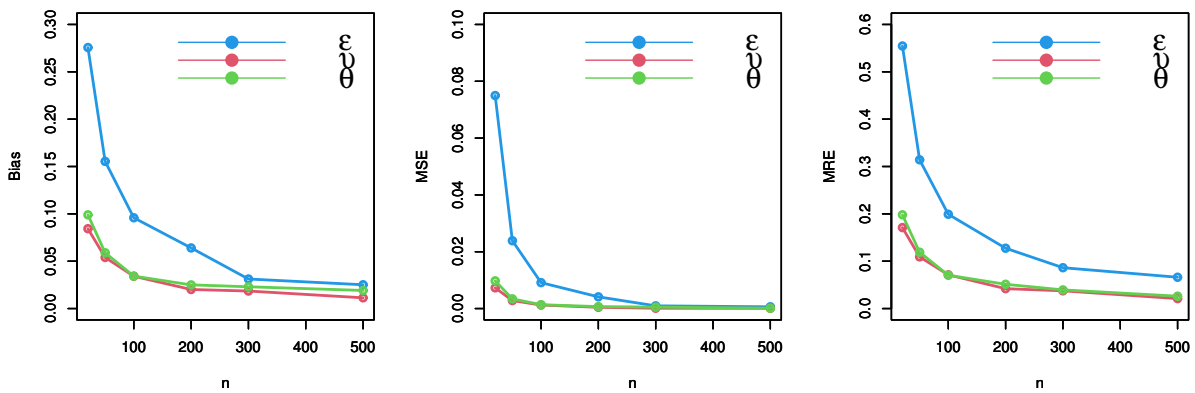

Figure 5. Simulation plots for the DOLoGeo parameters "$\varepsilon = 0.8$, $\vartheta = 0.5$, $\theta = 0.3$".

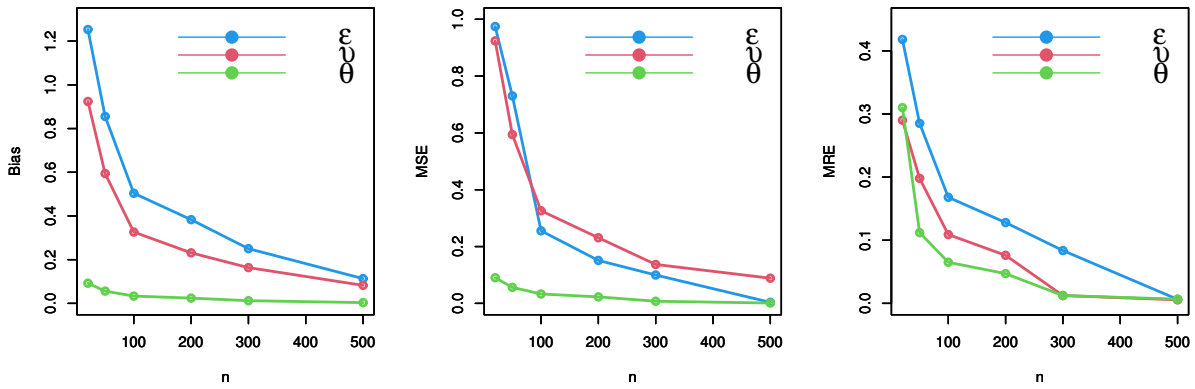

Figure 6. Simulation plots for the DOLoGeo parameters "$\varepsilon = 1.5$, $\vartheta = 1.3$, $\theta = 0.8$ ".

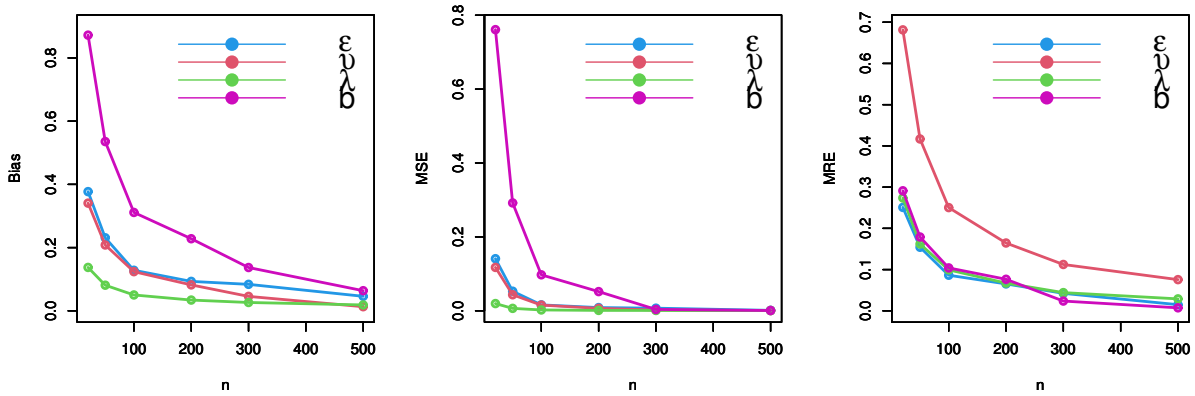

Figure 7. Simulation plots for the DOLoW parameters "$\varepsilon = 0.8$, $\vartheta = 0.5$, $\lambda = 0.3$, $b = 0.7$".

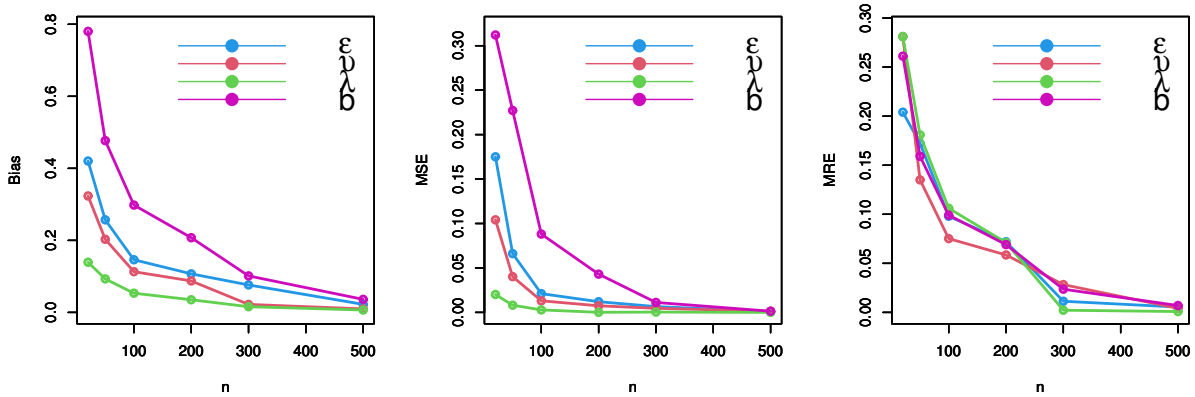

Figure 8. Simulation plots for the DOLoW parameters "$\varepsilon = 1.5$, $\vartheta = 1.3$, $\lambda = 0.8$, $b = 1.5$".

7. Data modeling with comparison study

In this section, we illustrate the applications of the DOLoGeo and DOLoW distributions using three real data sets. The fitted models are compared using some criteria, namely, the negative log-likelihood ($-L$), Akaike information criterion (AIC), corrected AIC (CAIC), Chi-square (χ^2) with degree of freedom (DF) and its p-value (P-V). We shall compare the DOLoGeo and DOLoW distributions with some competitive models described in Table 9.

Table 9. The competitive models.

Distribution	Abbreviation	Author(s)
One parameter discrete flexible	DFx-I	[20]
Discrete Lomax	DLo	[21]
Discrete log-logistic	DLogL	[22]
Poisson	Poi	[23]
Discrete Burr-Hatke	DBH	[24]
Discrete Burr	DB	[10]
Discrete Pareto	DPa	[10]
Geometric	Geo	-
Generalized geometric	GGeo	[25]
Discrete inverse Weibull	DIW	[26]
Discrete inverse Rayleigh	DIR	[11]
One parameter discrete Lindley	DLi-I	[27]
Two parameters discrete Lindley	DLi-II	[28]
Three parameters discrete Lindley	DLi-III	[13]
Discrete Rayleigh	DR	[29]
Discrete Weibull	DW	[30]
Exponentiated discrete Weibull	EDW	[12]
Discrete Gompertz-Weibull	DGzW	[14]
Discrete odd Weibull-inverse Weibull	DOWIW	[15]

7.1. Data set I: Cysts of kidneys (COK)

This data represents the counts of COK utilizing steroids (see, [17]). In this section, the MxLk estimators of the DOLoGeo parameters are calculated based on the COK data. Before analyzing the COK data, visualization of this data must be drawn to know its behavior. Figure 9 displays some non-parametric plots. As we can see, there are some extreme and outlier observations with positively skewed shape.

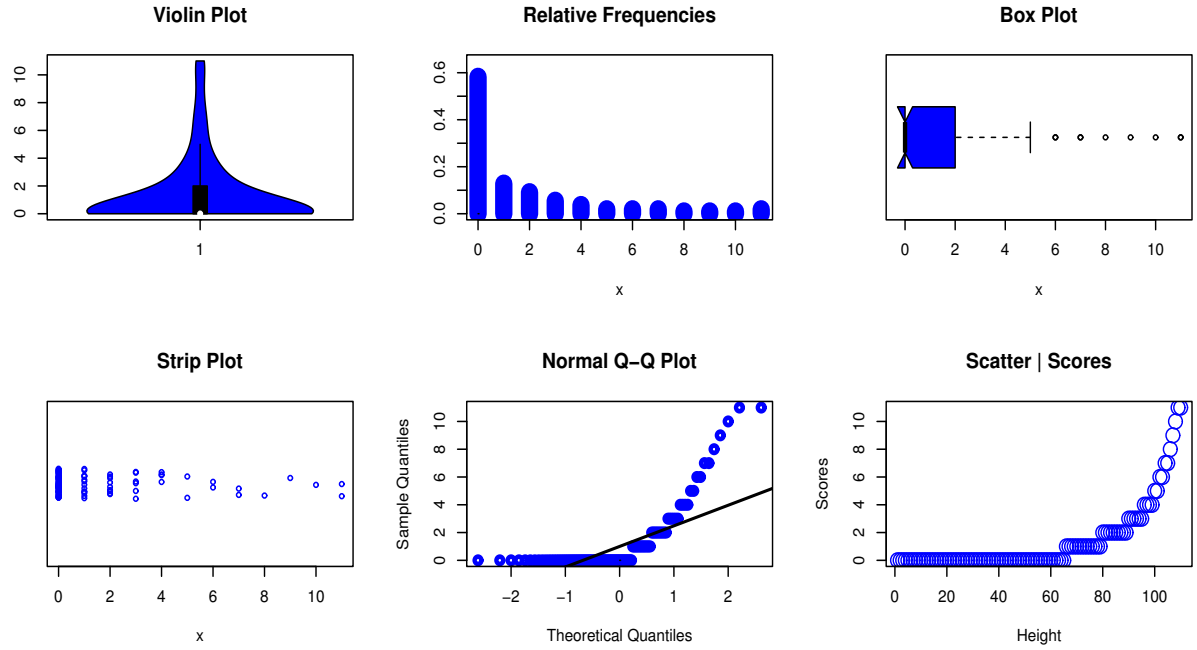

Figure 9. Non-parametric plots of the COK data.

The MxLk estimators with their corresponding standard errors (SDES) for the parameter(s) as well as goodness-of-fit test (GOFT) are presented in Tables 10, 11 and 12, respectively.

Table 10. The MxLk estimators and SDES for the COK data.

Parameter⟶	θ		ϑ		ε		γ	
Model ↓	MxLk	SDES	MxLk	SDES	MxLk	SDES	MxLk	SDES
DOLoGeo	0.282	0.489	0.084	0.167	0.259	0.134	–	–
DGzW	0.490	0.073	0.320	0.290	0.670	0.690	1.630	0.021
DLi-I	0.436	0.026	–	–	–	–	–	–
DFx-I	0.623	0.031	–	–	–	–	–	–
DPa	0.268	0.034	–	–	–	–	–	–
DLo	0.152	0.098	1.830	0.952	–	–	–	–
Poi	1.390	0.112	–	–	–	–	–	–
Geo	–	–	–	–	0.582	0.030	–	–
DR	–	–	–	–	0.901	0.009	–	–
DIR	–	–	–	–	0.554	0.049	–	–
GGeo	–	–	0.188	0.089	0.800	0.064	–	–
DIW	–	–	1.049	0.146	0.581	0.048	–	–
DLi-II	0.581	0.045	0.001	0.058	–	–	–	–
DLi-III	0.582	0.005	358.728	11863.370	0.001	20.698	–	–
DB	0.278	0.045	1.053	0.167	–	–	–	–
DLogL	0.780	0.136	1.208	0.159	–	–	–	–

Table 11. The GOFT of the COK data "part I".

	Observed	Expected Frequencies							
X	Frequencies	DOLoGeo	DGzW	DLi-I	DFx-I	DPa	DLo	Poi	Geo
0	65	64.995	64.242	40.286	45.256	65.842	61.615	27.398	45.980
1	14	14.413	15.441	29.834	29.094	18.267	21.023	38.084	26.760
2	10	8.869	9.180	18.357	16.508	8.164	9.687	26.468	15.575
3	6	6.137	6.073	10.336	8.893	4.513	5.275	12.264	9.064
4	4	4.372	4.201	5.523	4.703	2.820	3.197	4.262	5.275
5	2	3.1397	2.981	2.851	2.489	1.909	2.088	1.1850	3.070
6	2	2.259	2.154	1.437	1.335	1.368	1.441	0.274	1.787
7	2	1.627	1.561	0.711	0.731	1.022	1.038	0.054	1.039
8	1	1.172	1.140	0.347	0.409	0.789	0.773	0.009	0.605
9	1	0.844	0.831	0.167	0.234	0.626	0.592	0.001	0.352
10	1	0.608	0.612	0.079	0.137	0.506	0.463	0.000	0.205
11	2	1.564	1.584	0.070	0.209	4.172	2.806	0.002	0.286
Total	110	110	110	110	110	110	110	110	110
$-L$		167.026	167.021	189.110	182.288	171.192	170.481	246.210	178.767
AIC		340.052	342.062	380.220	366.575	344.384	344.961	494.420	359.533
CAIC		340.278	342.443	380.257	366.612	344.421	345.073	494.457	359.570
χ^2		0.569	0.567	34.635	31.702	3.430	3.238	89.277	19.109
DF		2	1	4	4	4	3	3	4
P-V		0.752	0.451	< 0.001	< 0.001	0.489	0.356	< 0.001	< 0.001

Table 12. The GOFT of the COK data "part II".

	Observed	Expected Frequencies							
X	Frequencies	DR	DIR	GGeo	DIW	DLi-II	DLi-III	DB	DLogL
0	65	10.889	60.888	62.738	63.910	46.026	46.008	64.743	63.192
1	14	26.618	33.989	19.665	20.699	26.768	26.765	19.177	20.101
2	10	29.448	8.123	9.439	8.053	15.568	15.570	8.484	8.644
3	6	22.296	3.004	5.436	4.234	9.054	9.058	4.632	4.656
4	4	12.629	1.420	3.463	2.599	5.266	5.268	2.863	2.864
5	2	5.539	0.779	2.349	1.754	3.062	3.066	1.920	1.921
6	2	1.914	0.473	1.663	1.261	1.781	1.783	1.365	1.368
7	2	0.526	0.308	1.213	0.949	1.036	1.037	1.013	1.019
8	1	0.116	0.212	0.904	0.739	0.602	0.604	0.777	0.786
9	1	0.020	0.152	0.685	0.592	0.350	0.351	0.613	0.623
10	1	0.003	0.112	0.525	0.485	0.204	0.204	0.494	0.504
11	2	0.004	0.538	1.918	4.723	0.282	0.284	3.917	4.321
Total	110	110	110	110	110	110	110	110	110
$-L$		277.778	186.547	168.655	172.935	178.767	178.767	171.139	171.717
AIC		557.556	375.094	341.310	349.869	361.534	363.533	346.278	347.43
CAIC		557.593	375.131	341.422	349.982	361.646	363.759	346.391	347.55
χ^2		306.515	40.456	2.444	6.445	19.091	19.096	2.587	4.033
DF		4	2	3	3	3	2	2	3
P-V		< 0.001	< 0.001	0.485	0.092	0.0003	< 0.001	0.274	0.258

As we can note, the DOLoGeo, DGzW, DPa, DLo, GGeo, DIW, DB and DLogL models work quite well for modeling the COK data, but the DOLoGeo distribution is the best among all competitive models. Figure 10 supports the results reported in Tables 11 and 12.

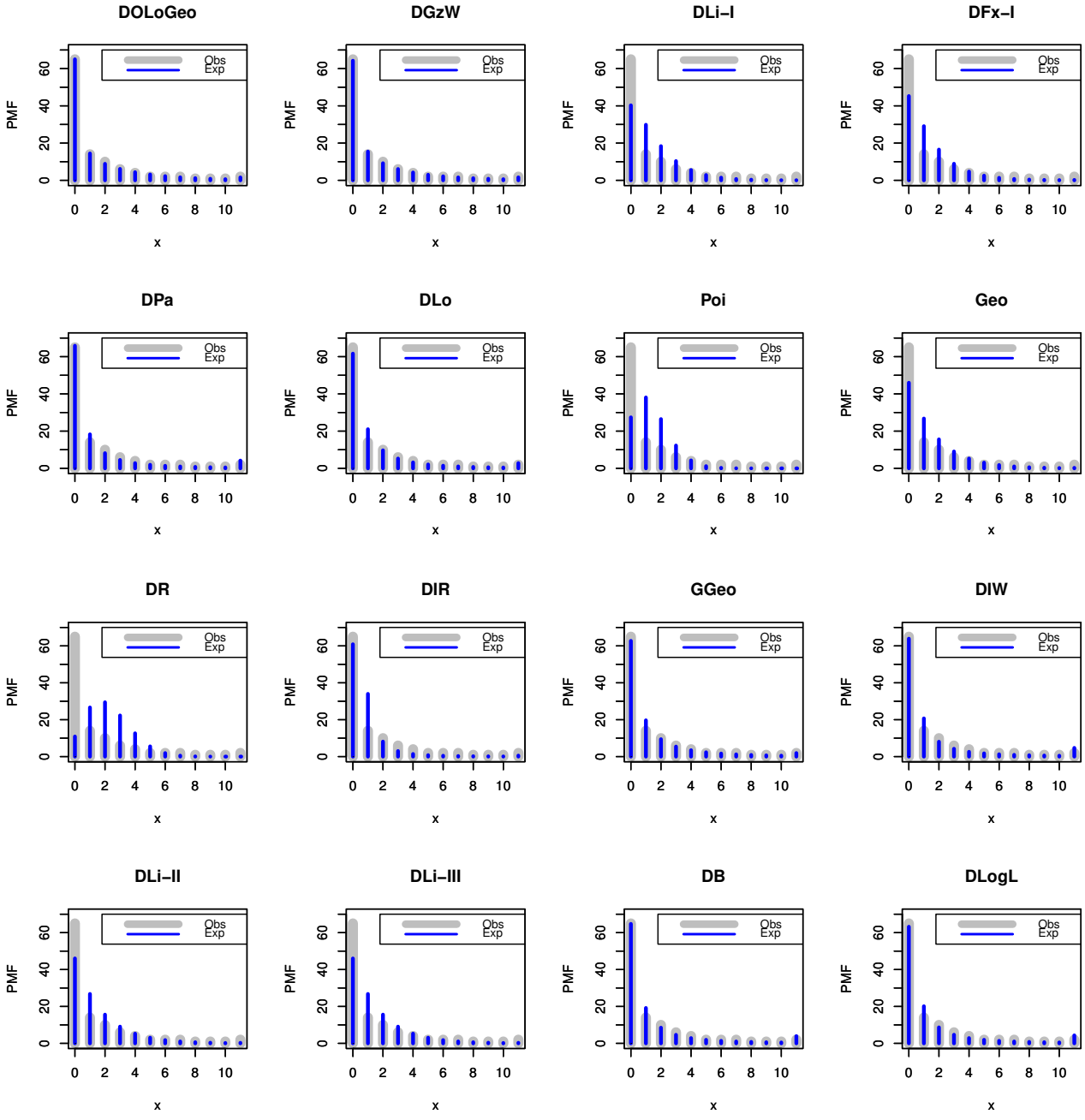

Figure 10. The observed and expected PMFs for the COK data.

Figure 11 shows the L profile of the parameters of the DOLoGeo model based on the COK data. It was found that the estimators are unique "unimodal-shaped".

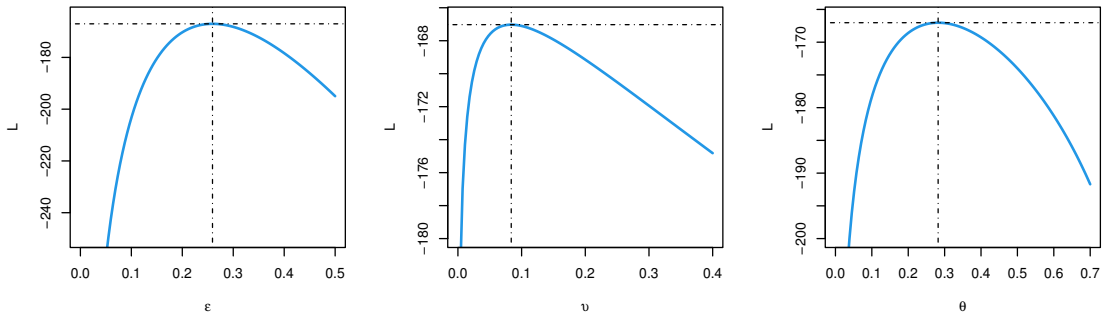

Figure 11. The L profile of the DOLoGeo parameters for the COK data.

Table 13 shows the difference between experimental and observational descriptive statistics, and the two measures are found to be approximately equal.

Table 13. Descriptive statistics of the COK data.

Model	Mean	Variance	Index of dispersion	Skewness	Kurtosis
Observed	1.39090	6.11184	4.39416	2.29259	8.17345
Empirical	1.39103	6.44578	4.63381	2.85099	14.2114

As we can see, the COK data is skewed to the right with leptokurtic shape. Moreover, it is suffering from over dispersion phenomena.

7.2. Data set II: COVID-19 in South Korea (COVID-19SK)

This data represents the daily new deaths of COVID-19SK from 15 February 2020 to 12 December 2020 (https://www.worldometers.info/coronavirus/country/south-korea/). In this subsection, the MxLk estimators of the DOLoW parameters are evaluated according to the COVID-19SK data. Figure 12 shows some non-parametric plots to discuss the behavior of this data. As we can note, there are some extreme observations with positively skewed-shaped.

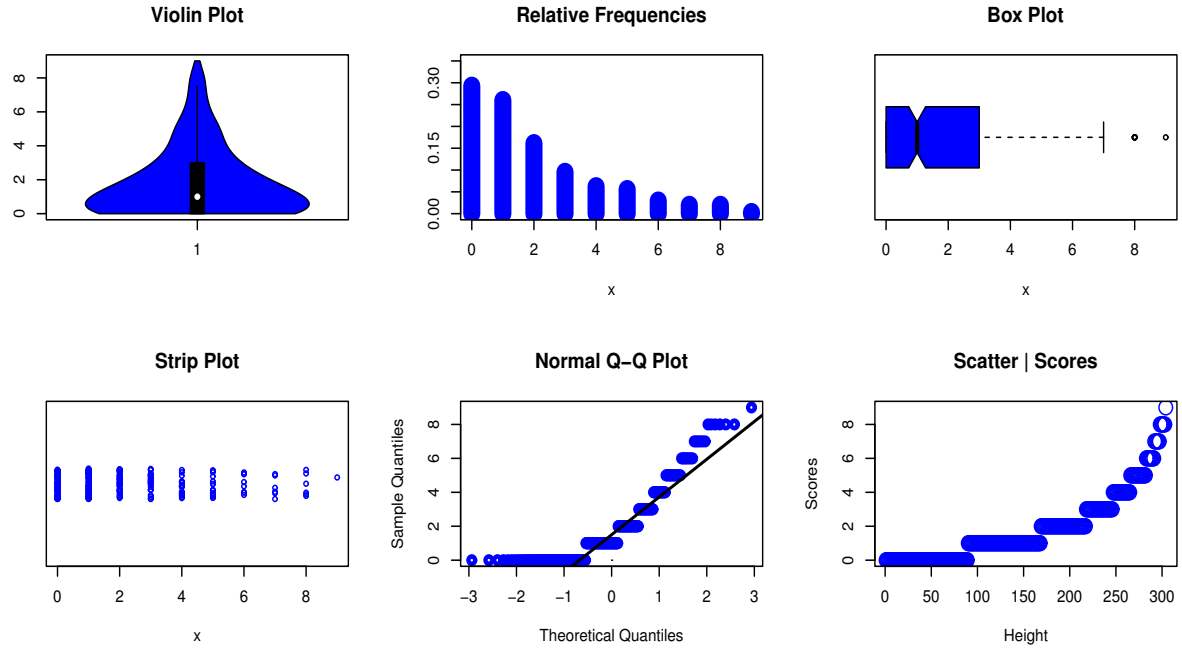

Figure12. Non-parametric plots of the COVID-19SK data.

The MxLk estimators with their corresponding SDES for the parameter(s) as well as GOFT are listed in Tables 14, 15 and 16, respectively.

Table 14. The MxLk and SDEs for the COVID-19SK data.

Parameter ⟶	ϑ		ε		λ		b	
Model ↓	MxLk	SDES	MxLk	SDES	MxLk	SDES	MxLk	SDES
DOLoW	0.0411	0.0485	0.3492	0.0828	0.9368	0.0568	2.3354	0.4929
DOWIW	0.9971	0.0002	1.0640	0.0690	0.9914	0.0033	1.0831	0.1025
DIW	–	–	–	–	0.2715	0.0252	1.4110	0.0831
DW	–	–	–	–	0.7070	0.0014	1.1543	0.0035
DGzW	0.1933	0.0124	0.3456	0.0251	0.1668	0.0471	1.0332	0.0964
EDW	–	–	0.7384	0.3401	0.8049	0.1393	1.3552	0.3492
DLogL	1.7161	0.0954	1.8783	0.1073	–	–	–	–
DBH	0.9043	0.0201	–	–	–	–	–	–
DPa	0.3770	0.0212	–	–	–	–	–	–
DIR	–	–	–	–	0.2294	0.0232	–	–
DB	0.5914	0.0313	2.4661	0.2483	–	–	–	–
Poi	1.9013	0.0795	–	–	–	–	–	–

Table 15. The GOFT of the COVID-19SK data "part I".

	Observed	Expected Frequencies					
X	Frequencies	DOLoW	DOWIW	DIW	DW	DGzW	EDW
0	89	88.1691	89.0402	82.3512	89.0833	91.4694	90.9439
1	79	82.7013	74.5281	103.7023	74.4215	70.2269	70.9587
2	49	44.4350	51.8622	44.3904	51.8637	50.4689	51.1197
3	29	29.1421	34.0593	22.3170	34.0858	34.7692	34.8180
4	19	21.4314	21.5954	12.9261	21.6184	22.9942	22.6380
5	17	15.7970	13.3512	8.2479	13.3642	14.5784	14.1296
6	9	10.6961	8.0920	5.6358	8.0974	8.8397	8.4988
7	6	6.3270	4.8250	4.0521	4.8256	5.1109	4.9410
8	6	3.2053	2.3872	3.0310	2.8355	2.8080	2.7836
9	1	2.0991	4.2594	17.3462	3.8047	2.7344	3.1688
Total	304	304	304	304	304	304	304
$-L$		561.961	564.625	586.855	564.625	564.009	564.429
AIC		1131.920	1137.250	1177.711	1133.250	1136.023	1134.859
CAIC		1132.061	1137.384	1177.751	1133.291	1136.151	1134.939
χ^2		1.839	2.792	41.868	2.811	3.799	3.618
DF		4	3	6	5	4	4
P-V		0.765	0.425	< 0.001	0.729	0.434	0.460

Table 16. The GOFT of the COVID-19SK data "part II".

	Observed	Expected Frequencies					
X	Frequencies	DLogL	DBH	DPa	DIR	DB	Poi
0	89	80.9312	166.601	149.3964	69.8901	92.8871	45.4081
1	79	92.7753	54.5982	50.5003	140.6133	97.7882	86.3362
2	49	51.4314	26.6673	25.4782	47.6852	42.6760	82.0743
3	29	27.3361	15.5404	15.3791	19.1251	21.1743	52.0144
4	19	15.5590	10.0170	10.3120	9.3330	12.1724	24.7250
5	17	9.5181	6.8854	7.3873	5.1984	7.7544	9.4024
6	9	6.1932	4.9533	5.5332	3.1623	5.3083	2.9793
7	6	4.2473	3.6822	4.4084	2.0981	3.8312	0.8092
8	6	3.0614	2.8081	3.4661	1.4290	2.8791	0.1921
9	1	12.9470	12.2471	32.1390	5.4655	17.5291	0.0590
Total	304	304	304	304	304	304	304
$-L$		577.011	620.466	633.531	606.870	587.652	621.098
AIC		1158.023	1242.932	1269.061	1215.740	1179.304	1244.195
CAIC		1158.063	1242.945	1269.075	1215.754	1179.344	1244.208
χ^2		25.019	109.333	128.631	92.204	44.784	115.896
DF		6	6	6	6	6	4
P-V		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

As we can see, the DOLoW, DOWIW, DW, DGzW and EDW distributions work quite well to discuss the COVID-19SK data, but the DOLoW model is the best among all tested distributions. Figure 13 supports the results reported in Tables 15 and 16.

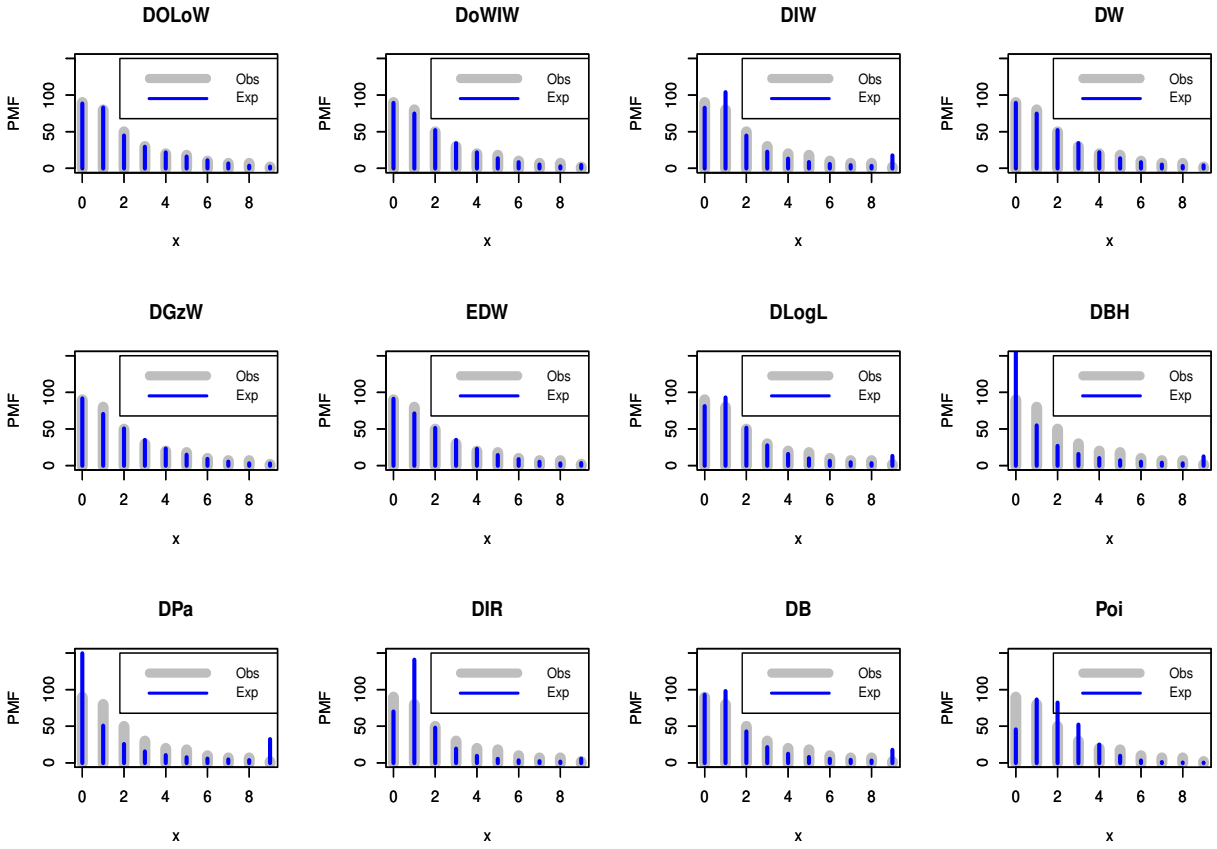

Figure 13. The observed and expected PMFs for the COVID-19SK data.

Figure 14 shows the L profile of the parameters of DOLoW model according to the COVID-19SK data. As we can see, the estimators are unique.

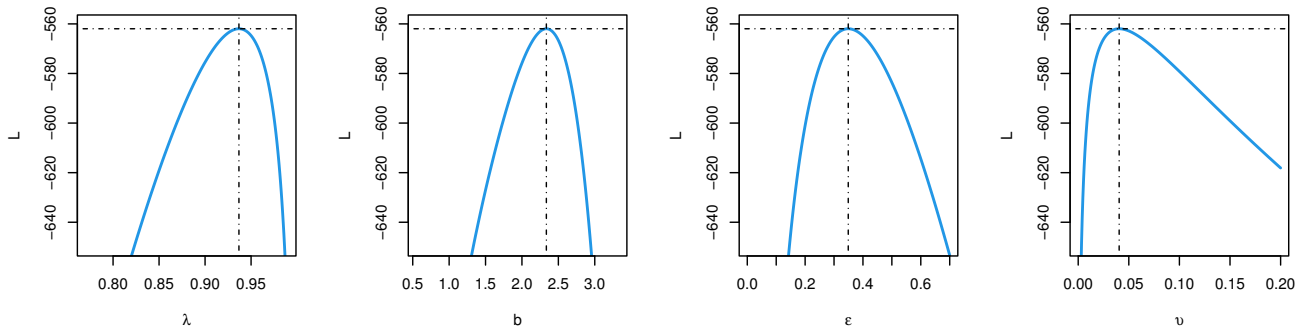

Figure 14. The L profile of the DOLoW parameters for the COVID-19SK data.

Table 17 lists the difference between experimental and observational descriptive statistics, and the two measures are found to be approximately equal.

Table 17. Descriptive statistics of the COVID-19SK data.

Model	Mean	Variance	Index of dispersion	Skewness	Kurtosis
Observed	1.85671	3.99782	2.15311	1.36570	4.23670
Empirical	1.90022	4.11995	2.16809	1.29282	4.31006

As we can note, the COVID-19SK data is skewed to the right with leptokurtic shape. Further, the data is suffering from over dispersion phenomena.

7.3. Data set III: COVID-19 in Angola (COVID-19A)

This data represents the daily new deaths of COVID-19A from 10 October 2020 to 29 November 2020 (see, https://www.worldometers.info/coronavirus/country/Angola/). In this segment, the MxLk estimators of the DOLoW parameters are derived for the COVID-19A data. Figure 15 shows some non-parametric plots to determine the behavior of this data. There are no extreme observations, and the shape of the data is positively skewed.

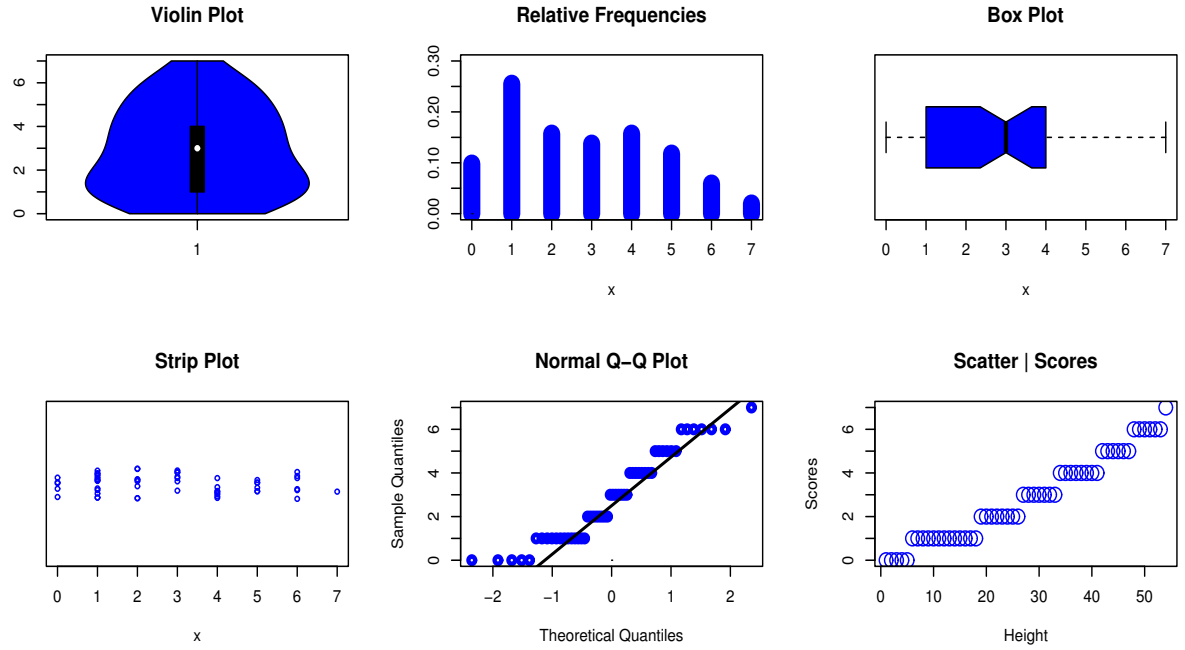

Figure 15. Non-parametric plots of the COVID-19A data.

The MxLk estimators with their corresponding SDES for the parameter(s) as well as GOFT are introduced in Tables 18, 19 and 20, respectively.

Table 18. The MxLk and SDES for the COVID-19A data.

Parameter ⟶	λ		ϑ		ε		b	
Model ↓	MxLk	SDES	MxLk	SDES	MxLk	SDES	MxLk	SDES
DOLoW	0.9890	0.0029	0.0111	0.0034	0.1503	0.0813	3.9260	0.1575
DGzW	0.6949	0.1236	0.3238	0.0414	0.2880	0.1741	1.3856	0.0417
DOWIW	2.5566	0.5300	0.1419	0.0194	12.16821	–	0.8795	0.0363
DW	0.9026	0.0046	–	–	–	–	1.7863	0.0300
EDW	0.9992	0.0006	–	–	0.2882	0.0319	4.1814	0.3591
DIW	3.1001×10^{-6}	1.713×10^{-6}	–	–	–	–	2.3032	0.0383
DBH	0.9361	0.0055	–	–	–	–	–	–
DB	0.8139	0.0095	–	–	–	–	4.8602	0.2668
Poi	2.6863	0.0321	–	–	–	–	–	–
DLogL	3.8403	0.0380	–	–	–	–	3.5747	0.0593
DPa	0.5289	0.0066	–	–	–	–	–	–
DIR	0.0001	0.00002	–	–	–	–	–	–

Table 19. The GOFT of the COVID-19A data "part I".

	Observed	Expected Frequencies					
X	Frequencies	DOLoW	DGzW	DOWIW	DW	EDW	DIW
0	5	5.0577	5.9272	4.4067	4.9679	6.5359	3.9033
1	13	12.7769	9.0886	11.4424	10.2179	8.5034	14.6702
2	8	8.2285	9.9893	11.055	11.2194	9.2502	11.7252
3	7	7.0635	9.4390	9.2179	9.5256	9.1879	7.0524
4	8	7.5566	7.6481	6.5447	6.7745	8.0207	4.2099
5	6	6.3829	5.1072	3.8596	4.1722	5.6742	2.6213
6	3	3.1138	2.6279	4.4701	2.2652	2.8576	1.7108
7	1	0.8201	1.1727	0.0038	1.8574	0.9702	5.1068
Total	51	51	51	51	51	51	51
$-L$		98.341	99.908	100.622	100.754	99.932	106.386
AIC		204.681	207.815	209.244	205.209	205.863	216.771
CAIC		205.551	208.685	210.114	205.759	206.384	217.021
χ^2		0.070	3.005	2.327	1.226	3.455	12.879
DF		2	1	0	2	2	2
P-V		0.966	0.083	< 0.001	0.542	0.178	0.002

Table 20. The GOFT of the COVID-19A data "part II".

	Observed	Expected Frequencies					
X	Frequencies	DBH	DB	Poi	DLogL	DPa	DIR
0	5	27.1287	6.7802	3.4758	3.7947	21.0495	4.9258
1	13	8.9736	18.9012	9.3360	12.1599	8.0133	13.1245
2	8	4.4381	8.3399	12.5383	12.4327	4.3483	10.3829
3	7	2.6264	4.2361	11.2259	8.3599	2.7699	6.6551
4	8	1.7225	2.5474	7.5382	4.9896	1.9361	4.2464
5	6	1.2075	1.6992	4.0495	2.9743	1.4382	2.8049
6	3	0.8869	1.2139	1.8128	1.8366	1.1154	1.9267
7	1	4.0163	7.2821	1.0234	4.4523	10.3294	6.9336
Total	51	51	51	51	51	51	51
$-L$		132.086	114.520	102.123	104.961	131.664	107.053
AIC		266.172	233.040	206.245	213.923	265.328	216.106
CAIC		266.254	233.290	206.327	214.173	265.409	216.188
χ^2		41.966	28.842	6.771	7.47185	24.7506	10.078
DF		2	2	3	2	2	3
P-V		< 0.001	< 0.001	0.079	0.024	< 0.001	0.018

As we can note, the DOLoW distribution is the best model among all competitive distributions. Figure 16 supports the results reported in Tables 19 and 20.

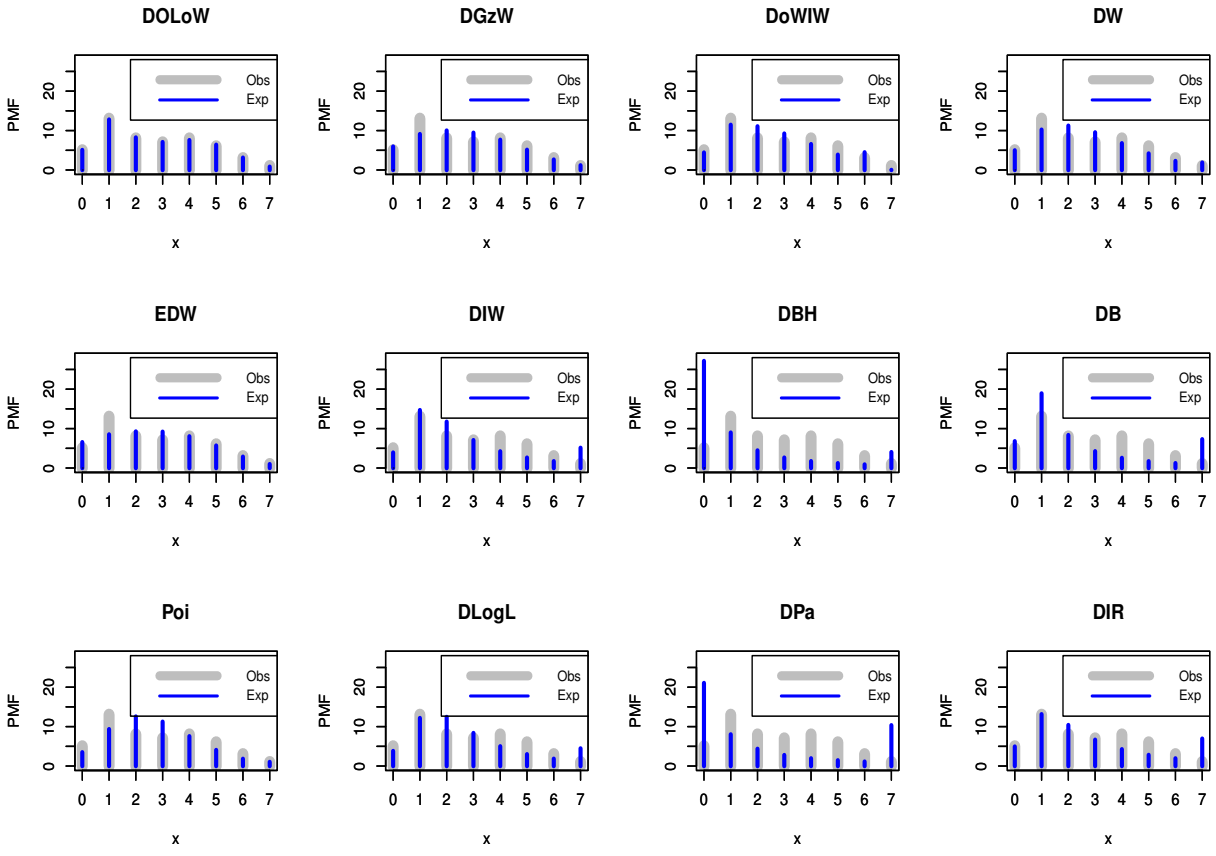

Figure 16. The observed and expected PMFs for the COVID-19A data.

The L profile of the DOLoW parameters can be displayed in Figure 17, and it was found that the derived estimators are unique.

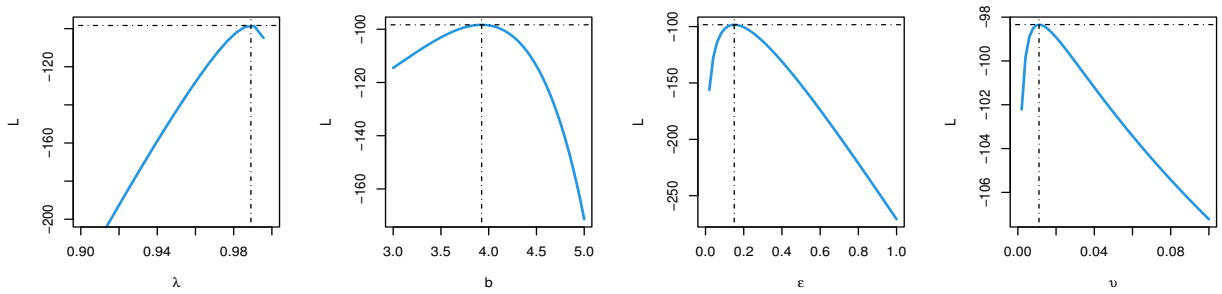

Figure 17. The L profile of the DOLoW parameters for the COVID-19A data.

Table 21 lists the difference between observational and experimental descriptive statistics, and the two measures are noted to be approximately equal.

Table 21. Descriptive statistics of the COVID-19A data.

Model	Mean	Variance	Index of dispersion	Skewness	Kurtosis
Observed	2.686275	3.430988	1.277229	0.3710481	2.117228
Empirical	2.687623	3.427940	1.275454	0.3659062	2.117569

As we can note from Table 21, the COVID-19A data is suffering from over dispersion phenomena. Furthermore, data is positively skewed with platykurtic shape. Figures 18, 19 and 20 show the contour plots based on datasets I, II, and III, respectively.

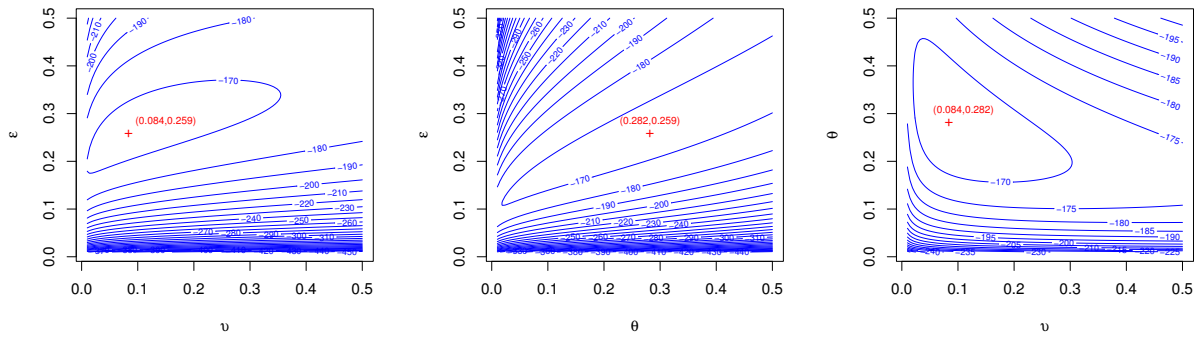

Figure 18. The contour plots for the COK data.

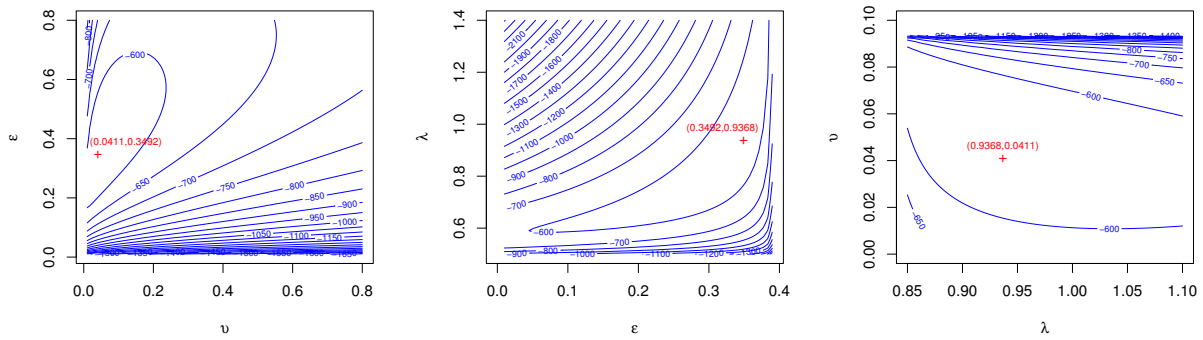

Figure 19. The contour plots for the COVID-19SK data.

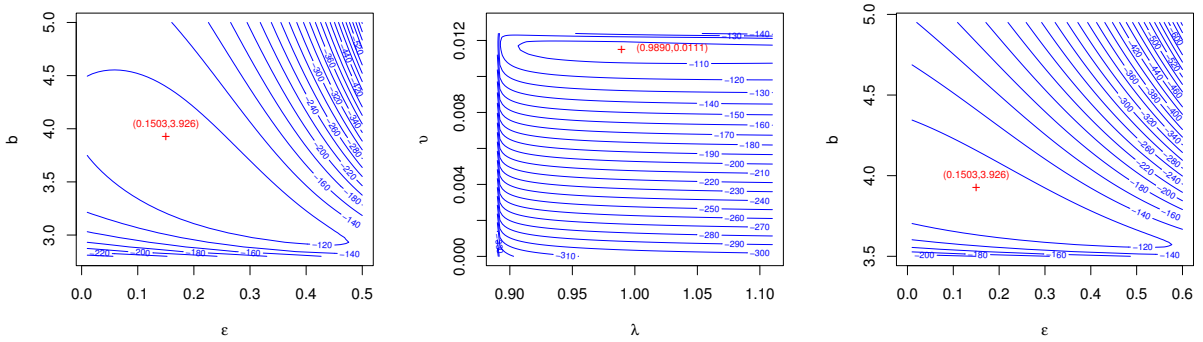

Figure 20. The contour plots for the COVID-19A data.

To discuss the behavior of risk modifier "hazard rates" shapes, a total time on test (TTT) diagram for the first, second, and third datasets is presented as a theoretical approach (see, Figure 21).

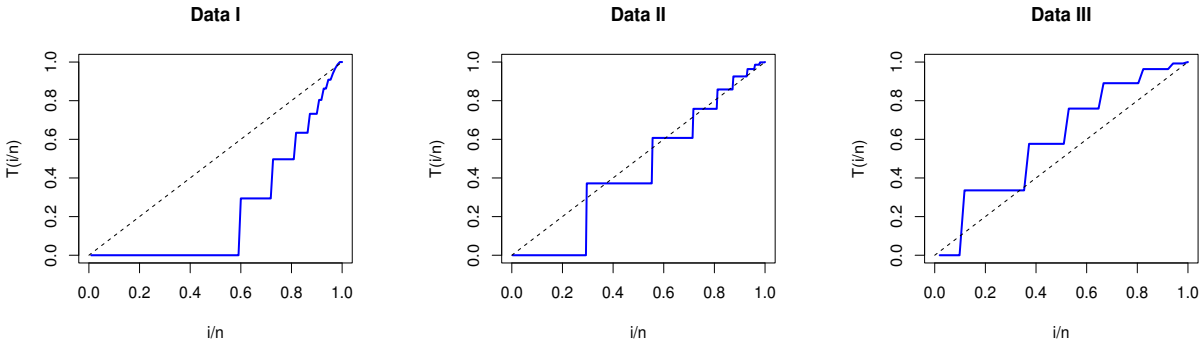

Figure 21. The TTT plots for datsets.

As we can see from theoretical and empirical "Section 4" plots, the DOLoGeo and DOLoW distributions will work well for discussing the failure rates shapes of the data under study.

8. Conclusions

In this article, a discrete analogue of odd Lomax generator of distributions has been derived and discussed in detail. After introducing the new class, two special discrete models have been investigated. Some distributional properties including, probability mass function, hazard rate function, quantile, crude moments, index of dispersion, entropies, order statistics, and L-moment statistics have been derived. It was found that the new probabilistic discrete class can be applied to discuss symmetric and asymmetric data under various kinds of kurtoses including platykurtic and leptokurtic shapes. It can be used to analyze over- (under-) dispersed count data with zero-inflated, extreme or outliers' observations. Moreover, it can be used to create many discrete models including different hazard rate forms. The class parameters have been estimated via the maximum likelihood approach. The performance of the estimation technique has been discussed utilizing MCMC simulation technique. Finally, to illustrate the flexibility of the new generator, three real data sets have been analyzed and discussed in detail. The results proved that the new class provides high efficiency in data modeling among all competing models. As future work, the authors will propose the bivariate extension of this family with its applications.

Acknowledgement. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444)

References

[1] Ghitany, M. E., Al-Awadhi, F. A., & Alkhalfan, L. (2007). Marshall–Olkin extended Lomax distribution and its application to censored data. Communications in Statistics: Theory and Methods, 36(10), 1855-1866.
[2] Ashour, S. K., & Eltehiwy, M. A. (2013). Transmuted Lomax distribution. American Journal of Applied Mathematics and Statistics, 1(6), 121-127.
[3] Kilany, N. M. (2016). Weighted Lomax distribution. SpringerPlus, 5(1), 1-18.
[4] Rady, E. H. A., Hassanein, W. A., & Elhaddad, T. A. (2016). The power Lomax distribution with an application to bladder cancer data. SpringerPlus, 5(1), 1-22.
[5] Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., Adejumo, A. O., & Odetunmibi, O. A. (2017). A new generalization of the Lomax distribution with increasing, decreasing, and constant failure rate. Modelling and Simulation in Engineering, 2017.
[6] Elbiely, M. M., & Yousof, H. M. (2018). A new extension of the Lomax distribution and its applications. Journal of Statistics and Applications, 2(1), 18-34.
[7] Hassan, A. S., & Abd-Allah, M. (2019). On the inverse power Lomax distribution. Annals of Data Science, 6(2), 259-278.
[8] Cordeiro, G. M., Afify, A. Z., Ortega, E. M., Suzuki, A. K., & Mead, M. E. (2019). The odd Lomax generator of distributions: Properties, estimation and applications. Journal of Computational and Applied Mathematics, 347, 222-237.
[9] Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. Metron, 71(1), 63-79.
[10] Krishna, H., & Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. Statistical methodology, 6(2), 177-188.
[11] Hussain, T., & Ahmad, M. (2014). Discrete inverse Rayleigh distribution. Pakistan Journal of Statistics, 30(2).
[12] Nekoukhou, V., & Bidram, H. (2015). The exponentiated discrete Weibull distribution. Sort, 39, 127-146.
[13] Eliwa, M. S., Altun, E., El-Dawoody, M., & El-Morshedy, M. (2020). A new three-parameter discrete distribution with associated INAR (1) process and applications. IEEE access, 8, 91150-91162.
[14] Eliwa, M. S., Alhussain, Z. A., & El-Morshedy, M. (2020). Discrete Gompertz-G family of distributions for over-and under-dispersed data with properties, estimation, and applications. Mathematics, 8(3), 358.
[15] El-Morshedy, M., Eliwa, M. S., & Tyagi, A. (2021). A discrete analogue of odd Weibull-G family of distributions: properties, classical and Bayesian estimation with applications to count data. Journal of Applied Statistics, https://doi.org/10.1080/02664763.2021.1928018.
[16] Eldeeb, A. S., Ahsan-ul-Haq, M., Eliwa, M. S., & Cell, Q. E. (2022). A discrete Ramos-Louzada distribution for asymmetric and over-dispersed data with leptokurtic-shaped: Properties and various estimation techniques with inference. AIMS Mathematics, 7(2), 1726-1741.
[17] Almetwally, E. M., Abdo, D. A., Hafez, E. H., Jawa, T. M., Sayed-Ahmed, N., & Almongy, H. M. (2022). The new discrete distribution with application to COVID-19 data. Results in Physics, 32, 104987.
[18] Chan, S. K., Riley, P. R., Price, K. L., McElduff, F., Winyard, P. J., Welham, S. J., ... & Long, D. A. (2010). Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as Indian hedgehog. American journal of physiology-renal physiology, 298(2), F346-F356.
[19] A Rényi, On measures of entropy and information, Math. Statist. Probab. 1 (1961), pp. 547–561.
[20] Eliwa, M. S., & El-Morshedy, M. (2021). A one-parameter discrete distribution for over-dispersed data: Statistical and reliability properties with applications. Journal of Applied Statistics, 49(10), 2467-2487.
[21] Para, B. A., & Jan, T. R. (2016). On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science. Biometrics and Biostatistics International Journal, 4(2), 1-15.

[22] Para, B. A., & Jan, T. R. (2016). Discrete version of log-logistic distribution and its applications in genetics. International Journal of Modern Mathematical Sciences, 14(4), 407-422.
[23] Poisson, S. D. (1837). Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilitiés. Paris, France: Bachelier, 1, 1837.
[24] El-Morshedy, M., Eliwa, M. S., & Altun, E. (2020). Discrete Burr-Hatke distribution with properties, estimation methods and regression model. IEEE access, 8, 74359-74370.
[25] Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. Test, 19(2), 399-415.
[26] Jazi, M. A., Lai, C. D., & Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. Statistical Methodology, 7(2), 121-132.
[27] Gómez-Déniz, E., & Calderín-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. Journal of statistical computation and simulation, 81(11), 1405-1416.
[28] Hussain, T., Aslam, M., & Ahmad, M. (2016). A two parameter discrete Lindley distribution. Revista Colombiana de EstadÝstica, 39(1), 45-61.
[29] Roy, D. (2004). Discrete Rayleigh distribution. IEEE Transactions on Reliability, 53(2), 255-260.
[30] Nakagawa, T., & Osaki, S. (1975). The discrete Weibull distribution. IEEE transactions on reliability, 24(5), 300-301.