axioms

MDPI

Article

A Discrete Linear-Exponential Model: Synthesis and Analysis with Inference to Model Extreme Count Data

Mahmoud El-Morshedy [1,2]

[1] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[2] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

Abstract: In this article, a novel probability discrete model is introduced for modeling overdispersed count data. Some relevant statistical and reliability properties including the probability mass function, hazard rate and its reversed function, moments, index of dispersion, mean active life, mean inactive life, and order statistics, are derived in-detail. These statistical properties are expressed in closed forms. The new model can be used to discuss right-skewed data with heavy tails. Moreover, its hazard rate function can be utilized to model the phenomena with a monotonically increasing failure rate shape. Different estimation approaches are listed to get the best estimator for modeling and reading the count data. A comprehensive comparison among techniques is performed in the case of simulated data. Finally, four real data sets are analyzed to prove the ability and notability of the new discrete model.

Keywords: linear-exponential model; survival discretization; Markov chain Monte Carlo; extreme observations

check for updates

Citation: El-Morshedy, M. A Discrete Linear-Exponential Model: Synthesis and Analysis with Inference to Model Extreme Count Data. *Axioms* **2022**, *11*, 531. https://doi.org/10.3390/axioms11100531

Academic Editor: Angel Ricardo Plastino

Received: 2 September 2022
Accepted: 30 September 2022
Published: 4 October 2022

1. Introduction

Data modeling in recent years has been very complicated due to the huge number of data sets which have been generated from different fields over time, especially, in engineering, medical, ecology, and renewable energy. The principal problem is when the data are suffering from overdispersion with different kinds of kurtosis including leptokurtic- or platykurtic-shaped. Therefore, it is important to model and analyze such data by utilizing a flexible probability distribution. Thus, several continuous probability models have been introduced and discussed in the statistical literature for this purpose. In several cases, the data need to be recorded on a discrete scale rather than on a continuous analogue. Due to the previous reason, the discretization of existing continuous distributions has received a wide attention because of the count data generated from various areas becoming more complex day-by-day. So, for modeling these count data, we need discrete probability models that are best suited for analytical studies of this multidimensional and complex phenomena. Many discrete distributions have been proposed and studied in detail such as the discrete Rayleigh (see Roy [1]), discrete Pareto (see Krishna and Pundir [2]), discrete exponential generalized-G family (see Eliwa et al. [3]), discrete Lindley (see Gommez-Déniz and Calderin-Ojeda [4]), discrete generalized exponentiated type two (see Nekoukhou et al. [5]), discrete exponentiated Weibull (see Nekoukhou and Bidram [6]), discrete Lindley-II (DLi-II) (see Hussain et al. [7]), discrete Burr XII (see Para and Jan [8]), discrete Gompertz-G class (see Eliwa el al. [9]), discrete exponentiated Lindley (see El-morshedy et al. [10]), discrete generalized Burr-Hatke (see Yousof et al. [11]), discrete Rayleigh-G family (see Ibrahim et al. [12]), binomial new Poisson-weighted exponential model (see Al-Bossly and Eliwa [13]), among others. Although there are a number of discrete models in the literature, there is still a lot of space left to propose a new discretized model that is suitable under various conditions.

Recently, Sah [14] introduced a novel one-parameter linear-exponential (NLE) distribution. The NLE distribution was based on the product of a linear function $x + \varsigma^2$, and an exponential function $\exp[-\varsigma x]$ with a single parameter $\varsigma > 0$. The cumulative distribution function (CDF) of the NLE model could be expressed as

$$G_\varsigma(x) = 1 - \frac{1 + \varsigma x + \varsigma^3}{1 + \varsigma^3} \exp[-\varsigma x]; \; x > 0, \tag{1}$$

where ς is a scale parameter. In this paper, a discrete analogue of the NLE model is presented under the abbreviation NDsLE. The nice feature of reporting the NDsLE model is that it stands with a single parameter which is to be listed so as to give a better alternative for some discrete distributions and to create another platform for researchers working on probability distribution theory. Other interesting features for the NDsLE model can be listed as follows:

- Its distributional statistics can be expressed in explicit terms.
- It can be used to model positively skewed count data.
- It can be utilized to discuss overdispersed count data.
- It can be applied to study count data which have a monotonically increasing hazard rate function (HRF).
- It can be used as a statistical tool to model extreme count data.

The rest of the paper is organized as follows, In Section 2, we introduce the NDsLE distribution. Various distributional statistics are derived in Section 3. In Section 4, the NDsLE parameter is estimated by using various techniques including maximum likelihood, proportion, moments, least squares, weighted least squares, and Cramér–von Mises criterion to get the best estimator for modeling data. A simulation study is presented in Section 5. Four real data sets are analyzed to show the flexibility of the NDsLE distribution in Section 6. Finally, Section 7 provides some conclusions.

2. The NDsLE Distribution: Mathematical Synthesis

Starting with (1) and utilizing the discretization concepts, the CDF of the NDsLE distribution can be expressed as

$$F_\beta(x) = 1 - \left[1 - \frac{(x+1)\ln(\beta)}{1 - [\ln(\beta)]^3}\right] \beta^{x+1}; \; x = 0, 1, 2, 3, \ldots, \tag{2}$$

where $\beta = \exp[-\varsigma]$ and $0 < \beta < 1$. The survival function corresponding to (2) can be proposed as

$$S_\beta(x) = \left[1 - \frac{(x+1)\ln(\beta)}{1 - [\ln(\beta)]^3}\right] \beta^{x+1}; \; x = 0, 1, 2, 3, \ldots. \tag{3}$$

The probability mass function (PMF) can be formulated as

$$f_\beta(x) = \left[1 - \frac{x\ln(\beta)}{1 - [\ln(\beta)]^3}\right] \beta^{x} - \left[1 - \frac{(x+1)\ln(\beta)}{1 - [\ln(\beta)]^3}\right] \beta^{x+1}; \; x = 0, 1, 2, 3, \ldots, \tag{4}$$

where the PMF of any discrete random variable (RV) can be derived as

$$f(x; \boldsymbol{\theta}) = S(x; \boldsymbol{\theta}) - S(x+1; \boldsymbol{\theta}); \;\; x = 0, 1, 2, 3, \ldots \tag{5}$$

where $\boldsymbol{\theta}$ is a vector of parameters. The PMF in (4) is log-concave, where $\frac{f_\beta(x+1)}{f_\beta(x)}$ is a decreasing function in x for all values of the NDsLE parameter. Depending on (4) and (3), the HRF can be derived as

$$h_\beta(x) = 1 - \left[1 - \frac{(x+1)\ln(\beta)}{1 - [\ln(\beta)]^3}\right]\left[1 - \frac{x\ln(\beta)}{1 - [\ln(\beta)]^3}\right]^{-1}\beta;\ x = 0, 1, 2, 3, \ldots, \tag{6}$$

where $h_\beta(x) = \frac{f_\beta(x)}{S_\beta(x-1)}$, whereas the reversed HRF (RHRF) can be presented as

$$r_\beta(x) = \frac{\left[1 - \frac{x\ln(\beta)}{1-[\ln(\beta)]^3}\right]\beta^x - \left[1 - \frac{(x+1)\ln(\beta)}{1-[\ln(\beta)]^3}\right]\beta^{x+1}}{1 - \left[1 - \frac{(x+1)\ln(\beta)}{1-[\ln(\beta)]^3}\right]\beta^{x+1}};\ x = 0, 1, 2, 3, \ldots, \tag{7}$$

where $r_\beta(x) = \frac{f_\beta(x)}{F_\beta(x)}$. In Figure 1, we give some $f_\beta(x)$, $h_\beta(x)$ and $r_\beta(x)$ plots of the NDsLE model under some selected parameter values.

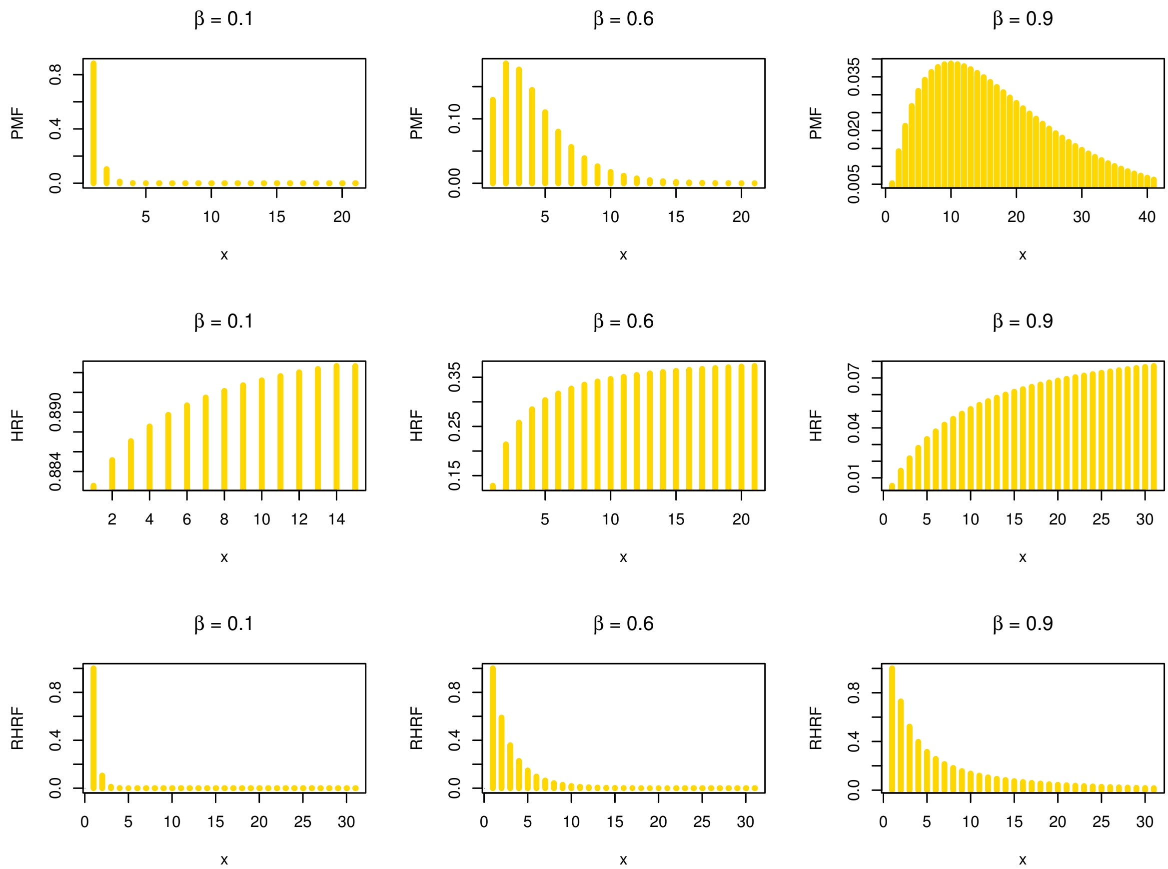

Figure 1. The PMF, HRF and RHRF plots of the NDsLE model.

Based on Figure 1, it is noted that the $f_\beta(x)$ of the NDsLE model is unimodal, and it can be used as a probability tool to model asymmetric data. Moreover, the NDsLE model is a proper approach for modeling some phenomena which have increasing HRF or decreasing RHRF. Suppose Y and Z are two independent NDsLE RVs with parameters β_1 and β_2, respectively. Then, the HRF of $T = \min(Y, Z)$ can be formulated as

$$
\begin{aligned}
h_T(x;\beta_1,\beta_2) &= \frac{\Pr(\min(Y,Z)=x)}{\Pr(\min(Y,Z)\geq x)} \\
&= \frac{\Pr(\min(Y,Z)\geq x)-\Pr(\min(Y,Z)\geq x+1)}{\Pr(\min(Y,Z)\geq x)} \\
&= \frac{\Pr(Y\geq x)\Pr(Z\geq x)-\Pr(Y\geq x+1)\Pr(Z\geq x+1)}{\Pr(Y\geq x)\Pr(Z\geq x)} \\
&= \frac{\Pr(Y\geq x)\Pr(Z=x)+\Pr(Y=x)\Pr(Z\geq x)-\Pr(Y=x)\Pr(Z=x)}{\Pr(Y\geq x)\Pr(Z\geq x)},
\end{aligned}
$$

then

$$
\begin{aligned}
h_T(x;\beta_1,\beta_2) &= \left\{1-\left[1-\frac{(x+1)\ln(\beta_1)}{1-[\ln(\beta_1)]^3}\right]\left[1-\frac{x\ln(\beta_1)}{1-[\ln(\beta_1)]^3}\right]^{-1}\beta_1\right\} \\
&\quad+\left\{1-\left[1-\frac{(x+1)\ln(\beta_2)}{1-[\ln(\beta_2)]^3}\right]\left[1-\frac{x\ln(\beta_2)}{1-[\ln(\beta_2)]^3}\right]^{-1}\beta_2\right\} \\
&\quad-\left\{1-\left[1-\frac{(x+1)\ln(\beta_1)}{1-[\ln(\beta_1)]^3}\right]\left[1-\frac{x\ln(\beta_1)}{1-[\ln(\beta_1)]^3}\right]^{-1}\beta_1\right\} \\
&\quad\times\left\{1-\left[1-\frac{(x+1)\ln(\beta_2)}{1-[\ln(\beta_2)]^3}\right]\left[1-\frac{x\ln(\beta_2)}{1-[\ln(\beta_2)]^3}\right]^{-1}\beta_2\right\}.
\end{aligned}
$$

The extra term $h_{\beta_1}(x)h_{\beta_2}(x)$ arises because in the discrete form, $\Pr(Y=x,Z=x)\neq 0$. Since the HRF of the two RVs Y and Z is increasing, then the HRF of $T=\min(Y,Z)$ is also increasing. Similarly, the HRF of $H=\max(Y,Z)$ can be expressed as

$$
h_H(x;\beta_1,\beta_2)=1-\frac{1-\left\{1-\left[1-\frac{(x+1)\ln(\beta_1)}{1-[\ln(\beta_1)]^3}\right]\beta_1^{x+1}\right\}\left\{1-\left[1-\frac{(x+1)\ln(\beta_2)}{1-[\ln(\beta_2)]^3}\right]\beta_2^{x+1}\right\}}{1-\left\{1-\left[1-\frac{x\ln(\beta_1)}{1-[\ln(\beta_1)]^3}\right]\beta_1^{x}\right\}\left\{1-\left[1-\frac{x\ln(\beta_2)}{1-[\ln(\beta_2)]^3}\right]\beta_2^{x}\right\}}.
$$

3. Main Statistical and Reliability Properties

3.1. Ordinary Moments and Descriptive Statistics

Let X be a non-negative RV, where $X\sim$NDsLE(β), then the rth moments can be expressed as

$$
\begin{aligned}
\mathbb{E}(X^r) &= \sum_{x=0}^{+\infty} x^r\left\{\left[1-\frac{x\ln(\beta)}{1-[\ln(\beta)]^3}\right]\beta^x-\left[1-\frac{(x+1)\ln(\beta)}{1-[\ln(\beta)]^3}\right]\beta^{x+1}\right\}. \qquad (8)\\
&= \frac{1}{(\log\beta)^3-1}\{-[\beta-1]\,\mathit{Hurwitzlerchphi}[\beta,-1-r,\theta]\log\beta \\
&\quad+\mathit{Hurwitzlerchphi}[\beta,-r,\theta][-1+\beta-\beta\log\beta-(\beta-1)(\log\beta)^3]\},
\end{aligned}
$$

where *Hurwitzlerchphi* represents the Hurwitz–Lerch transcendental function which can be proposed in the form $\Phi(z,s,a)=\sum_{k=0}^{\infty} z^k(k+a)^{-s}$. Setting $r=1,2,3,4$ in (8), the first four moments of the RV X can be derived in closed forms as

$$\mathbb{E}(X) = \frac{\left[(\beta-1)[\ln(\beta)]^3 - \ln(\beta) - \beta + 1\right]\beta}{\left(1-[\ln(\beta)]^3\right)(\beta-1)^2},$$
$$\mathbb{E}\left(X^2\right) = \frac{\left[(\beta^2-1)[\ln(\beta)]^3 - (3\beta+1)\ln(\beta) - \beta^2 + 1\right]\beta}{(\ln(\beta)-1)\left([\ln(\beta)]^2+\ln(\beta)+1\right)(\beta-1)^3},$$
$$\mathbb{E}\left(X^3\right) = \frac{\left[(\beta^3+3\beta^2-3\beta-1)\left([\ln(\beta)]^3-1\right) - \left(7\beta^2+10\beta+1\right)\ln(\beta)\right]\beta}{(\ln(\beta)-1)\left([\ln(\beta)]^2+\ln(\beta)+1\right)(\beta-1)^4},$$

and

$$\mathbb{E}\left(X^4\right) = \frac{\left[(\beta^4+10\beta^3-10\beta-1)\left([\ln(\beta)]^3-1\right) - \left(15\beta^3+55\beta^2+25\beta+1\right)\ln(\beta)\right]\beta}{(\ln(\beta)-1)\left([\ln(\beta)]^2+\ln(\beta)+1\right)(\beta-1)^5}.$$

According to the previous four moments, the variance "$\mathbb{V}(X)$", skewness "$\mathbb{S}(X)$", and kurtosis "$\mathbb{K}(X)$" can be expressed in closed forms where

$$\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2,$$

$$\mathbb{S}(X) = \frac{\mathbb{E}(X^3) - 3\mathbb{E}(X^2)\mathbb{E}(X) + 2[\mathbb{E}(X)]^3}{[\mathbb{V}(X)]^{3/2}}$$

and

$$\mathbb{K}(X) = \frac{\mathbb{E}(X^4) - 4\mathbb{E}(X^2)\mathbb{E}(X) + 6\mathbb{E}(X^2)[\mathbb{E}(X)]^2 - 3[\mathbb{E}(X)]^4}{[\mathbb{V}(X)]^2}.$$

The index of dispersion, say $\mathbb{I}(X)$, is defined as the $\mathbb{V}(X)$ to $\mathbb{E}(X)$ ratio; it indicates whether a certain model is suitable for under or overdispersed data sets. If $\mathbb{I}(X) < (> 1)$, the model is under- (over)dispersed. Further, the coefficient of variation, say $\mathbb{C}(X)$, is reported. The $\mathbb{I}(X)$ can be formulated in closed expression as

$$\mathbb{I}(X) = \frac{-(\beta-1)^2 - (\beta^2-1)\log\beta + \beta(\log\beta)^2 + 2(\beta-1)^2(\log\beta)^3 + (\beta^2-1)(\log\beta)^4 - (\beta-1)^2(\log\beta)^6}{(\beta-1)^2[(\log\beta)^3-1][1-\beta-\log\beta+(\beta-1)(\log\beta)^3]}. \quad (9)$$

Table 1 lists some numerical results of some descriptive statistics (DS) for the NDsLE model for different values of the parameter β.

Table 1. Some DS for the NDsLE model under various values of β.

β	$\mathbb{E}(X)$	$\mathbb{V}(X)$	$\mathbb{S}(X)$	$\mathbb{K}(X)$	$\mathbb{I}(X)$	$\mathbb{C}(X)$
0.05	0.058583	0.061944	4.478550	25.910824	1.057374	4.248412
0.1	0.132634	0.149298	3.236004	17.278912	1.125648	2.913229
0.2	0.347302	0.448986	2.352445	14.582675	1.292781	1.929338
0.3	0.697084	1.038812	1.916297	15.730504	1.490225	1.462121
0.4	1.242090	2.122653	1.638590	19.126849	1.708937	1.172969
0.5	2.039962	4.038364	1.483025	24.722690	1.979628	0.985101
0.6	3.190287	7.654077	1.421220	31.709619	2.399182	0.867195
0.7	4.987059	15.773291	1.408363	38.660981	3.162846	0.796373
0.8	8.413829	40.242572	1.410342	44.651470	4.782909	0.753962
0.9	18.47137	180.249242	1.413247	49.635682	9.758301	0.726838
0.95	38.48882	760.249742	1.413981	51.867544	19.752485	0.716380
0.99	198.498224	19,800.249960	1.414205	53.578961	99.750262	0.708889

From Table 1, the NDsLE model is appropriate for modeling overdispersed data sets where $\mathbb{I}(X) > 1$. Further, the NDsLE distribution is capable of modeling positively skewed data under various shapes of kurtosis. Figure 2 lists the results which are including in Table 1.

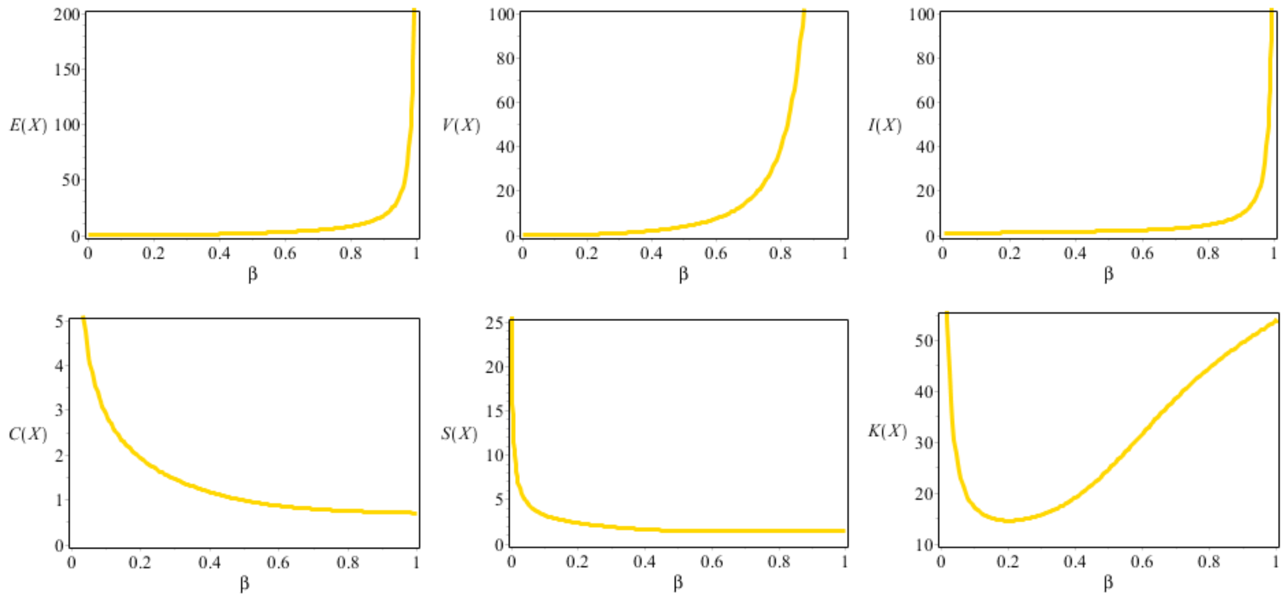

Figure 2. Some DS plots for the NDsLE model.

3.2. Mean Active and Inactive Life Measures

To study the aging behavior of a component, several reliability measures have been defined in the survival analysis (SA) literature. One of them is called the mean active life (MAL). The MAL is a helpful reliability tool to analyze the burn-in and maintenance policies. In the discrete setting, the MAL can be defined as $\Theta(i) = \mathbb{E}(X - i | X \geq i)$ for $i \in \mathbb{N}_0$ and $\mathbb{N}_0 = \{0, 1, 2, 3, ...\}$. Assume the RV $X \sim$NDsLE(β), then the MAL can be formulated as

$$\Theta_\beta(i) = \frac{-\beta^{i+1}}{\Xi(\beta, i)} \Big\{ (\beta - 1)[\ln(\beta)]^3 + (i\beta - i - 1)\ln(\beta) - \beta + 1 \Big\},$$

where

$$\Xi(\beta, i) = \left\{ 1 + \frac{i \ln(\beta)}{[\ln(\beta)]^3 - 1} \right\} \beta^i ([\ln(\beta)]^3 - 1)(\beta - 1)^2.$$

Another reliability concept of interest in SA is the mean inactive life (MIAL), which measures the time elapsed since the failure of X given that the component has failed some time before i. The MIAL, say $\Theta^*(i)$, is defined as $\Theta^*(i) = \mathbb{E}(i - X | X < i)$ for $i \in \mathbb{N}_0 - \{0\}$ and $\Theta^*(0) = 0$. Let $X \sim$NDsLE(β), then the MIAL can be expressed as

$$\Theta^*_\beta(i) = \Xi^*(\beta, i) \left[1 - \left\{ 1 + \frac{i \ln(\beta)}{[\ln(\beta)]^3 - 1} \right\} \beta^i \right]^{-1},$$

where

$$\Xi^*(\beta, i) = i - \frac{\beta^{i+1} - \beta}{\beta - 1} - \frac{\beta^{i+1}\{(i+1)\beta - i - 1 - \beta\}\ln(\beta) + \beta \ln(\beta)}{([\ln(\beta)]^3 - 1)(\beta - 1)^2}.$$

For $i \in \mathbb{N}_0$, we get $\Theta^*(i) \leq i$. Let X be a NDsLE RV, then the CDF can be proposed by the MIAL as

$$F_\beta(k) = F_\beta(0) \prod_{i=1}^{k} \left[\frac{\Theta^*(i)}{\Theta^*(i+1) - 1} \right]; \; k \in \mathbb{N}_0 - \{0\},$$

where $F_\beta(0) = \left(\prod_{i=1}^{d} \left[\frac{\Theta^*(i)}{\Theta^*(i+1)-1} \right] \right)^{-1}$ and $0 < d < \infty$. The mean of the NDsLE distribution can be expressed as

$$\text{Mean} = i - \Theta^*(i) F_\beta(i-1) + \Theta(i)\left[1 - F_\beta(i-1)\right]; \; i \in \mathbb{N}_0 - \{0\}.$$

The RHRF and the MIAL are related as

$$r_\beta(i) = \frac{1 - \Theta^*(i+1) + \Theta^*(i)}{\Theta^*(i)}; \; i \in \mathbb{N}_0 - \{0\}.$$

3.3. Order Statistics (ORSS) and L-Moment (L-M) Statistics

Let $X_1, X_2, ..., X_p$ be a random sample (RS) from the NDsLE distribution, and let $X_{1:p}, X_{2:p}, ..., X_{p:p}$ be their corresponding ORSS. Then, the CDF of the ith ORSS is

$$\begin{aligned} F_{\beta,i:p}(x) &= \sum_{k=i}^{p} \binom{p}{k} \left[F_{\beta,i}(x)\right]^k \left[1 - F_{\beta,i}(x)\right]^{p-k} \\ &= \sum_{k=i}^{p} \sum_{j=0}^{p-k} \Lambda_{(p)}^{(k)} F_{\beta,i}(x, k+j), \end{aligned}$$

where $\Lambda_{(p)}^{(k)} = (-1)^j \binom{p}{k} \binom{p-k}{j}$ and $F_{\beta,i}(x, k+j)$ represents the CDF of the exponentiated NDsLE model with power parameter $k+j$. Further, the corresponding PMF of the ith ORSS can be expressed as

$$f_{\beta,i:p}(x) = \sum_{k=i}^{p} \sum_{j=0}^{p-k} \Lambda_{(p)}^{(k)} f_{\beta,i}(x, k+j), \tag{10}$$

where $f_{\beta,i:p}(x) = F_{\beta,i:p}(x) - F_{\beta,i:p}(x-1)$. The c^{th} moments of $X_{i:p}$ is

$$\mathbb{E}(X_{i:p}^c) = \sum_{x=0}^{\infty} \sum_{k=i}^{p} \sum_{j=0}^{p-k} \Lambda_{(p)}^{(k)} x_{i:p}^c f_{\beta,i}(x, k+j). \tag{11}$$

Based on (11), the L-Ms can be derived from the following relation

$$\lambda_q = \frac{1}{q} \sum_{j=0}^{q-1} \Upsilon(j,q) \mathbb{E}\left(X_{q-j:q}\right), \tag{12}$$

where $\Upsilon(j,q) = (-1)^j \binom{q-1}{j}$. Utilizing (12), we can introduce some statistical measures such as the L-M of the mean $= \lambda_1$, the L-M coefficient of variation $= \frac{\lambda_2}{\lambda_1}$, the L-M coefficient of skewness $= \frac{\lambda_3}{\lambda_2}$ and the L-M coefficient of kurtosis $= \frac{\lambda_4}{\lambda_2}$.

4. Estimation Approaches

4.1. Maximum Likelihood Estimation (MLE)

In this section, we list the MLEs of the NDsLE parameter. Let $X_1, X_2, ..., X_p$ be an RS of size p from the NDsLE model. The log-likelihood function (L) can be formulated as

$$\begin{aligned} L_\beta(x_1, x_2, ..., x_p; \beta) &= \ln(\beta) \sum_{i=1}^{p} x_i + \sum_{i=1}^{p} \ln\Big[[\ln(\beta)]^3 (1-\beta) + \ln(\beta)(x_i - \beta x_i - \beta) + \beta - 1\Big] \\ &- \sum_{i=1}^{p} \ln[\ln(\beta - 1] - \sum_{i=1}^{p} \ln\Big[[\ln(\beta)]^2 + \ln(\beta) + 1\Big]. \end{aligned} \tag{13}$$

By differentiating (13) with respect to the parameter β, we get the normal nonlinear likelihood equation as follows

$$\begin{aligned} \frac{\partial L_\beta(x_1, x_2, ..., x_p; \beta)}{\partial \beta} &= \frac{1}{\beta} \sum_{i=1}^{p} x_i - \frac{1}{\beta} \sum_{i=1}^{p} \frac{\beta[\ln(\beta)]^3 + 3(\beta-1)[\ln(\beta)]^2 + \beta(x_i+1)\ln(\beta) - x_i + \beta x_i + 2\beta}{[\ln(\beta)]^3(1-\beta) + \ln(\beta)(x_i - \beta x_i - \beta) + \beta - 1} \\ &- \frac{1}{\beta} \sum_{i=1}^{p} \frac{1}{\ln(\beta) - 1} - \frac{1}{\beta} \sum_{i=1}^{p} \frac{2\ln(\beta) + 1}{[\ln(\beta)]^2 + \ln(\beta) + 1}. \end{aligned}$$

The resulted equation cannot be solved analytically. Thus, an iterative procedure such as the Newton–Raphson method is required to solve it numerically.

4.2. Proportion Estimation (ProE)

Assume $X_1, X_2, ..., X_p$ is an RS of size p from the NDsLE model. We define an indicator as follows

$$\mathbf{\Omega}(x_i) = \begin{cases} 1 & \text{if } x_i = 0 \\ 0 & \text{if otherwise.} \end{cases} \tag{14}$$

Let $\mathbf{O} = \sum_{i=1}^{p} \mathbf{\Omega}(x_i)$ stand for the number of zeros in the RS. Using the CDF of the NDsLE model as well as (14), we get $\Pr(X \le 0) = \frac{\mathbf{O}}{p}$. So, the parameter β is estimated by solving the following equation

$$1 - \widehat{\beta} + \frac{\widehat{\beta}\ln(\widehat{\beta})}{1 - \Big[\ln(\widehat{\beta})\Big]^3} - \frac{\mathbf{O}}{p} = 0,$$

where the $\widehat{\beta}$ of β is an unbiased and consistent estimator. In some cases, we could not get the zeros in the RS. Thus, we replace zeros by ones or by any observation inside the sample to get the estimator.

4.3. Moment's Estimation (MoE)

Let $X_1, X_2, ..., X_p$ be an RS of size p from the NDsLE distribution. Based on the approach of moments for estimating the parameter β, we can derive the estimator $\widehat{\beta}$ by solving the following equation

$$\frac{\Big[(\widehat{\beta} - 1)\Big[\ln(\widehat{\beta})\Big]^3 - \ln(\widehat{\beta}) - \widehat{\beta} + 1\Big]\widehat{\beta}}{\Big(1 - \Big[\ln(\widehat{\beta})\Big]^3\Big)\Big(\widehat{\beta} - 1\Big)^2} - \frac{1}{p} \sum_{i=1}^{p} x_i = 0, \tag{15}$$

with respect to β. A symbolic program should be utilized to solve (15) numerically according to data observations $x_i; i = 1, 2, 3, ..., p$.

4.4. Least Squares and Weighted Least Squares Estimations

Let $X_{(1)}, X_{(2)}, \cdots, X_{(p)}$ be the ORSS of the RS of size p from the NDsLE model. The least squares estimator, say LSE, of the NDsLE parameter can be derived by minimizing

$$B_{\beta}\left(x_{(1)}, x_{(2)}, ..., x_{(p)}\right) = \sum_{i=1}^{p}\left[F_{\beta}(x_{(i)}) - \frac{i}{p+1}\right]^2, \tag{16}$$

with respect to β, while the weighted LSE, say WLSE, of the NDsLE parameter can be proposed by minimizing

$$B^{*}_{\beta}\left(x_{(1)}, x_{(2)}, ..., x_{(p)}\right) = \sum_{i=1}^{p}\frac{(p+1)^2(p+2)}{i(p-i+1)}\left[F_{\beta}(x_{(i)}) - \frac{i}{p+1}\right]^2, \tag{17}$$

also with respect to β.

4.5. Cramér–Von Mises Minimum Distance (MD) Estimation

Cramér–von Mises estimator, say CVME, is a type of MD estimator and has less bias than the other MD estimators. Assume $X_{(1)}, X_{(2)}, \cdots, X_{(p)}$ is the ORSS of the RS of size p from the NDsLE distribution. Then, the CVME of the NDsLE parameter is listed by minimizing

$$M_{\beta}\left(x_{(1)}, x_{(2)}, ..., x_{(p)}\right) = \frac{1}{12p} + \sum_{i=1}^{p}\left[F_{\beta}(x_{(i)}) - \frac{2i-1}{2p}\right]^2, \tag{18}$$

with respect to β.

5. Simulations: Comparing Various Estimators (CVE)

A general form to generate an RV X from the NDsLE model is to first generate the value Z from the NLE distribution, and then to discretize this value to get X, where $X = [Z]$ is the largest integer less than or equal to Z. In this section, we assess the performance of the MLE, MoE, ProE, LSE, WLSE, and CVME estimators with respect to sample size p using R software. For CVE, Markov chain Monte Carlo simulations were performed based on various schemes. The assessment was based on a simulation study:

1 Generate $P = 10,000$ samples of different sizes "p_i; $i = 1, 2, 3, 4$" from the NDsLE model as follows
   - scheme I: $\beta = 0.1 \mid p_1 = 20,\ p_2 = 50,\ p_3 = 150,\ p_4 = 300,\ p_5 = 500,\ p_6 = 700.$
   - scheme II: $\beta = 0.6 \mid p_1 = 20,\ p_2 = 50,\ p_3 = 150,\ p_4 = 300,\ p_5 = 500,\ p_6 = 700.$
   - scheme III: $\beta = 0.9 \mid p_1 = 20,\ p_2 = 50,\ p_3 = 150,\ p_4 = 300,\ p_5 = 500,\ p_6 = 700.$

2 Compute the MLE, MoE, ProE, LSE, WLSE, and CVME for the 10,000 samples, say $\widehat{\beta}_j$ for $j = 1, 2, ..., 10,000$.

We calculated the bias "BS", mean squared errors (NDS), mean relative errors (NES) for $P = 10,000$ samples as

$$\left|BS(\widehat{\beta})\right| = \frac{1}{P}\sum_{j=1}^{P}\left|\widehat{\beta}_j - \beta_j\right|,\ NDS(\beta) = \frac{1}{P}\sum_{j=1}^{P}(\widehat{\beta}_j - \beta_j)^2,\ NES(\beta) = \frac{1}{P}\sum_{j=1}^{P}\frac{\left|\widehat{\beta}_j - \beta_j\right|}{\beta_j}.$$

The results of the simulations are listed in Tables 2–4 and provided via Figures 3–5. Based on the reported tables and figures, the BS approached to zero when the sample size p increased. Similarly, the NDS and NES of the parameter approached zero when p increased. These results revealed the unbiasedness, efficiency, consistency properties of the MLE, MoE, ProE, LSE, WLSE, and CVME estimators. Thus, we can conclude that all estimation techniques worked quite well under different sizes of samples.

Table 2. The performance of different estimators.

p	Criteria	MLE	MoE	ProE	LSE	WLSE	CVME
				$\beta = 0.1$			
20	**|BS|**	0.13725150	0.13901581	0.14074285	0.16108567	0.14705658	0.17605784
	NDS	0.01902380	0.01890457	0.01986388	0.02667562	0.02235636	0.03025188
	NES	0.27358819	0.27527452	0.28235828	0.32115804	0.29125118	0.35453818
50	**|BS|**	0.08171928	0.09025780	0.09645521	0.10278538	0.08905878	0.12285608
	NDS	0.00702353	0.00812688	0.00914528	0.01125328	0.00789680	0.01506367
	NES	0.16055877	0.18026840	0.19147852	0.20177478	0.17886973	0.24406986
150	**|BS|**	0.05062535	0.05194580	0.05474531	0.05834522	0.05112831	0.06955535
	NDS	0.00236687	0.00295055	0.00378943	0.00312597	0.00298364	0.00586742
	NES	0.09963255	0.10575505	0.11091282	0.11386744	0.10456843	0.14020875
300	**|BS|**	0.03456352	0.03577827	0.04237865	0.04452802	0.03756683	0.05192556
	NDS	0.00101222	0.00131788	0.00186584	0.00205870	0.00147848	0.00298087
	NES	0.06685143	0.07125485	0.07974215	0.08128451	0.07428947	0.10697707
500	**|BS|**	0.00563654	0.00572998	0.00745885	0.00936984	0.00652036	0.02363264
	NDS	0.00045415	0.00048236	0.00069854	0.00091994	0.00062365	0.00079102
	NES	0.00397664	0.00413697	0.00748885	0.00896674	0.00513036	0.00972369
700	**|BS|**	0.00023611	0.00029699	0.00042369	0.00082366	0.00040236	0.00093698
	NDS	0.00008835	0.00012631	0.00042835	0.00075526	0.00063697	0.00091236
	NES	0.00011447	0.00033695	0.00091023	0.00302587	0.00056369	0.00503667

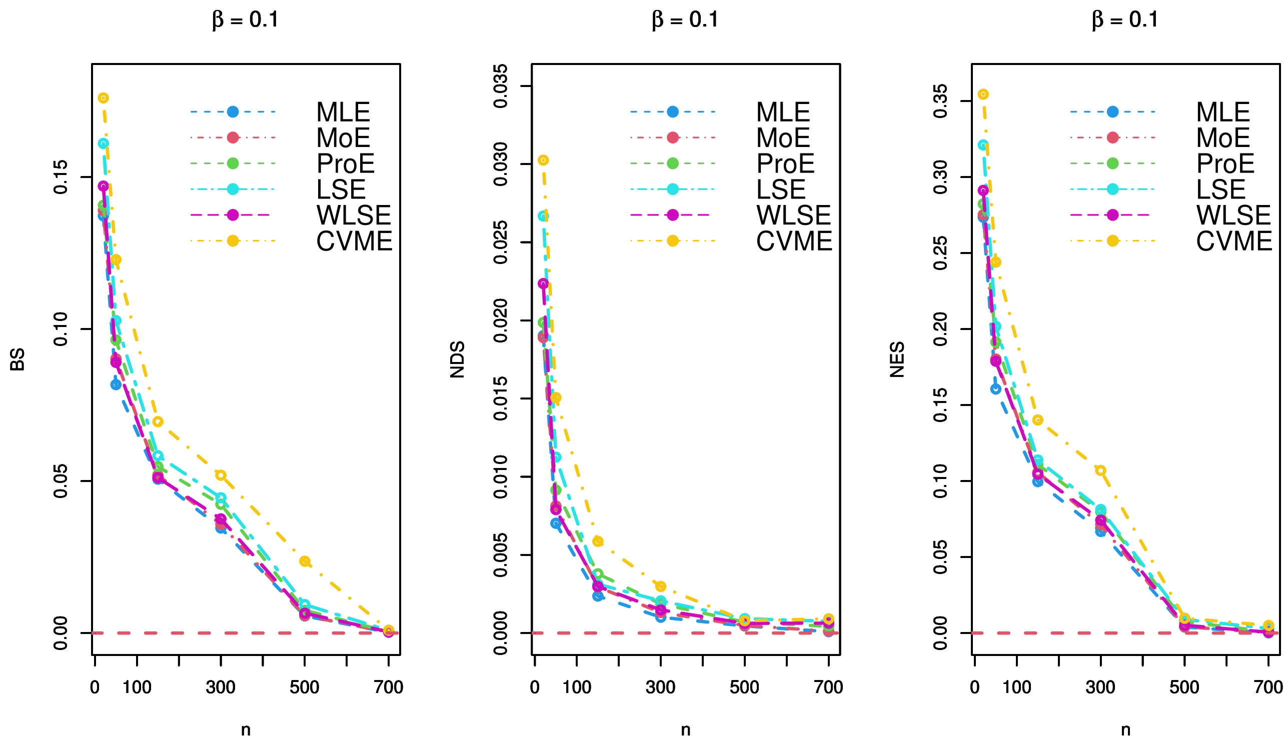

Figure 3. Simulation results for $\beta = 0.1$ under various estimation techniques.

Table 3. The performance of different estimators.

p	Criteria	MLE	MoE	ProE	LSE	WLSE	CVME
				$\beta = 0.6$			
20	$\lvert$**BS**$\rvert$	0.16739517	0.19123657	0.19256701	0.21455452	0.21103266	0.22795025
	NDS	0.02896965	0.03569438	0.03787250	0.04565745	0.04434692	0.05116176
	NES	0.33136947	0.38294372	0.38442251	0.42756685	0.42136644	0.44823182
50	$\lvert$**BS**$\rvert$	0.10426959	0.11701633	0.11922424	0.12896521	0.12836432	0.14841683
	NDS	0.01696430	0.01731907	0.01795875	0.01775027	0.01736494	0.02125485
	NES	0.21236497	0.23410265	0.23745930	0.25856052	0.25736642	0.29865753
150	$\lvert$**BS**$\rvert$	0.05763954	0.06421943	0.06665305	0.07428507	0.07465820	0.08778526
	NDS	0.00436199	0.00549613	0.00570231	0.00754276	0.00634669	0.00855237
	NES	0.11436947	0.12734927	0.13425205	0.14827045	0.14871807	0.17554575
300	$\lvert$**BS**$\rvert$	0.04236943	0.04436940	0.04802549	0.04982567	0.04845854	0.06448586
	NDS	0.00236497	0.00231961	0.00362864	0.00245550	0.00341525	0.00493678
	NES	0.08369456	0.08731057	0.09415758	0.09892364	0.09929757	0.12796238
500	$\lvert$**BS**$\rvert$	0.01236686	0.01412569	0.01523694	0.02136951	0.01632554	0.02853226
	NDS	0.00096544	0.00123999	0.00250255	0.00336951	0.00273036	0.00412559
	NES	0.00366941	0.00413254	0.00441293	0.00604412	0.00463351	0.00715669
700	$\lvert$**BS**$\rvert$	0.00083258	0.00093105	0.00099649	0.00203697	0.00102369	0.00234170
	NDS	0.00001556	0.00007301	0.00009641	0.00085994	0.00009923	0.00091025
	NES	0.00013697	0.00053087	0.00074694	0.00421588	0.00079652	0.00472698

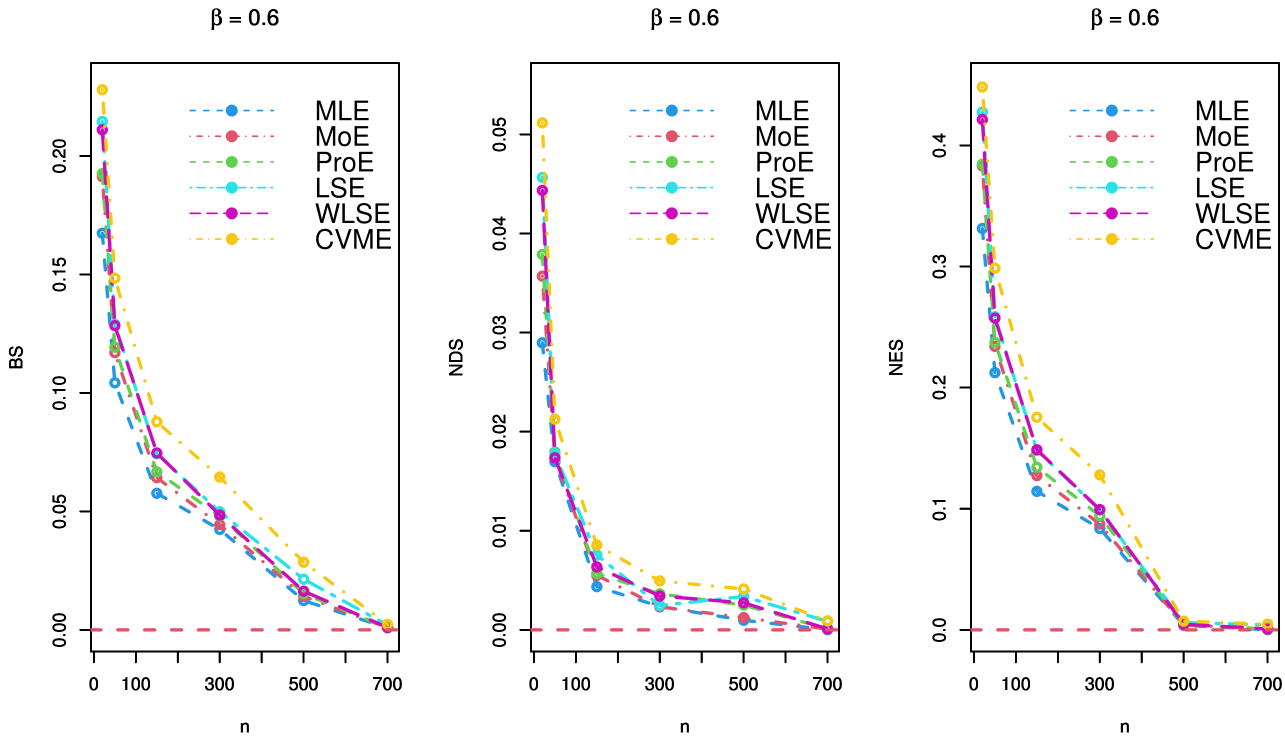

Figure 4. Simulation results for $\beta = 0.6$ under various estimation techniques.

Table 4. The performance of different estimators.

p	Criteria	MLE	MoE	ProE	LSE	WLSE	CVME
				$\beta = 0.9$			
20	**|BS|**	0.55453838	0.55688935	0.57945845	0.68858486	0.60385178	0.71258585
	NDS	0.30771875	0.30893134	0.33786846	0.47102581	0.36361889	0.53195785
	NES	0.36985850	0.37569634	0.38686858	0.45938674	0.40222177	0.48563588
50	**|BS|**	0.34567787	0.36515095	0.36688197	0.42327560	0.39353507	0.49595470
	NDS	0.12122575	0.13123588	0.13453566	0.17963740	0.15409750	0.24712803
	NES	0.23215873	0.24102733	0.24454586	0.28255808	0.26235343	0.33013842
150	**|BS|**	0.19361808	0.21388834	0.20184873	0.24152838	0.22387776	0.31747785
	NDS	0.03868834	0.04653909	0.04151596	0.05877181	0.04958287	0.09896941
	NES	0.13014480	0.14287981	0.13786848	0.16112668	0.14905561	0.20913451
300	**|BS|**	0.13589894	0.15115909	0.14515782	0.18074655	0.14912518	0.22012085
	NDS	0.01903822	0.02358783	0.02486380	0.03357548	0.02298567	0.04784828
	NES	0.08989239	0.10083913	0.09718184	0.12161379	0.09790818	0.14589624
500	**|BS|**	0.02369435	0.05362884	0.05012889	0.07412699	0.05369966	0.09102589
	NDS	0.00236898	0.00363669	0.00389025	0.00441258	0.00341025	0.00992369
	NES	0.00985754	0.02155836	0.01023369	0.02410398	0.01127036	0.04036955
700	**|BS|**	0.00412583	0.00720369	0.00623697	0.00922547	0.00710255	0.01230358
	NDS	0.00044584	0.00062690	0.00066302	0.00070369	0.00059366	0.00088603
	NES	0.00054863	0.00080369	0.00063694	0.00094780	0.00071458	0.00203669

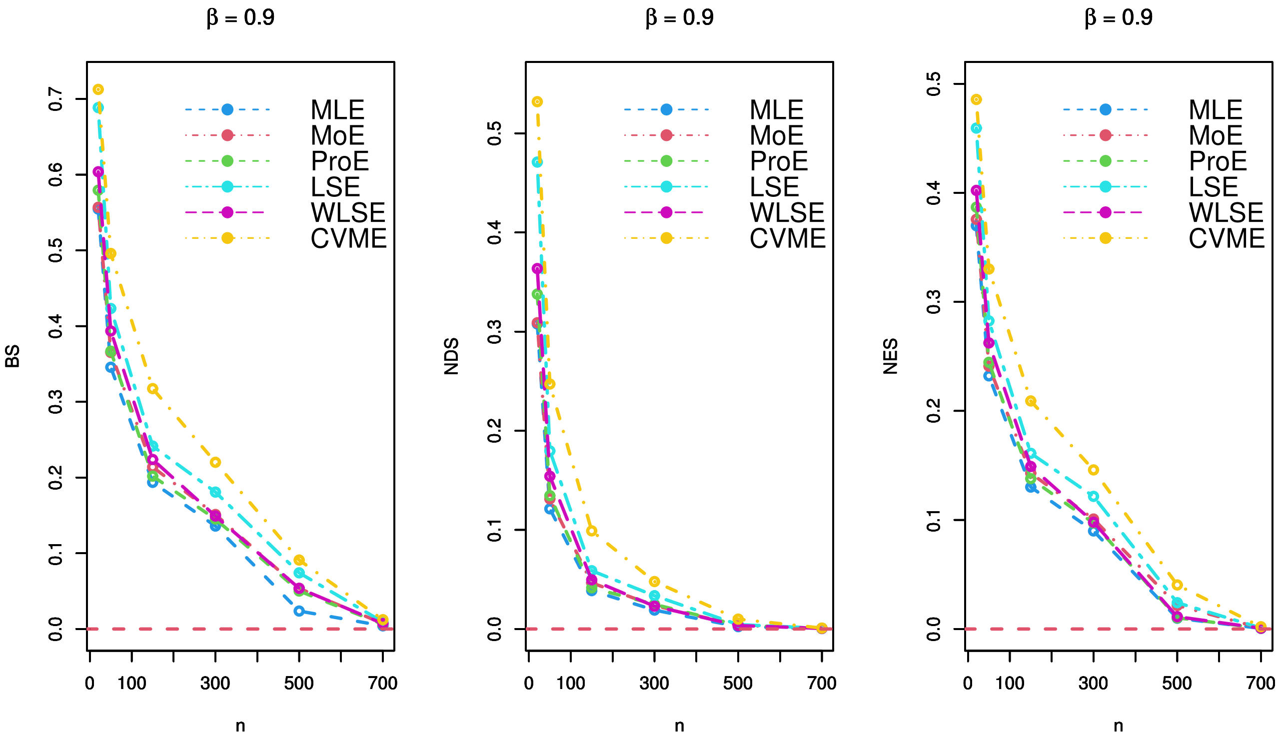

Figure 5. Simulation results for $\beta = 0.9$ under various estimation techniques.

6. A Comparative Study to Model Extreme and Outliers Observations

In this Section, we test the fitting capability of the NDsLE distribution. The fitting of the distributions were compared utilizing some well-known statistical measures, namely, $-L$, the Akaike information criterion (A^*), the correct Akaike information criterion (CA^*), the Hannan–Quinn information criterion (H^*), and the Kolmogorov–Smirnov (K–S) test as well as the chi-square (χ^2) test with its degree of freedom (DF), and the associated p-value (PV). The competitive models (CMs) are provided in Table 5.

Table 5. The CMs of the NDsLE model.

Distribution	Abbreviation	Author(s)
Discrete inverse Rayleigh	DsIR	Hussain and Ahmad [15]
Discrete Pareto	DsPa	Krishna and Pundir [2]
Poisson	Poi	Poisson [16]
Discrete Burr-Hatke	DsBH	El-Morshedy et al. [17]
Discrete Burr type II	DsB-II	Para and Jan [8]
Discrete inverse Weibull	DsIW	Jazi et al. [18]
Discrete log-logistic	DsLog-L	Para and Jan [19]

6.1. Data Set I

This data set represents the number of European red mites on apple leaves (see Chakraborty and Chakravarty [20]). The initial mass shape is reported using nonparametric approaches such as strip, box, violin, and QQ plots in Figure 6. It is noted that the data are asymmetric and some extreme observations are found. Table 4 reports the MLEs with its standard errors (Std-er), and confidence interval (C.I) for the NDsLE parameter and other CMs. Further, the goodness-of-fit (GOF) measures as well as expected (observed) frequency, say EF (OF), have been listed in the same Table.

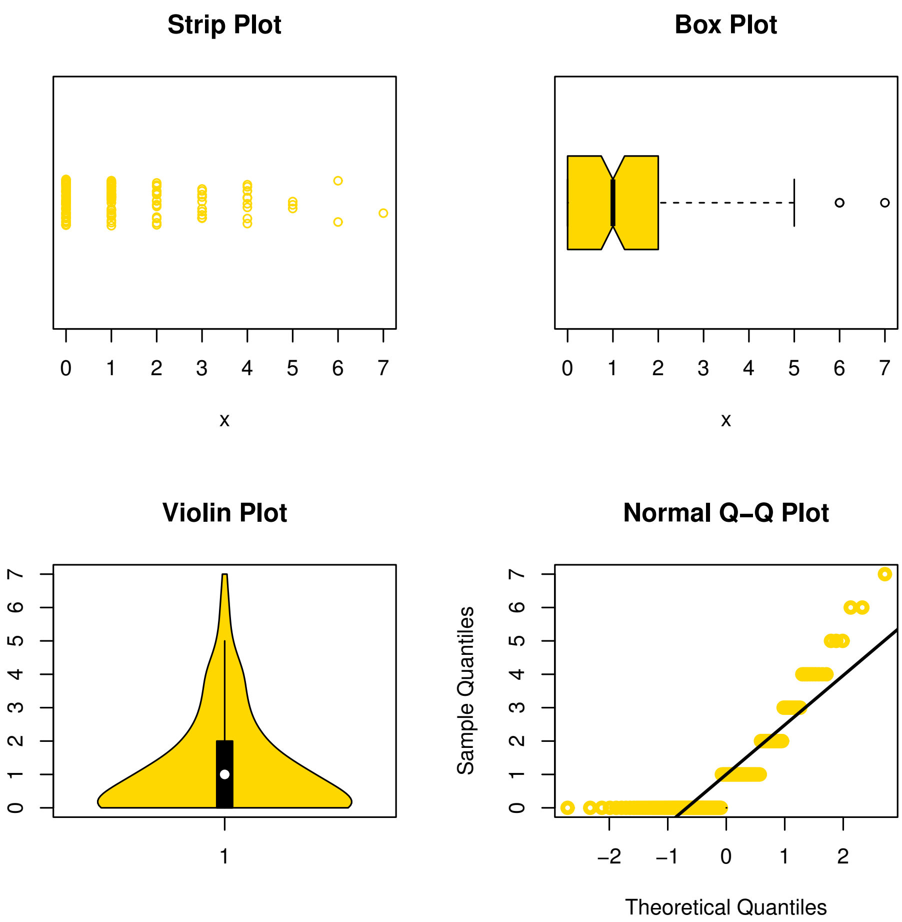

Figure 6. Nonparametric plots for data set I.

From Table 6, the NDsLE model provides the best fit among all CMs because it has the smallest value among $-LL$, A*, CA*, B*, H*, and χ^2 as well as the highest PV. The empirical PMF, PP, and CDF plots for data set I are displayed in Figures 7, 8, and 9, respectively, which indicates that the data set I plausibly came from the NDsLE model.

Table 6. The MLEs, Std-er, C.I, and GOF measures for data set I.

No.		EF							
X	**OF**	**NDsLE**	**DsIR**	**DsPa**	**Poi**	**DsBH**	**DsB-II**	**DsIW**	**DsLogL**
0	70	63.608	65.658	88.308	47.654	88.938	70.469	68.411	67.527
1	38	42.291	56.351	25.005	54.643	27.919	43.053	45.814	44.099
2	17	23.019	14.835	11.314	31.329	12.905	16.214	15.307	17.266
3	10	11.412	5.608	6.312	11.975	7.056	7.364	6.935	7.874
4	9	5.359	2.673	3.972	3.433	4.238	3.924	3.777	4.167
5	3	2.428	1.473	2.705	0.787	2.702	2.338	2.311	2.458
6	2	1.073	0.895	1.948	0.150	1.795	1.509	1.530	1.569
7	1	0.810	2.507	10.436	0.029	4.447	5.129	5.915	5.040
Total	**150**	**150**	**150**	**150**	**150**	**150**	**150**	**150**	**150**
MLE for β		0.382	0.438	0.278	1.147	0.814	0.400	0.456	1.116
Std-er		0.018	0.041	0.029	0.087	0.040	0.044	0.041	0.096
95% C.I	**Lower**	0.347	0.358	0.219	0.975	0.735	0.315	0.376	0.927
	Upper	0.417	0.518	0.336	1.318	0.893	0.486	0.536	1.304
MLE for α		–	–	–	–	–	1.882	1.527	1.829
Std-er		–	–	–	–	–	0.251	0.157	0.177
95% C.I	**Lower**	–	–	–	–	–	1.391	1.219	1.482
	Upper	–	–	–	–	–	2.373	1.834	2.176
$-L$		223.799	233.142	238.832	242.809	230.552	227.727	229.333	227.265
A*		449.597	468.284	479.663	487.619	463.103	459.454	462.666	458.531
CA*		449.624	468.311	479.690	487.647	463.130	459.536	462.747	458.613
B*		452.608	471.295	482.674	490.631	466.114	465.476	468.687	464.552
H*		450.820	469.507	480.886	488.843	464.326	461.901	465.112	460.977
χ^2		5.764	17.376	26.916	26.646	15.573	8.829	11.306	7.843
DV		3	3	4	2	4	2	3	2
PV		0.124	$\leq$0.001	$\leq$0.001	$\leq$0.001	0.004	0.0121	0.010	0.019

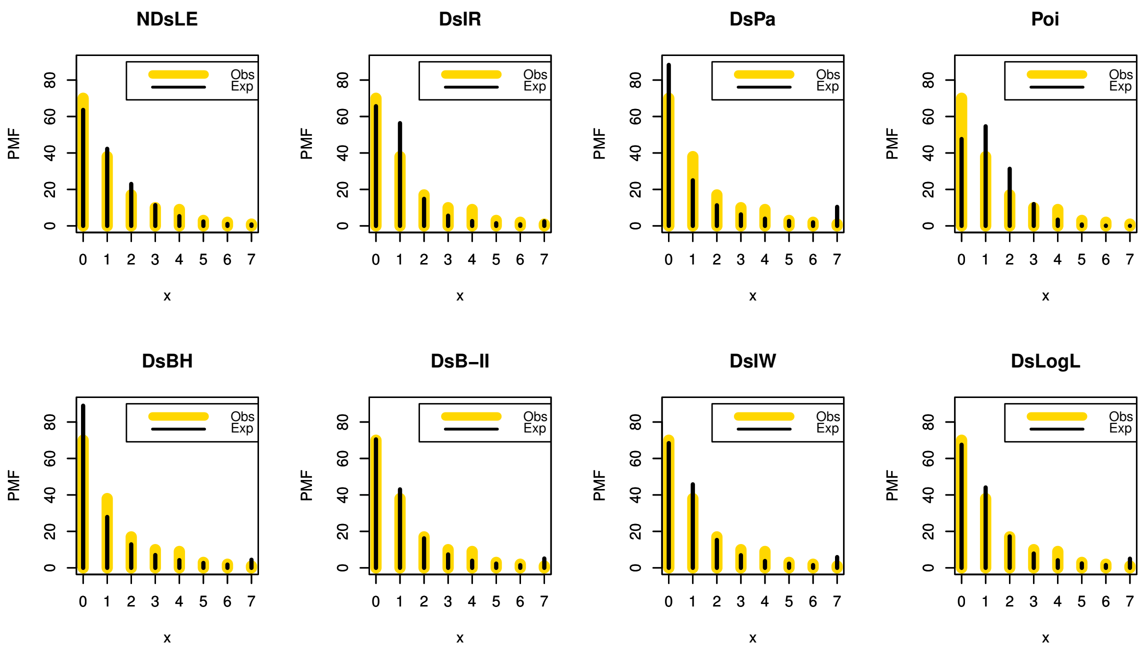

Figure 7. The empirical PMFs plots for data set I.

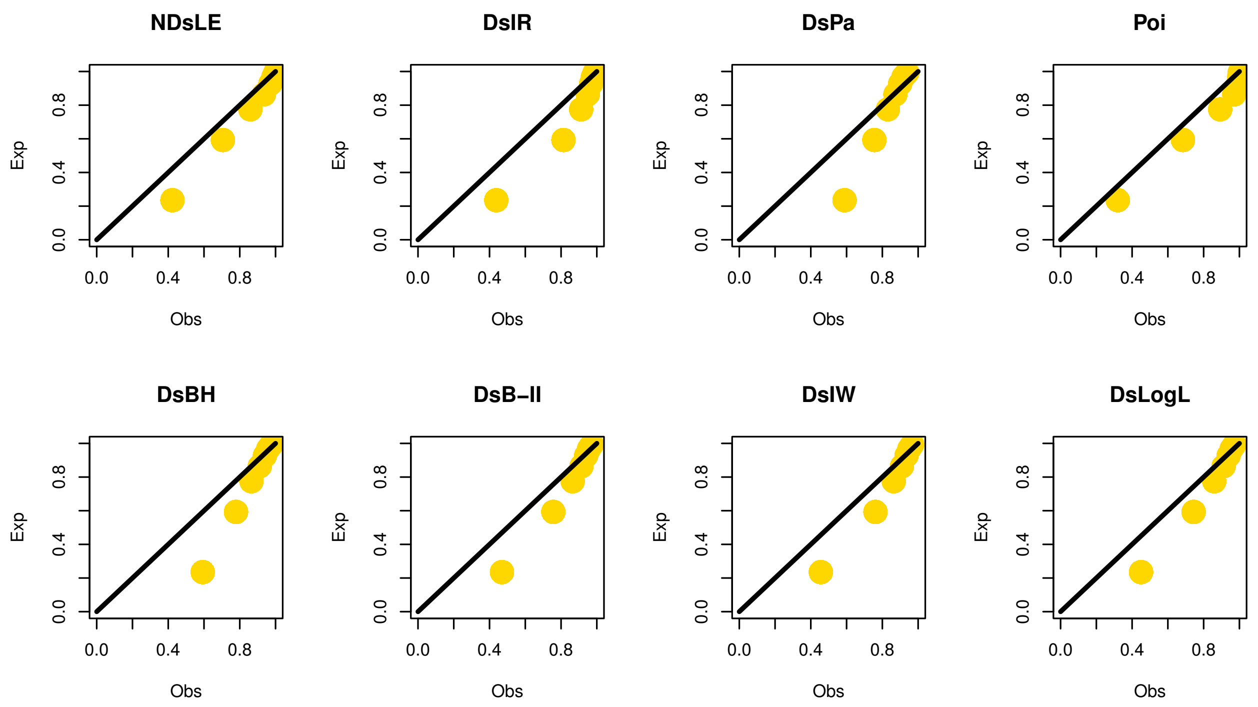

Figure 8. The empirical PPs plots for data set I.

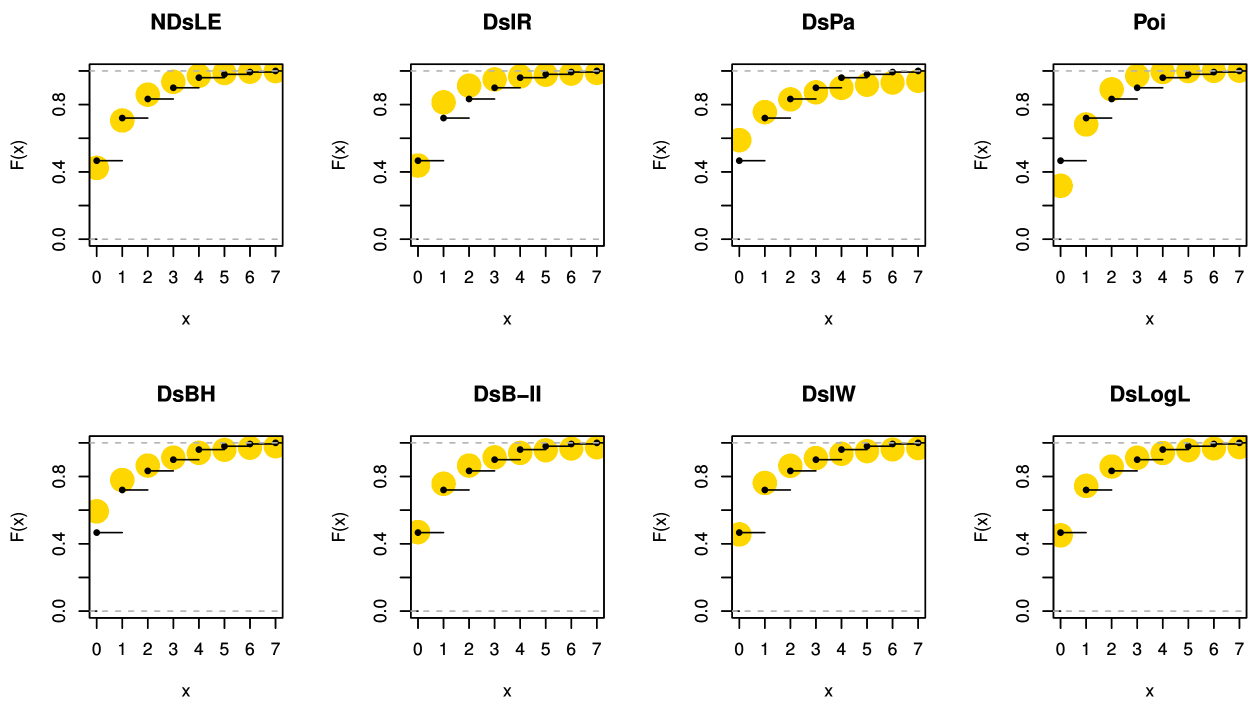

Figure 9. The empirical CDFs plots for data set I.

Table 7 lists different estimators for data set I, and it was found that the MLE and MoE techniques worked quite well for modeling data set I.

Table 7. CVE for data set I.

No.		EF					
X	**OF**	**MLE**	**MoE**	**ProE**	**LSE**	**WLSE**	**CVME**
0	70	63.608	62.767	70.094	42.507	46.324	42.507
1	38	42.291	42.297	41.970	39.659	40.598	39.659
2	17	23.019	23.265	21.034	27.850	27.237	27.850
3	10	11.412	11.645	9.669	17.399	16.328	17.399
4	9	5.359	5.518	4.223	10.195	9.196	10.195
5	3	2.428	2.523	1.783	5.736	4.977	5.736
6	2	1.073	1.125	0.735	3.138	2.620	3.138
7	1	0.810	0.86	0.492	3.516	2.720	3.516
Total	**150**	**150**	**150**	**150**	**150**	**150**	**150**
β		0.382	0.385	0.357	0.469	0.452	0.469
χ^2		5.764	5.657	9.501	28.677	20.378	28.677
DF		3	3	3	5	4	5
PV		0.124	0.130	0.023	$<$0.001	$<$0.001	$<$0.001

Table 8 lists some numerical accounts of empirical and theoretical descriptive statistics. It is noted that all scales were approximately equal.

Table 8. Descriptive statistics for data set I.

	$\mathbb{E}(X)$	$\mathbb{V}(X)$	$\mathbb{I}(X)$	$\mathbb{C}(X)$	$\mathbb{S}(X)$
Data	1.146667	2.273647	1.982831	1.314996	1.544539
MLE	1.127071	1.879304	1.667423	1.216318	1.679096
MoE	1.145691	1.918175	1.674252	1.208862	1.672065
ProE	0.980188	1.579744	1.611674	1.282283	1.742105
LSE	1.761452	3.321649	1.885745	1.034680	1.519391
WLSE	1.621238	2.981079	1.838767	1.064976	1.543652
CVME	1.761452	3.321649	1.885745	1.034680	1.519391

6.2. Data Set II

This data set was given by Karlis and Xekalaki ([21]) and represents the numbers of fires in Greece for the period from 1 July 1998 to 31 August 1998. The strip, box, violin, and QQ plots are displayed in Figure 10, and we can see that data set II is asymmetric and has some outlier observations. Table 9 introduces the MLEs with its Std-er, and C.I for the model parameter and other CMs. Moreover, the GOF measures are shown in the same Table.

From Table 6, the NDsLE model provided the best fit among all CMs. The empirical PP and CDF plots for data set II are displayed in Figures 11 and 12, respectively.

Table 10 reports various estimators for data set II, and it was found that the MLE, MoE, LSE, WLSE, and CVME methods worked quite well for modeling data set II, but the MoE approach was the best.

Table 11 reports some numerical accounts of empirical and theoretical descriptive statistics. It was found that all scales were approximately equal except that of the ProE approach.

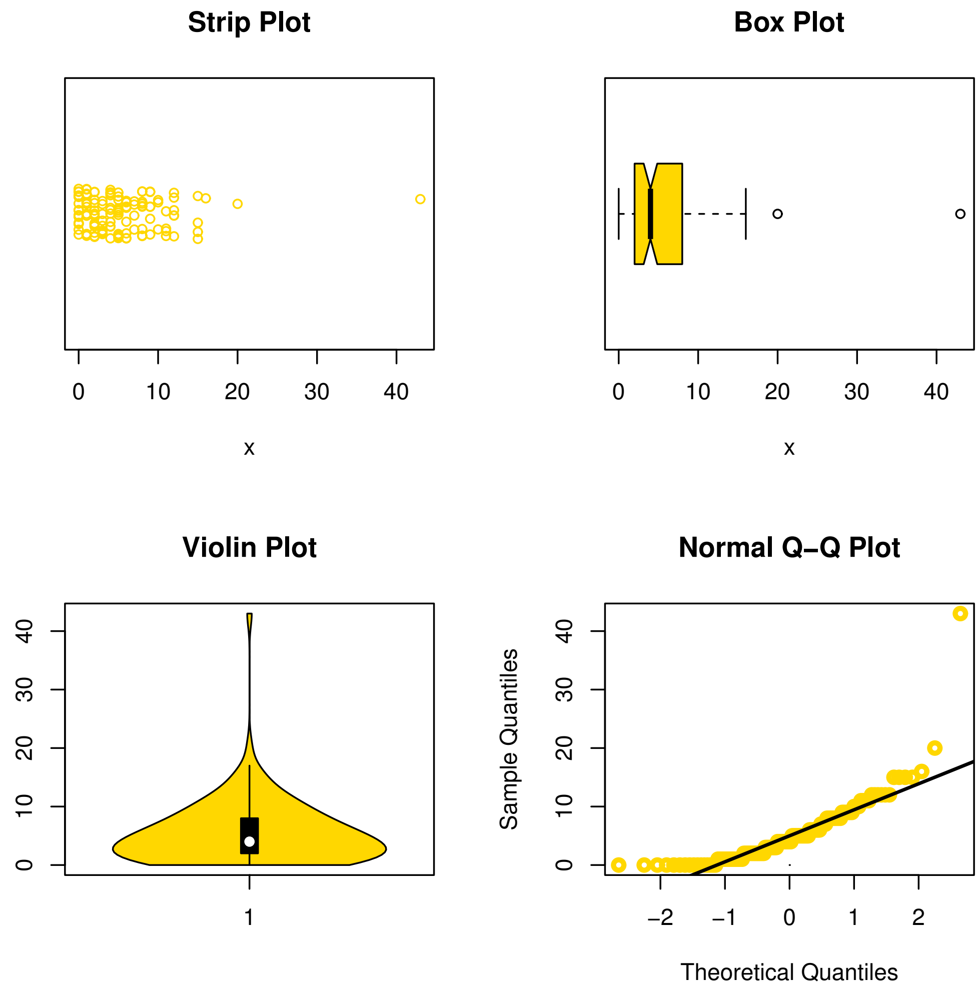

Figure 10. Nonparametric plots for data set II.

Table 9. The MLEs, Std-er, C.I, and GOF measures for data set II.

		NDsLE	DsIR	DsPa	Poi	DsBH	DsB-II	DsIW	DsLogL
MLE for β		0.714	0.018	0.546	5.398	0.984	0.761	0.079	4.226
Std-er		0.015	0.007	0.029	0.209	0.013	0.043	0.022	0.389
95% C.I	**Lower**	0.685	0.004	0.488	4.988	0.959	0.677	0.035	3.462
	Upper	0.743	0.033	0.605	5.809	1	0.845	0.123	4.989
MLE for α		–	–	–	–	–	2.503	1.035	1.717
Std-er		–	–	–	–	–	0.487	0.079	0.138
95% C.I	**Lower**	–	–	–	–	–	1.548	0.881	1.446
	Upper	–	–	–	–	–	3.457	1.189	1.988
$-L$		346.003	412.72	389.635	467.827	407.155	373.393	361.155	346.890
$\mathbf{A}^*$		694.005	827.440	781.271	937.655	816.311	750.786	726.310	697.780
$\mathbf{CA}^*$		694.038	827.473	781.304	937.688	816.344	750.886	726.410	697.880
$\mathbf{B}^*$		696.817	830.252	784.083	940.467	819.123	756.410	731.934	703.405
$\mathbf{H}^*$		695.148	828.582	782.413	938.797	817.453	753.071	728.595	700.065
K–S		0.089	0.429	0.355	0.2547	0.547	0.299	0.208	0.149
PV		0.281	$0 \leq 0.001$	$0 \leq 0.001$	$\leq$0.001	$0 \leq 0.001$	$\leq$0.0010	$\leq$0.0010	0.009

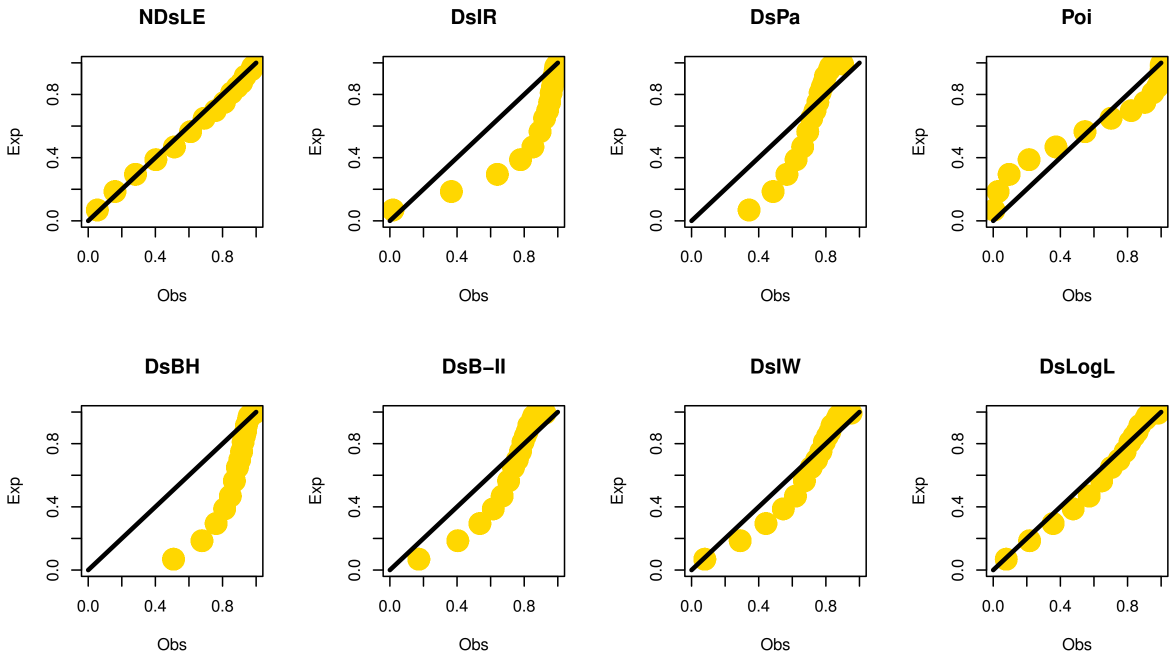

Figure 11. The empirical PPs plots for data set II.

Table 10. CVE for data set II.

	MLE	MoE	ProE	LSE	WLSE	CVME
β	0.714	0.717	0.599	0.729	0.733	0.729
K–S	0.089	0.082	0.318	0.093	0.099	0.093
PV	0.281	0.376	$\leq$0.001	0.236	0.173	0.236

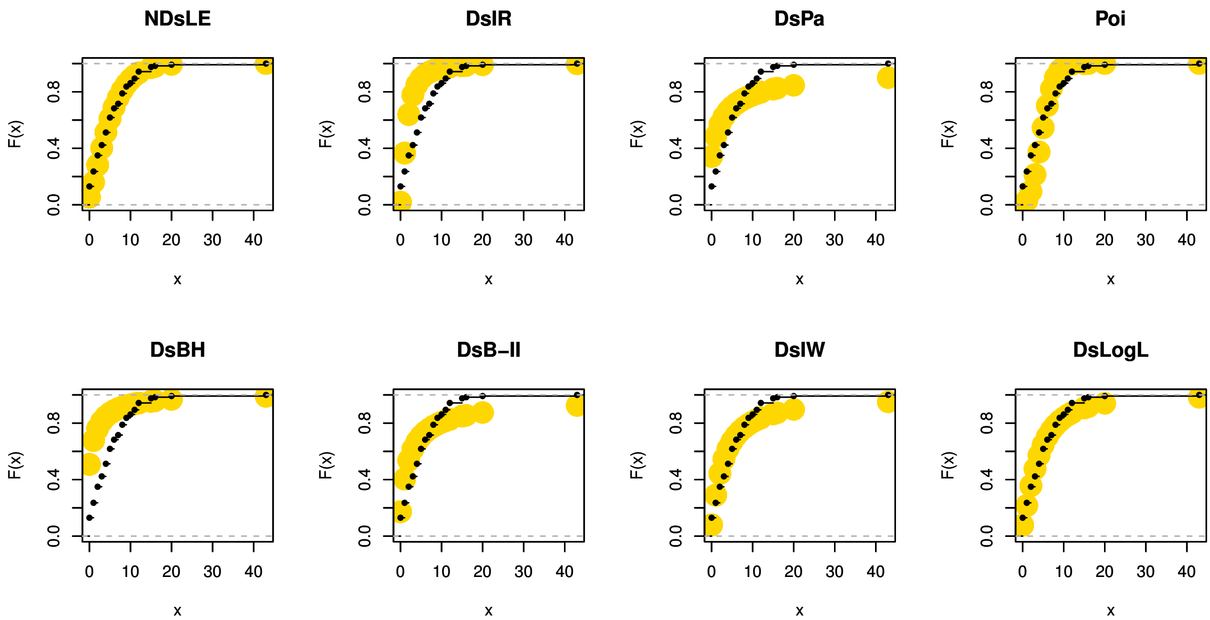

Figure 12. The empirical CDFs plots for data set II.

Table 11. Descriptive statistics for data set II.

	$\mathbb{E}(X)$	$\mathbb{V}(X)$	$\mathbb{I}(X)$	$\mathbb{C}(X)$	$\mathbb{S}(X)$
Data	5.398373	30.044915	5.565550	1.015366	3.028614
MLE	5.328796	17.681101	3.318029	0.789088	1.408254
MoE	5.406128	18.129156	3.353445	0.787594	1.408258
ProE	3.176367	7.603379	2.393734	0.868106	1.421522
LSE	5.731521	20.080603	3.503538	0.781841	1.408353
WLSE	5.846112	20.793291	3.556773	0.779999	1.408410
CVME	5.731521	20.080603	3.503538	0.781841	1.408353

6.3. Data Set III

The data were reported in https://www.worldometers.info/coronavirus/country/south-korea/ (accessed on 14 February 2022) and represent the daily new deaths in South Korea for COVID-19 from 15 February to 12 December 2020. In Figure 13, we can see the data are asymmetric, and some extreme observations are present. Table 12 lists the MLEs with its Std-er, GOF, and C.I for the NDsLE parameter and other CMs.

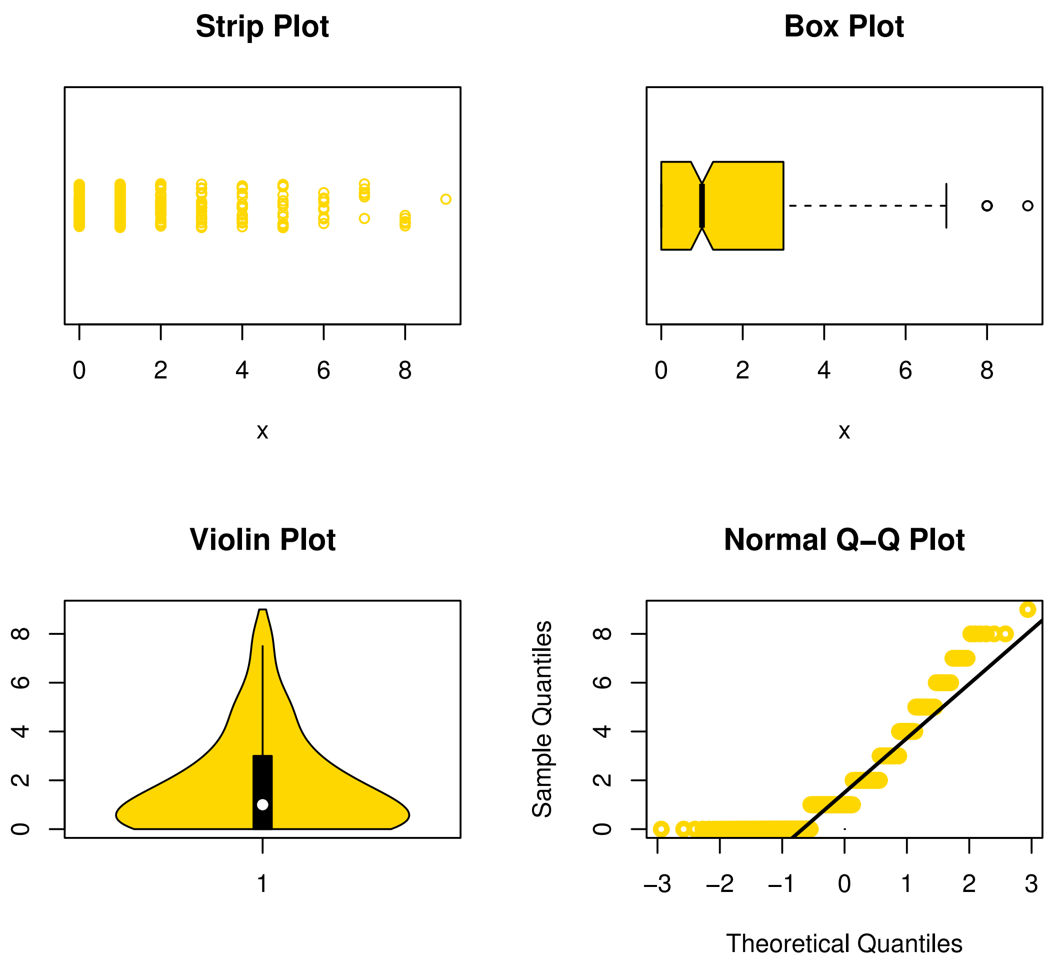

Figure 13. Nonparametric plots for data set III.

Table 12. The MLEs, Std-er, C.I, and GOF measures for data set III.

No. X		OF	NDsLE	DsIR	DsPa	Poi	DsBH	DsB-II	DsIW	DsLogL
			EF							
0		89	80.319	69.890	149.356	45.408	166.600	92.887	82.351	80.931
1		79	78.626	140.613	50.500	86.336	54.598	97.788	103.702	92.775
2		49	57.163	47.685	25.478	82.074	26.667	42.676	44.390	51.431
3		29	36.851	19.125	15.379	52.014	15.540	21.174	22.317	27.336
4		19	22.251	9.333	10.312	24.725	10.017	12.172	12.926	15.559
5		17	12.891	5.198	7.387	9.402	6.885	7.754	8.248	9.518
6		9	7.259	3.162	5.533	2.979	4.953	5.308	5.636	6.193
7		6	4.004	2.098	4.408	0.809	3.682	3.831	4.052	4.247
8		6	2.173	1.429	3.466	0.192	2.808	2.879	3.031	3.061
9		1	2.463	5.467	32.181	0.061	12.25	17.531	17.347	12.949
Total		**304**	**304**	**304**	**304**	**304**	**304**	**304**	**304**	**304**
MLE for β			0.482	0.229	0.377	1.901	0.904	0.591	0.271	1.716
Std-er			0.012	0.023	0.021	0.079	0.020	0.031	0.025	0.095
95% C.I	**Lower**		0.459	0.184	0.335	1.746	0.864	0.529	0.221	1.529
	Upper		0.506	0.276	0.419	2.056	0.944	0.653	0.321	1.902
MLE for α			–	–	–	–	–	2.466	1.411	1.878
Std-er			–	–	–	–	–	0.248	0.083	0.107
95% C.I	**Lower**		–	–	–	–	–	1.979	1.248	1.668
	Upper		–	–	–	–	–	2.953	1.575	2.087
$-L$			566.579	606.870	633.531	621.098	620.466	587.652	586.855	577.011
A*			1135.157	1215.740	1269.061	1244.195	1242.932	1179.304	1177.711	1158.023
CA*			1135.171	1215.754	1269.075	1244.208	1242.945	1179.344	1177.751	1158.063
B*			1138.874	1219.457	1272.778	1247.912	1246.649	1186.738	1185.145	1165.457
H*			1136.644	1217.227	1270.548	1245.682	1244.419	1182.278	1180.684	1160.997
χ^2			8.181	92.204	128.631	115.896	109.333	44.784	41.868	25.019
DV			6	6	7	4	6	6	6	6
PV			0.225	$\leq$0.001	$\leq$0.001	$\leq$0.001	$\leq$0.001	$\leq$0.001	$\leq$0.001	$\leq$0.001

From Table 12, the NDsLE model provides the best fit among all CMs. The empirical PMF, PP, and CDF plots for data set III are displayed in Figures 14, 15, and 16, respectively.

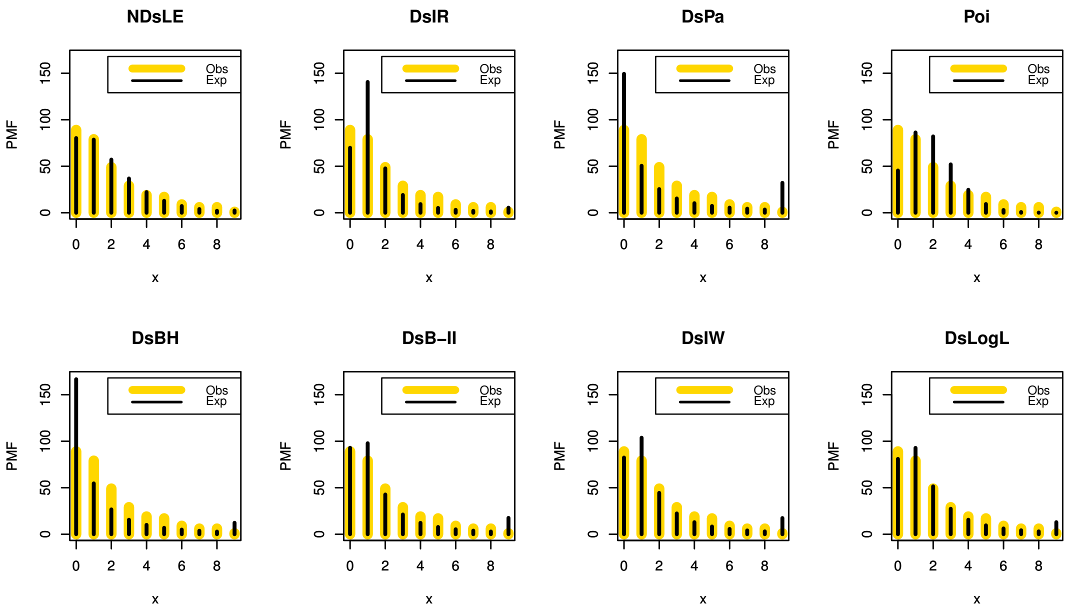

Figure 14. The empirical PMFs plots for data set III.

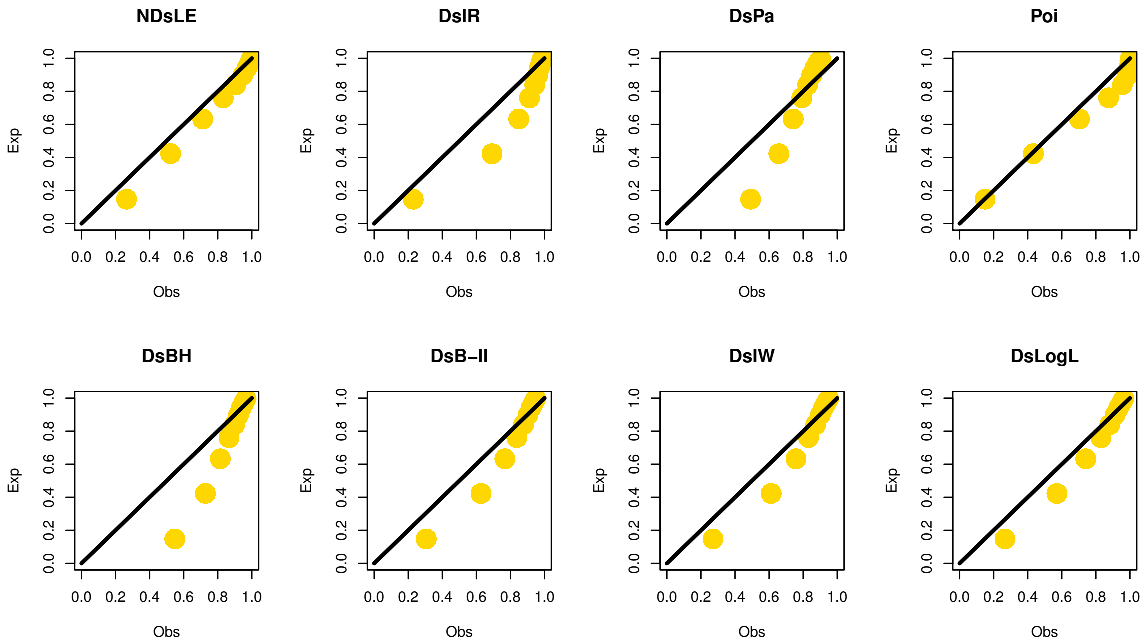

Figure 15. The empirical PPs plots for data set III.

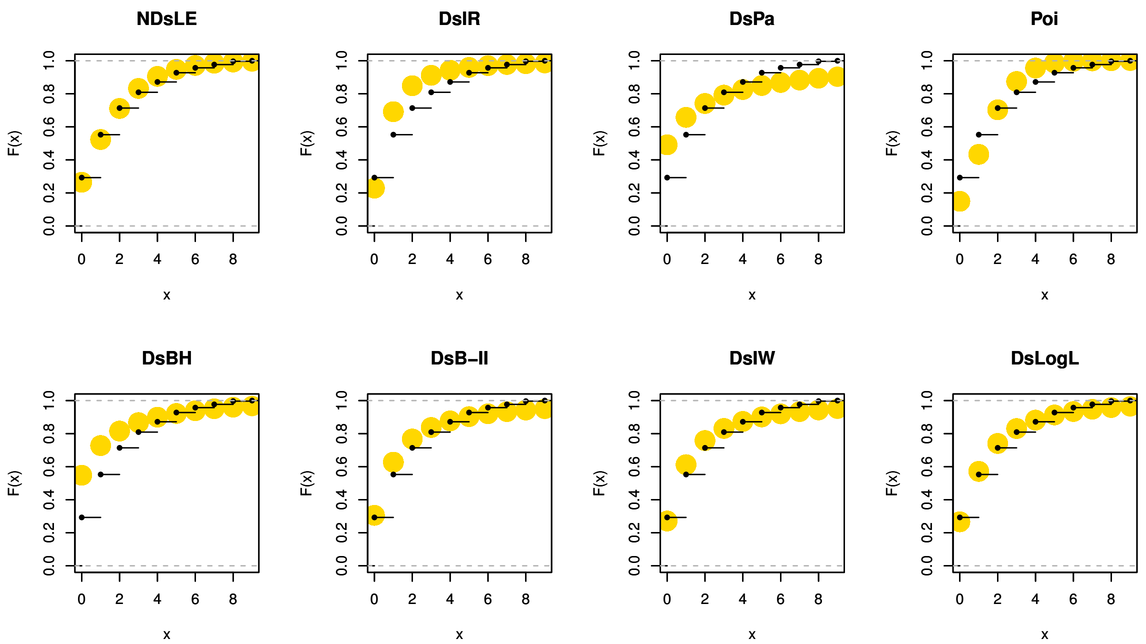

Figure 16. The empirical CDFs plots for data set III.

Table 13 lists various estimators for data set III; it was found that the MLE and MoE techniques worked quite well for modeling data set III, but the MLE method was the best.

Table 14 listed some numerical accounts of empirical and theoretical descriptive statistics. It is clear that all scales were approximately equal.

Table 13. CVE for data set III.

No.		EF					
X	**OF**	**MLE**	**MoE**	**ProE**	**LSE**	**WLSE**	**CVME**
0	89	80.319	80.467	88.839	57.517	60.347	57.517
1	79	78.626	78.674	81.090	68.763	70.282	68.763
2	49	57.163	57.147	56.045	57.568	57.777	57.568
3	29	36.851	36.812	34.515	42.163	41.624	42.163
4	19	22.251	22.210	19.946	28.771	27.957	28.771
5	17	12.891	12.858	11.071	18.790	17.977	18.790
6	9	7.259	7.235	5.976	11.910	11.222	11.910
7	6	4.004	3.988	3.161	7.387	6.855	7.387
8	6	2.173	2.163	1.646	4.507	4.119	4.507
9	1	2.463	2.446	1.711	6.624	5.840	6.624
Total	**304**	**304**	**304**	**304**	**304**	**304**	**304**
β		0.482	0.485	0.463	0.541	0.533	0.541
χ^2		8.181	8.209	13.017	30.135	24.197	30.135
DV		6	6	6	7	7	7
PV		0.225	0.223	0.042	$\leq$0.001	0.001	$\leq$0.001

Table 14. Descriptive statistics for data set III.

	$\mathbb{E}(X)$	$\mathbb{V}(X)$	$\mathbb{I}(X)$	$\mathbb{C}(X)$	$\mathbb{S}(X)$
Data	1.901316	4.122242	2.168099	1.067856	1.263883
MLE	1.874529	3.606015	1.923690	1.013028	1.502944
MoE	1.901373	3.674830	1.932725	1.008211	1.499399
ProE	1.710994	3.197583	1.868845	1.045110	1.527590
LSE	2.459262	5.227293	2.125553	0.929680	1.448801
WLSE	2.372434	4.969632	2.094739	0.939653	1.454364
CVME	2.459263	5.227294	2.125553	0.929680	1.448801

6.4. Data Set IV

The data set represents the leukemia remission times (in weeks) for 20 patients (see Damien and Walker, [22]). The data are asymmetric-shaped and contain some extreme values (see Figure 17).

Table 15 introduces the MLEs with its Std-er, GOF, and C.I for the NDsLE parameter and other CMs.

Table 15. The MLEs, Std-er, C.I, and GOF measures for data set IV.

		NDsLE	DsIR	DsPa	Poi	DsBH	DsB-II	DsIW	DsLogL
MLE for β		0.869	3.374×10^{-7}	0.655	13.750	0.997	0.996	0.004	9.602
Std-er		0.019	0.035	0.0619	0.829	0.012	0.098	0.007	2.326
95% C.I	**Lower**	0.831	0.000	0.534	12.125	0.973	0.804	0.000	5.043
	Upper	0.907	0.069	0.777	15.375	1.000	1.000	0.018	14.160
MLE for α		–	–	–	–	–	103.479	1.007	1.639
Std-er		–	–	–	–	–	1.516	0.175	0.302
95% C.I	**Lower**	–	–	–	–	–	100.507	0.664	1.047
	Upper	–	–	–	–	–	106.451	1.351	2.231
$-L$		74.276	85.086	84.582	145.432	94.635	79.973	74.797	74.269
A*		150.551	172.171	171.165	292.865	191.269	163.947	153.593	152.538
CA*		150.773	172.394	171.387	293.087	191.492	164.653	154.299	153.244
B*		151.547	173.167	172.161	293.861	192.265	165.938	155.585	154.529
H*		150.745	172.366	171.359	293.059	191.464	164.335	153.982	152.927
K–S		0.143	0.482	0.372	0.379	0.669	0.334	0.197	0.189
PV		0.809	<0.001	0.008	0.006	<0.001	0.013	0.402	0.498

From Table 15, the NDsLE model provided the best fit among all tested models. The empirical PPs and CDFs plots for data set IV are displayed in Figures 18 and 19, respectively.

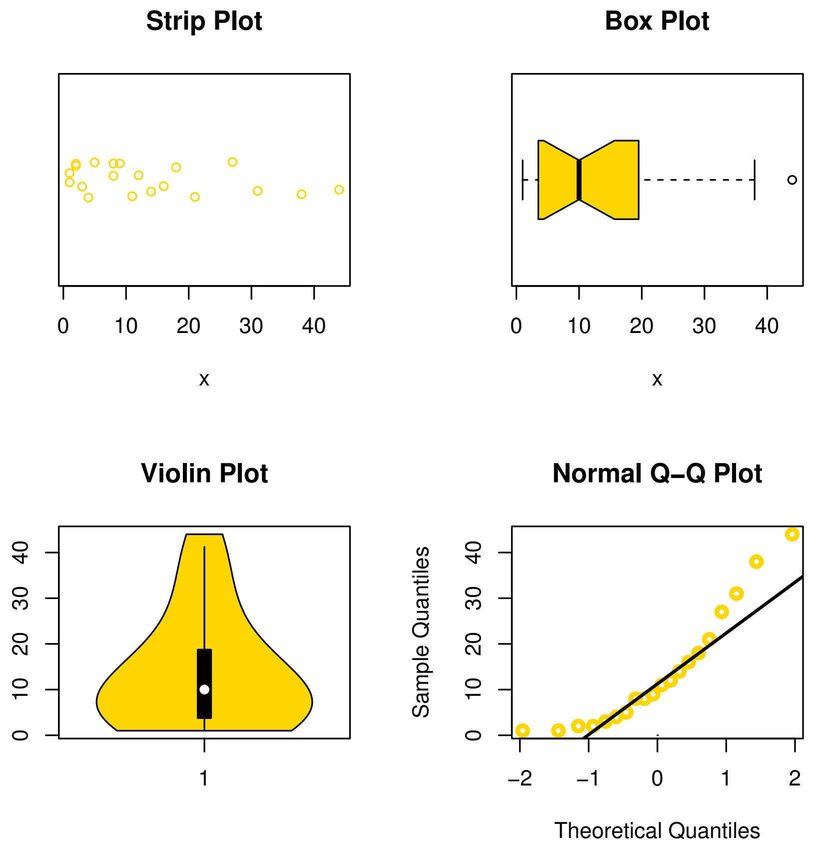

Figure 17. Nonparametric plots for data set IV.

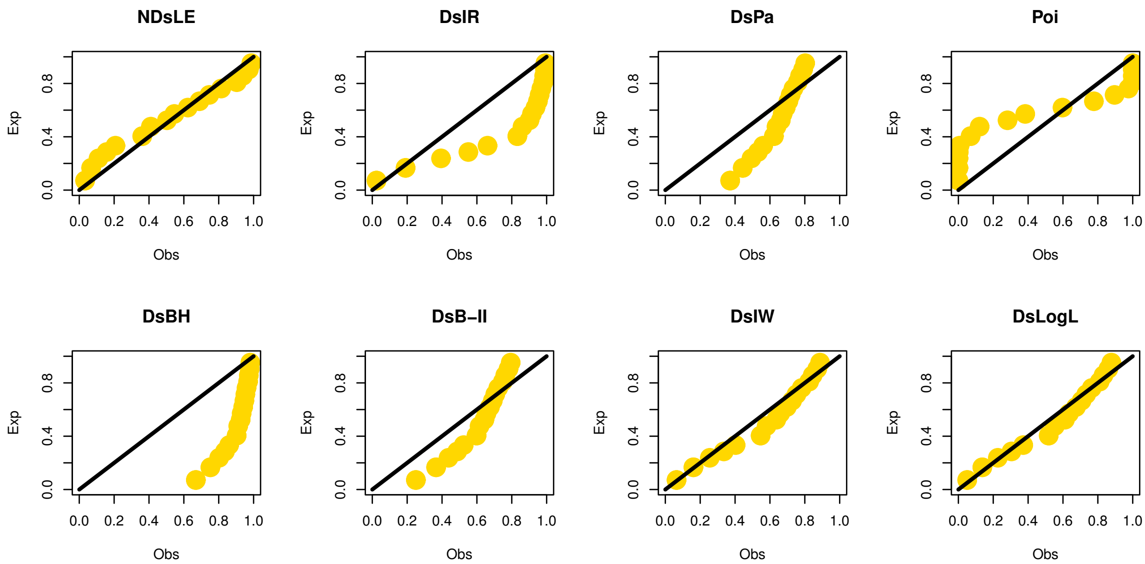

Figure 18. The empirical PPs plots for data set IV.

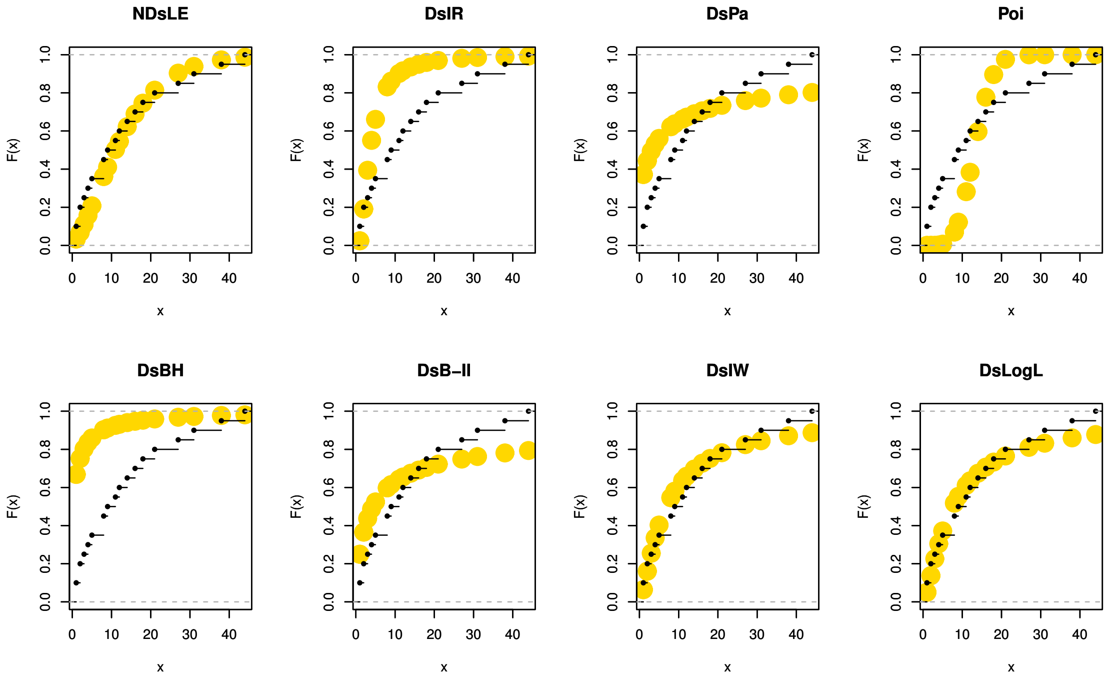

Figure 19. The empirical CDFs plots for data set IV.

Table 16 lists different estimators for data set IV; it was found that the MLE, MoE, ProE, LSE, WLSE and CVME methods worked quite well for modeling data set IV, but the MLE approach was the best.

Table 16. CVE for data set IV.

	MLE	MoE	ProE	LSE	WLSE	CVM
β	0.869	0.849	0.773	0.844	0.856	0.843
K–S	0.143	0.161	0.227	0.156	0.169	0.155
PV	0.809	0.676	0.253	0.718	0.613	0.726

Table 17 report some numerical accounts of empirical and theoretical descriptive statistics. It is noted that all scales are approximately equal.

Table 17. Descriptive statistics for data set IV.

	$\mathbb{E}(X)$	$\mathbb{V}(X)$	$\mathbb{I}(X)$	$\mathbb{C}(X)$	$\mathbb{S}(X)$
Data	1.901316	4.122242	2.168099	1.067856	1.263883
MLE	1.874529	3.606015	1.923690	1.013028	1.502944
MoE	1.901373	3.674830	1.932725	1.008211	1.499399
ProE	1.710994	3.197583	1.868845	1.045110	1.527590
LSE	2.459262	5.227293	2.125553	0.929680	1.448801
WLSE	2.372434	4.969632	2.094739	0.939653	1.454364
CVME	2.459263	5.227294	2.125553	0.929680	1.448801

The profiles of L functions for data sets I, II, III and IV are displayed in Figure 20; it was found that the estimator was a unique "unimodal function".

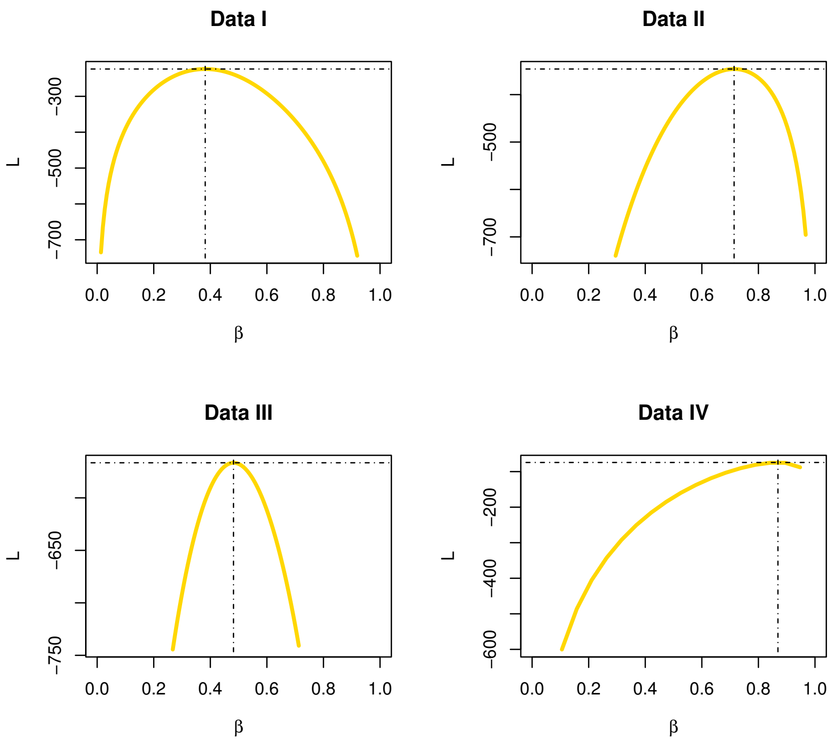

Figure 20. The profiles of L functions for data sets I, II, III and IV.

7. Conclusions

In this paper, a novel discrete model with one parameter called the discrete linear-exponential (NDsLE) model, was introduced. Its various statistical features were derived in detail. It was found that the NDsLE model was a proper model for right-skewed data sets, especially those having extreme observations. Moreover, the NDsLE model provided a wide variation in the shape of the kurtosis, and consequently it could be utilized for modeling different kinds of data. The NDsLE parameter was estimated using different estimation techniques, namely, MLE, MoE, ProE, LSE, WLSE, and CVME. Simulation studies were performed based on different sample sizes, and it was found that the six methods worked quite effectively in estimating the NDsLE parameter. Four data sets were analyzed to illustrate and prove the notability of the NDsLE model. Finally, the NDsLE model would be a better alternative to other lifetime models available in the existing literature, especially, in extreme values fields.

Funding: The author extends their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number IF-PSAU-2021/01/18291.

Data Availability Statement: The four data sets are available in the paper.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Roy, D. Discrete Rayleigh distribution. *IEEE Trans. Reliab.* **2004**, *53*, 255–260. [CrossRef]
2. Krishna, H.; Pundir, P.S. Discrete Burr and discrete Pareto distributions. *Stat. Methodol.* **2009**, *6*, 177–188. [CrossRef]
3. Eliwa, M.S.; El-Morshedy, M.; Yousof, H.M. A discrete exponential generalized-G family of distributions: Properties, with Bayesian and Non-Bayesian estimators to model medical, engineering, and agriculture data. *Mathematics* **2022**, *10*, 3348. [CrossRef]
4. Gómez-Déniz, E.; Calderín-Ojeda, E. The discrete Lindley distribution: Properties and applications. *J. Stat. Comput. Simul.* **2011**, *81*, 1405–1416. [CrossRef]
5. Nekoukhou, V.; Alamatsaz, M.H.; Bidram, H. Discrete generalized exponential distribution of a second type. *Statistics* **2013**, *47*, 876–887. [CrossRef]

6. Nekoukhou, V.; Bidram, H. The exponentiated discrete Weibull distribution. *Sort* **2015**, *39*, 127–146.
7. Hussain, T.; Aslam, M.; Ahmad, M. A two-parameter discrete Lindley distribution. *Rev. Colomb. Estad.* **2016**, *39*, 45–61. [CrossRef]
8. Para, B.A.; Jan, T.R. On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science. *Biom. Biostat. J.* **2016**, *4*, 1–15.
9. Eliwa, M.S.; Alhussain, Z.A.; El-Morshedy, M. Discrete Gompertz-G family of distributions for over-and under-dispersed data with properties, estimation, and applications. *Mathematics* **2020**, *8*, 358. [CrossRef]
10. El-Morshedy, M.; Eliwa, M.S.; Nagy, H. A new two-parameter exponentiated discrete Lindley distribution: Properties, estimation and applications. *J. Appl. Stat.* **2020**, *47*, 354–375. [CrossRef] [PubMed]
11. Yousof, H.M.; Chesneau, C.; Hamedani, G.G.; Ibrahim. M. A new discrete distribution: Properties, characterizations, modeling real count data, Bayesian and non-Bayesian estimations. *Statistica* 2021, *forthcoming*.
12. Ibrahim, M.; Ali, M.M.; Yousof, H.M. The discrete analogue of the Weibull G family: Properties, different applications, Bayesian and non-Bayesian estimation methods. *Ann. Data Sci.* **2021**, 1–38. [CrossRef]
13. Al-Bossly, A.; Eliwa, M.S. Asymmetric probability mass function for count data based on the binomial technique: Synthesis and analysis with inference. *Symmetry* **2022**, *14*, 826. [CrossRef]
14. Sah, B.K. One-parameter linear-exponential distribution. *Maths* **2021**, *6*, 6–15. [CrossRef]
15. Hussain, T.; Ahmad, M. Discrete inverse Rayleigh distribution. *Pak. J. Stat.* **2014**, *30*, 203–222.
16. Poisson, S.D. *Probabilité des Jugements en Mati ère Criminelle et en Matière Civile, Précédées Des rè gles Générales du Calcul des Probabilitiés*; Bachelier: Paris, France, 1837; Volume 1, p. 1837.
17. El-Morshedy, M.; Eliwa, M.S.; Altun, E. Discrete Burr-Hatke distribution with properties, estimation methods and regression model. *IEEE Access* **2020**, *8*, 74359–74370. [CrossRef]
18. Jazi, M.A.; Lai, C.D.; Alamatsaz, M.H. A discrete inverse Weibull distribution and estimation of its parameters. *Stat. Methodol.* **2010**, *7*, 121–132. [CrossRef]
19. Para, B.A.; Jan, T.R. Discrete version of log-logistic distribution and its applications in genetics. *Int. Mod. Math. Sci.* **2016**, *14*, 407–422.
20. Chakraborty, S.; Chakravarty, D. Discrete gamma distributions: Properties and parameter estimations. *Commun. Stat. Theory Methods* **2012**, *41*, 3301–3324. [CrossRef]
21. Karlis, D.; Xekalaki, E.; Lipitakis, E.A. On some discrete valued time series models based on mixtures and thinning. In Proceedings of the Fifth Hellenic-European Conference on Computer Mathematics and Its Applications, Athens, Greece, 20–22 September 2001; pp. 872–877.
22. Damien, P.; Walker, S. A Bayesian non-parametric comparison of two treatments. *Scand. J. Stat.* **2002**, *29*, 51–56. [CrossRef]