axioms

Article

Discrete Single-Factor Extension of the Exponential Distribution: Features and Modeling

Mahmoud El-Morshedy [1,2,*], Hend S. Shahen [2], Bader Almohaimeed [3] and Mohamed S. Eliwa [2,4,5]

[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
[3] Department of Mathematics, College of Science, Qassim University, Buraydah 51482, Saudi Arabia
[4] Department of Statistics and Operation Research, College of Science, Qassim University, Buraydah 51482, Saudi Arabia
[5] Department of Mathematics, International Telematic University Uninettuno, I-00186 Rome, Italy
* Correspondence: m.elmorshedy@psau.edu.sa

Abstract: The importance of counting data modeling and its applications to real-world phenomena has been highlighted in several research studies. The present study focuses on a one-parameter discrete distribution that can be derived via the survival discretization approach. The proposed model has explicit forms for its statistical properties. It can be applied to discuss asymmetric "right skewed" data with long "heavy" tails. Its failure rate function can be used to discuss the phenomena with a monotonically decreasing or unimodal failure rate shape. Further, it can be utilized as a probability tool to model and discuss over- and under-dispersed data. Various estimation techniques are reported and discussed in detail. A simulation study is performed to test the property of the estimator. Finally, three real data sets are analyzed to prove the notability of the introduced model.

Keywords: discretization technique; various estimation techniques; simulation; count data

MSC: 62E99; 62E15

check for updates

Citation: El-Morshedy, M.; Shahen, H.S.; Almohaimeed, B.; Eliwa, M.S. Discrete Single-Factor Extension of the Exponential Distribution: Features and Modeling. *Axioms* **2022**, *11*, 737. https://doi.org/10.3390/axioms11120737

Academic Editors: Jiajuan Liang and Kaitai Fang

Received: 2 November 2022
Accepted: 10 December 2022
Published: 16 December 2022

1. Introduction

The modeling of count data in recent years has been very complicated due to the huge number of data sets generated from various fields over time, particularly in ecology, renewable energy, engineering, and medicine. The main problem occurs when the data suffer from excessive scattering with different types of dispersion forms. To solve this problem, statisticians have introduced flexible probability models that have different types of dispersions to model such data. The moment exponential (MEx) distribution is one of the most popular models for this purpose, especially in the case of an over- or under-dispersed shape for the hydrological processes. The random variable (RV) X is said to follow the MEx distribution if its survival function (SF) is given by

$$S(x;\lambda) = (\lambda x + 1)e^{-\lambda x};\ x > 0, \tag{1}$$

where $\lambda > 0$ is a scale parameter. The MEx model is flexible enough to accommodate monotonic failure rates. On account of the resilience of the MEx model, many statisticians have sought to derive many extension "modifications" from this distribution with its applications in diverse fields. Examples include: Burr XII-MEx (see Bhatti et al. [1]), generalized exponentiated MEx (see Iqbal et al. [2]), Poisson MEx (see, Ahsan-ul-Haq [3]), Topp-Leone MEx (see Abbas et al. [4]), order statistics of exponentiated MEx (see Akhter et al. [5]), Weibull-MEx (see Hashmi et al. [6]), statistical inference of the lower record values based on exponentiated MEx (see Kumar et al. [7]), slashed MEx (see Iriarte et al. [8]), and others.

In many cases, the data need to be recorded or listed on a discrete scale rather than on a continuous analog scale. Due to this reason, the discretization of continuous probability models has received much attention as the census data produced from various regions become more complex by the day. Therefore, to model these counting data, discrete probability models are required for analytical studies of these multidimensional "complex" phenomena. Discretization of a continuous probability model can be derived via different techniques. The most widely utilized approach is the survival discretization method in which the probability mass function (PMF) of the RV X can be formulated as

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

Based on Equation (2), several discrete models have been reported and discussed, i.e., the discrete analog of the Weibull-G class (see Ibrahim et al. [9]), the discrete exponential generalized-G family (see Eliwa et al. [10]), the discrete generalized Burr–Hatke (see Yousof et al. [11]), the discrete Bilal (see Altun [12]), the discrete generalized Lindley (see El-Morshedy [13]), the discrete alpha power inverse Lomax (see Almetwally [14]), the discrete Perks (see Tyagi et al. [15]), the discrete Lindley (see Bakouch et al. [16]), and others.

Although the statistical literature contains a lot of discrete models, more and more discrete probability distributions are needed to discuss complex phenomena with sparse observations. In this paper, we propose the discrete analog of Equation (1). In this paper, the discrete version of the MEx model is derived from the abbreviated version called DMEx. The great advantage of the DMEx model is that it stands with one parameter that has been listed to give a better alternative to some discrete distributions and reports another tribune for statisticians working in the field of data analysis. Other interesting features of the DMEx distribution can be reported as follows: Its distributive properties can be formulated in explicit forms. It can be applied to discuss asymmetric "positively skewed" count data. It can be utilized to discuss the dispersion of "over- and under-shaped" count data, and it can be used to analyze count data that have a monotonic unimodal or increased hazard rate shape.

2. The Structure of the DMEx Model

Using Equations (1) and (2), the SF of the DMEx distribution can be expressed as

$$S(x) = [-(x+1)\ln\beta + 1]\beta^{x+1};\ x \in \mathbb{N}_0, \tag{3}$$

where $0 < \beta = e^{-\lambda} < 1$ and $\mathbb{N}_0 = 0, 1, 2, 3, \dots$. The corresponding CDF and PMF to Equation (3) can be proposed as

$$F(x;\beta) = 1 - [-(x+1)\ln\beta + 1]\beta^{x+1};\ x \in \mathbb{N}_{0,} \tag{4}$$

and

$$\Pr(X = x;\beta) = [-x\ln\beta + \beta(x+1)\ln\beta - \beta + 1]\beta^{x};\ x \in \mathbb{N}_{0,} \tag{5}$$

respectively, where β controls the shape of the distribution. Figure 1 shows the PMF plots for various values of the parameter β.

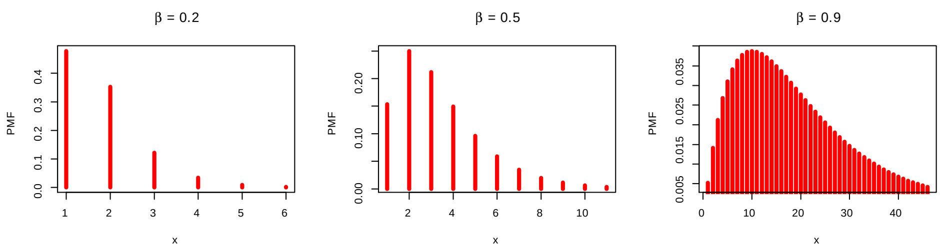

Figure 1. PMF visualization plots for the DMEx model.

As can be seen, the PMF is unimodally shaped, and it can be applied to discuss the positively skewed count data. The HRF can be proposed as

$$h(x;\beta) = \frac{-x\ln\beta + \beta(x+1)\ln\beta - \beta + 1}{-x\ln\beta + 1};\ x \in \mathbb{N}_0. \tag{6}$$

The reversed hazard rate function (RHRF) is given by

$$r(x;\beta) = \frac{\beta^x[-x\ln\beta + \beta(x+1)\ln\beta - \beta + 1]}{1 - [-(x+1)\ln\beta + 1]\beta^{x+1}};\ x \in \mathbb{N}_0. \tag{7}$$

Figure 2 shows the HRF and RHRF plots for different values of the distribution parameter β.

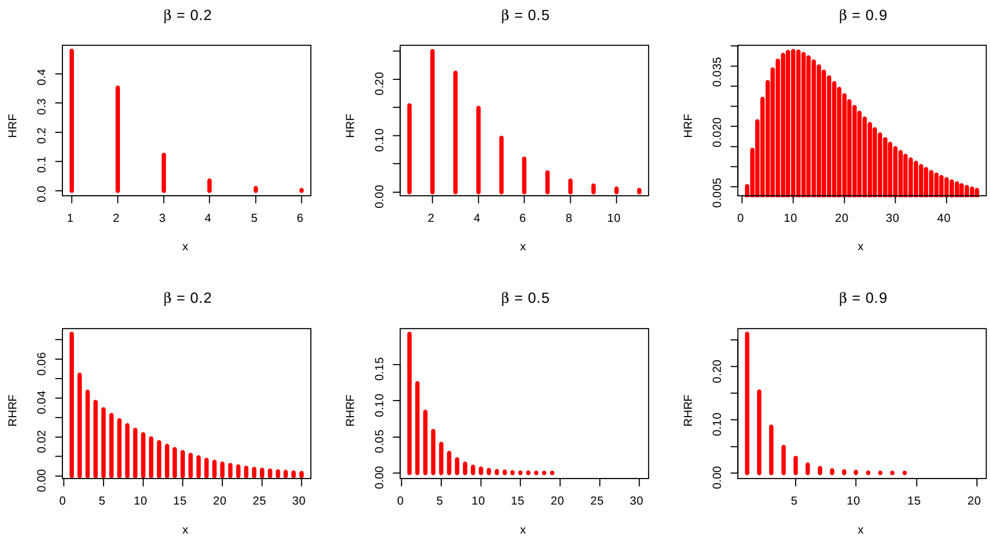

Figure 2. HRF and RHRF visualization plots for the DMEx model.

Both the HRF and its reversed function can be used effectively to model decreasing or unimodal failure modes.

2.1. Moments of a Statistical Distribution

Assume X is a DMEx RV, the probability generating function (PGF), say $\Pi_X(s)$, can be derived in a closed form as

$$\begin{aligned} \Pi_X(s) &= \sum_{x=0}^{\infty} s^x \Pr(X = x;\beta) \\ &= \sum_{x=0}^{\infty} [-x\ln\beta + \beta(x+1)\ln\beta - \beta + 1](s\beta)^x \\ &= \frac{-\beta(-1+s)\ln\beta + (-1+\beta)(\beta s - 1)}{(\beta s - 1)^2}, \end{aligned} \tag{8}$$

where the power series converges at least for all complex numbers s with $|s| \leq 1$. Equation (8) can be derived utilizing the Maple software program. Thus, the first four moments of the DMEx distribution can be expressed as

$$E(X) = \frac{-\beta(\beta - 1 + \ln\beta)}{(-1+\beta)^2}, \tag{9}$$

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

$$E\left(X^3\right) = \frac{-\beta\left[(7\beta^2+10\beta+1)\ln\beta + \beta^3 + 3\beta^2 - 3\beta - 1\right]}{(-1+\beta)^4}, \tag{11}$$

and

$$E\left(X^4\right) = \frac{\beta\left[(15\beta^3+55\beta^2+25\beta+1)\ln\beta + \beta^4 + 10\beta^3 - 10\beta - 1\right]}{(-1+\beta)^5}. \tag{12}$$

Using Equations (9)–(12), the $Var(X)$, $Sk(X)$, and $Ku(X)$ can be derived in closed forms where

$$Var(X) = E(X^2) - [E(X)]^2, \tag{13}$$

$$Sk(X) = \frac{E(X^3) - 3E(X^2)E(X) + 2[E(X)]^3}{[Var(X)]^{3/2}},$$

and

$$Ku(X) = \frac{E(X^4) - 4E(X)E(X^3) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4}{[Var(X)]^2}. \tag{14}$$

The moment generating function (MGF) can be expressed as

$$\begin{aligned}
\Pi_X^*(s) &= \sum_{x=0}^{\infty} e^{xs}\Pr(X = x; \beta) \\
&= \sum_{x=0}^{\infty} [-x\ln\beta + \beta(x+1)\ln\beta - \beta + 1](e^s\beta)^x \\
&= \frac{-\beta(-1+e^s)\ln\beta + (-1+\beta)(\beta e^s - 1)}{(\beta e^s - 1)^2}.
\end{aligned}$$

Table 1 lists some computational statistics (CS) of the DMEx model based on various values of the parameter β. All results given in Table 1 are reported in Figure 3.

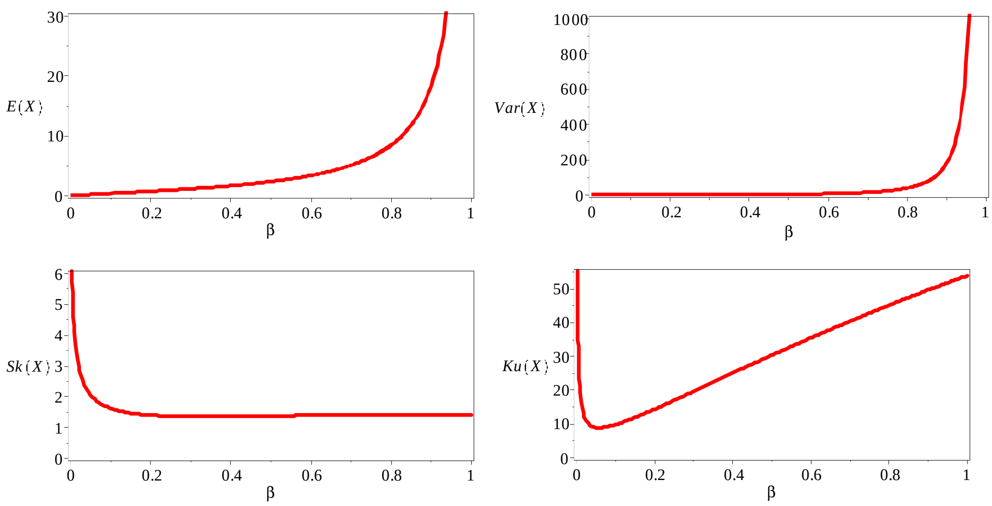

Figure 3. The plots of descriptive measures of the DMEx distribution.

Table 1. Some CS of the DMEx distribution.

$\beta \longrightarrow$ Measure $\downarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E(X)$	0.3954	0.7529	1.1657	1.6848	2.3863	3.4156	5.1075	8.4629	18.4824
$Var(X)$	0.3901	0.8139	1.4378	2.4502	4.2371	7.7429	15.802	40.249	180.2500
$Sk(X)$	1.6228	1.4038	1.3664	1.3697	1.3821	1.3943	1.4037	1.4099	1.4132
$Ku(X)$	9.8387	14.3813	19.6580	25.0539	30.3715	35.5222	40.4667	40.4667	49.7002

As can be seen, both the $E(X)$ and $Var(X)$ of the DMEx distribution increase when β grows to one. Further, the DMEx model is capable of modeling positively skewed count data under a leptokurtic shape.

2.2. Dispersion and Variation Measures

The index of dispersion, say $D(X)$, is related to the coefficient of variation, say $C(X)$. The $D(X)$ is also referred to as the coefficient of dispersion, which can be utilized to decide the possible over "$D(X) > 1$" or under "$D(X) < 1$" dispersion in the used data set. The $C(X)$ measure is generally applied for comparison with independent samples based on their variability. A higher $C(X)$ value indicates a higher variability. Let X be a DMEx RV. Then, the $D(X)$ and $C(X)$ can be formulated as

$$D(X) = \frac{-(3\beta+1)\ln(\beta) - \beta^2 + 1}{(-1+\beta)(\ln\beta + \beta - 1)]} - \frac{\beta(\ln\beta + \beta - 1)}{(-1+\beta)^2}, \tag{15}$$

and

$$C(X) = \sqrt{\frac{[(3\beta+1)\ln(\beta) + \beta^2 - 1](-1+\beta)}{\beta(\ln\beta + \beta - 1)^2} - 1}, \tag{16}$$

respectively, where $D(X) = \frac{Var(X)}{|E(X)|}$ and $C(X) = \frac{1}{|E(X)|}\sqrt{Var(X)}$. Table 2 shows some numerical calculations for $D(X)$ and $C(X)$ of the DMEx model based on the different values of the parameter β. All results given in Table 2 are reported in Figure 4.

Table 2. The $D(X)$ and $C(X)$ statistics of the DMEx distribution.

$\beta \longrightarrow$ Measure $\downarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$D(X)$	0.9866	1.0810	1.2335	1.4543	1.7756	2.2669	3.0939	4.7559	9.7525
$C(X)$	1.5797	1.1982	1.0287	0.9291	0.8626	0.8147	0.7783	0.7497	0.7264

As can be seen, for $\beta \to 1$, the $D(X)$ of the DMEx distribution increases whereas the $C(X)$ decreases. Moreover, the DMEx model is appropriate for modeling over- and under-dispersed count data.

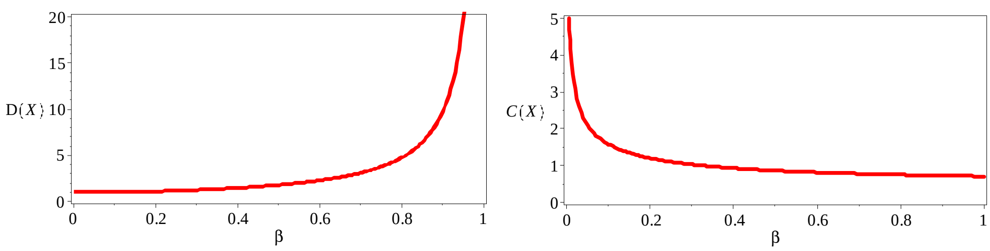

Figure 4. Plots of the $D(X)$ and $C(X)$ statistics of the DMEx distribution.

2.3. L-Moment Statistics

Suppose that n random variables $X_1, X_2, \ldots, X_n$ are ordered in non-decreasing magnitude and written as $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$. In the definition of order statistics (OS), there is no restriction on whether $X_i^{\prime}$s are independent or identically distributed (IID); however, many well-known results about OS are under the classical assumption that $X_i^{\prime}$s are IID. Let the RV X have the DMEx distribution. Then, the CDF of the ith OS can be expressed as

$$\begin{aligned} F_{i:n}(x;\beta) &= \sum_{k=i}^{n} \binom{n}{k} [F_i(x;\beta)]^k [1 - F_i(x;\beta)]^{n-k} \\ &= \sum_{k=i}^{n} \sum_{j=0}^{n-k} \Phi_m^{(n,k)} [F_i(x;\beta)]^{k+j}, \end{aligned} \tag{17}$$

where $\Phi_m^{(n,k)} = (-1)^j \binom{n}{k}\binom{n-k}{j}$. Furthermore, the corresponding PMF of the ith OS can be proposed as

$$f_{i:n}(x;\beta) = F_{i:n}(x;\beta) - F_{i:n}(x-1;\beta) = \sum_{k=i}^{n} \sum_{j=0}^{n-k} \Phi_m^{(n,k)} [f_i(x;\beta)]^{k+j}.$$

Thus, the rth moments of $X_{i:n}$, say $E(X_{i:n}^r)$, can be expressed as

$$E(X_{i:n}^r) = \sum_{x=0}^{\infty} \sum_{k=i}^{n} \sum_{j=0}^{n-k} \Psi_m^{(n,k)} x^r [f_i(x;\beta)]^{k+j}. \tag{18}$$

The L-moments (L-M), say ϑ_τ, are summary statistics for probability models. They are analogous to ordinary moments but are computed from linear functions of the ordered data values. The L-M of the RV X can be proposed as

$$\vartheta_\tau = \frac{1}{\tau} \sum_{i-0}^{\tau-1} (-1)^i \binom{\tau-1}{i} E(X_{\tau-i:\tau}). \tag{19}$$

Using Equation (19), some statistical measures based on the L-M statistics can be computed, such as $E(X) = \vartheta_1$, $C(X) = \frac{\vartheta_2}{\vartheta_1}$, $Sk(X) = \frac{\vartheta_3}{\vartheta_2}$, and $Ku(X) = \frac{\vartheta_4}{\vartheta_2}$.

3. Different Estimation Techniques

3.1. Maximum Likelihood Estimation (MLE)

In this subsection, estimation of the DMEx parameter is discussed using the method of maximum likelihood based on a complete sample. Let $X_1, X_2, \ldots, X_n$ be a random sample (RS) from a DMEx distribution. Then, the log-likelihood, say L, a function of the DMEx may be expressed as

$$L = \ln\beta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \ln[-x_i \ln\beta + \beta(x_i+1)\ln\beta - \beta + 1]. \tag{20}$$

Differentiating Equation (20) with respect to the parameter β and setting the result equal to 0, we obtain

$$\frac{\partial L}{\partial \beta} = \frac{1}{\beta} \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \frac{-\frac{x_i}{\beta} + (x_i+1)(\ln\beta + 1) - 1}{-x_i \ln\beta + \beta(x_i+1)\ln\beta - \beta + 1}. \tag{21}$$

Since Equation (21) cannot be derived in the closed form as a function of only the data "$x_i; i = 1, 2, 3, \ldots, n$", a numerical iterative procedure is required to solve it numerically "Newton-Raphson as an example".

3.2. Moment Estimation (ME)

Using the moment approach to estimate β, we must first equate the population moment to the corresponding sample moment and then solve the non-linear equation

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{-\beta(\ln\beta+\beta-1)}{(-1+\beta)^2}, \tag{22}$$

with respect to β.

3.3. Proportion Estimation (PE)

Let $X_1, X_2, \ldots, X_n$ be an RS from the DMEx distribution. Since we have one unknown parameter, one indicator function is defined as follows

$$T(x_i) = \begin{cases} 1; & \text{if } x_i = 0 \\ 0; & \text{if } x_i \neq 0. \end{cases} \tag{23}$$

Assume $Z = \sum_{i=1}^{n} T(x_i)$ denotes the number of zeroes in the sample. Using Equations (4) and (23), we obtain $P(X \leq 0) = \frac{Z}{n}$. Hence, we obtain the estimation of the parameter β by solving the following equation

$$\widehat{\beta}\ln\widehat{\beta} - \widehat{\beta} + 1 - \frac{Z}{n} = 0. \tag{24}$$

Since $\frac{Z}{n}$ is an unbiased and consistent empirical estimator of probability $P(X \leq 0)$, the $\widehat{\beta}$ is also an unbiased and consistent estimator of β.

4. Comparing Different Estimators (CDEs): A Simulation Study

In this segment, we assess the performance of the MLE, ME, and PE with respect to the sample size n utilizing R software. For CDEs, MCMC simulations are performed according to different schemes. The assessment is according to a simulation study:

1. Generate N = 10,000 samples of various sizes "$n_i; i = 1,2,3,4$" from the DMEx model as follows
   - Scheme I: $\beta = 0.15 \mid n_1 = 25,\ n_2 = 50,\ n_3 = 100,\ n_4 = 250,\ n_5 = 400,\ n_6 = 600.$
   - Scheme II: $\beta = 0.35 \mid n_1 = 25,\ n_2 = 50,\ n_3 = 100,\ n_4 = 250,\ n_5 = 400,\ n_6 = 600.$
   - Scheme III: $\beta = 0.85 \mid n_1 = 25,\ n_2 = 50,\ n_3 = 100,\ n_4 = 250,\ n_5 = 400,\ n_6 = 600.$
2. Compute the MLE, ME, and PE for the 10,000 samples, say $\widehat{\beta}_k$ for $k = 1,2,\ldots,10{,}000$.
3. Calculate the bias "BS", mean squared errors (MSE), and mean relative errors (MRE) for N = 10,000 samples as

$$|\mathrm{BS}(\beta)| = \frac{1}{N}\sum_{k=1}^{N}\left|\widehat{\beta}_k - \beta_k\right|,\ \mathrm{MSE}(\beta) = \frac{1}{N}\sum_{k=1}^{N}(\widehat{\beta}_k - \beta_k)^2,\ \mathrm{MRE}(\beta) = \frac{1}{N}\sum_{k=1}^{N}\frac{\left|\widehat{\beta}_k - \beta_k\right|}{\beta_k}.$$

4. The empirical results of the simulations are reported in Tables 3–5 and provided via Figures 5–7.

As can be seen, the BS of the parameter β approaches 0 when the sample size n grows. Similarly, both the MSE and MRE of the DMEx parameter approach 0 when the sample size n increases. These results reveal the consistency property of the derived estimators. Thus, we can conclude that all estimation methods work quite well under various sizes of samples.

Table 3. Simulation results for β = 0.15.

n	Criteria	MLE	ME	PE
25	|BS|	0.71072536	0.68736626	1.01333569
	MSE	0.50610243	0.48899665	0.62541235
	MRE	0.47496366	0.46614520	0.67532023
50	|BS|	0.39401774	0.38114550	0.59444182
	MSE	0.10203639	0.09469954	0.17296732
	MRE	0.26463284	0.25413209	0.39819026
100	|BS|	0.27723665	0.25923302	0.39711141
	MSE	0.07412203	0.06774263	0.15838302
	MRE	0.18212052	0.17334963	0.26733503
250	|BS|	0.18533699	0.18241249	0.29293434
	MSE	0.03441458	0.03211288	0.08771857
	MRE	0.12229566	0.12014121	0.1964149
400	|BS|	0.13130521	0.11796928	0.19905126
	MSE	0.01796344	0.01441778	0.03983039
	MRE	0.08722563	0.07803269	0.1327154
600	|BS|	0.04210295	0.03471560	0.09222016
	MSE	0.00429679	0.00413072	0.01223098
	MRE	0.00831602	0.00726341	0.02431204

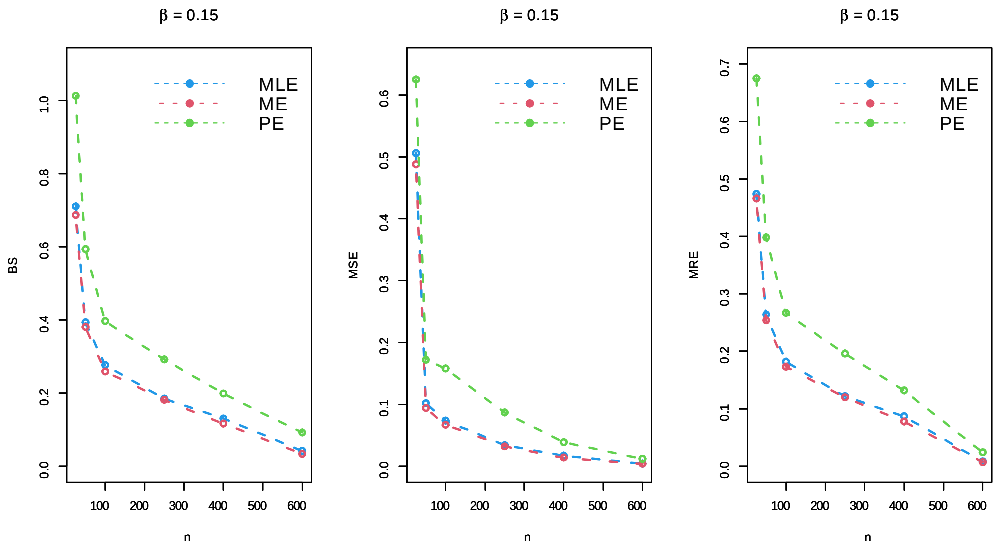

Figure 5. Simulation visualization plots for $\beta = 0.15$.

Table 4. Simulation results for β = 0.35.

n	Criteria	MLE	ME	PE
25	|BS|	0.46033471	0.45230219	0.49111028
	MSE	0.21236964	0.20903288	0.24800326
	MRE	0.93014126	0.91314177	0.99711412
50	|BS|	0.31830325	0.31130925	0.41790345
	MSE	0.10344125	0.09841329	0.17471516
	MRE	0.63899659	0.62332954	0.82130958
100	|BS|	0.23141239	0.22213412	0.31815142
	MSE	0.05210236	0.04930287	0.10274589
	MRE	0.46266369	0.44310965	0.63937195

Table 4. *Cont.*

n	Criteria	MLE	ME	PE
250	|BS|	0.16141257	0.15810236	0.24730864
	MSE	0.02695256	0.02596985	0.05886243
	MRE	0.32242856	0.31519732	0.48810236
400	|BS|	0.11463142	0.09866367	0.17199896
	MSE	0.01386537	0.00980015	0.02914120
	MRE	0.22720103	0.19710414	0.33409875
600	|BS|	0.08795636	0.07296985	0.12110286
	MSE	0.00877157	0.00627420	0.00923698
	MRE	0.11209537	0.09830987	0.21008025

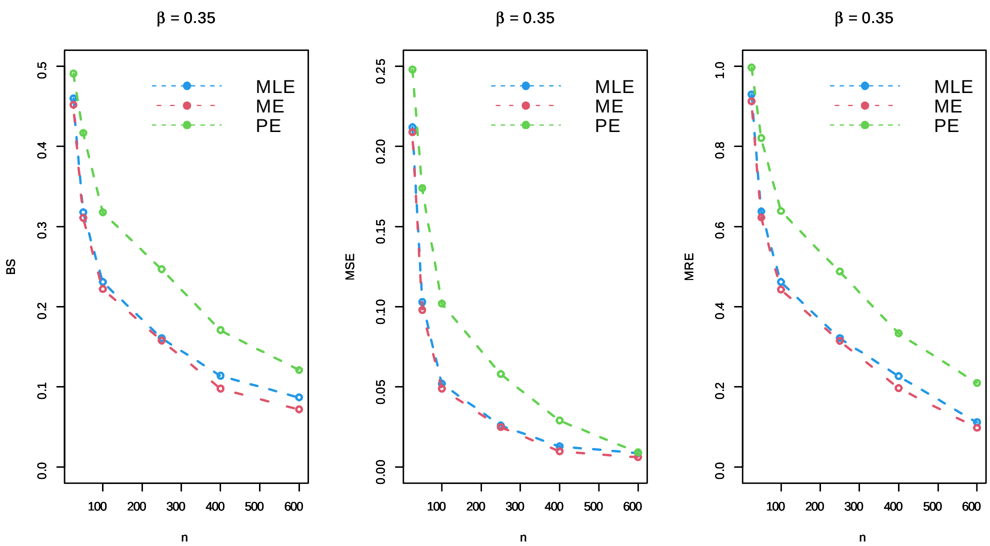

Figure 6. Simulation visualization plots for $\beta = 0.35$.

Table 5. Simulation results for $\beta = 0.85$.

n	Criteria	MLE	ME	PE
25	|BS|	0.46296336	0.48641129	0.49813368
	MSE	0.21374125	0.23632653	0.24841515
	MRE	0.92396326	0.97291764	0.99433695
50	|BS|	0.31103258	0.37031526	0.42356964
	MSE	0.09774623	0.13274859	0.17233696
	MRE	0.62174120	0.74533626	0.81241852
100	|BS|	0.23163949	0.26933026	0.31233635
	MSE	0.05208856	0.07341259	0.09810221
	MRE	0.46241203	0.53233636	0.61233982
250	|BS|	0.16196336	0.19810775	0.22185236
	MSE	0.02641205	0.03901486	0.04941212
	MRE	0.32163955	0.39910213	0.44330352
400	|BS|	0.11300125	0.13433636	0.16241252
	MSE	0.01322585	0.01841021	0.02744125
	MRE	0.23474694	0.27117655	0.33139625
600	|BS|	0.03332357	0.09811494	0.12141254
	MSE	0.00894112	0.00903661	0.01230225
	MRE	0.10541453	0.12322396	0.18338552

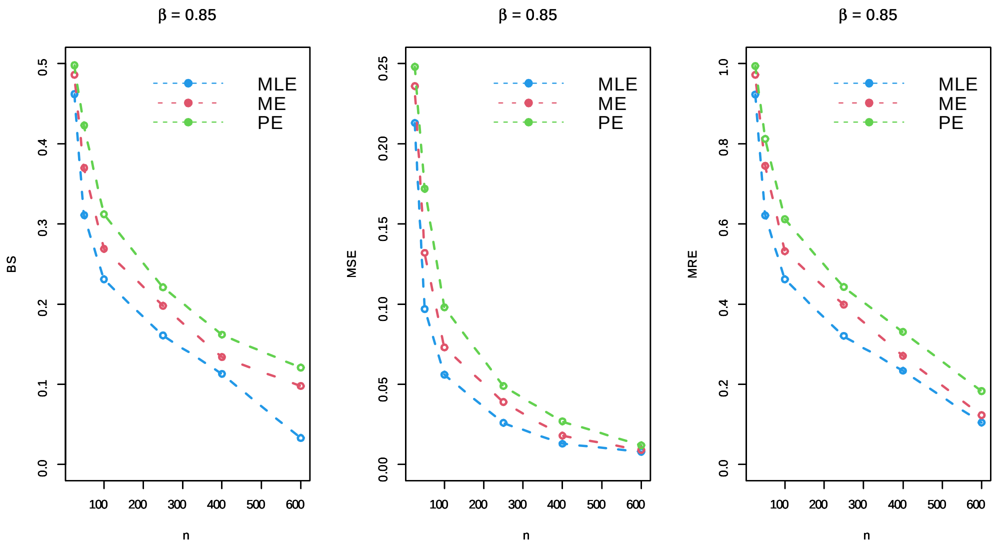

Figure 7. Simulation visualization plots for $\beta = 0.85$.

5. Data Modeling: Competitive Models and Statistical Criteria

In this segment, the importance of the DMEx distribution is discussed utilizing data sets from various areas. We shall compare the fits of the DMEx model with diverse competitive distributions, such as the Poisson (Poi), the discrete Pareto (DPa), the discrete Burr-XII (DBXII), the discrete Rayleigh (DR), the discrete Burr–Hatke (DBH), and the discrete inverse Rayleigh (DIR) models. The tested distributions are compared to some criteria, such as $-L$, the Hannan–Quinn information criterion (HQIC), the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the corrected AIC (CAIC), and the Kolmogorov–Smirnov (KS) test with its p-value.

5.1. Data Set I: Electronic Components

The first set of data represents the failure times for a sample of 15 electronic components (EC) in an accelerated life test (see, Johnston [17]). The observed descriptive statistics (ODS) include: mean = 27.533, variance = 631.980, skewness = 1.534, and kurtosis = 6.924. Non-parametric plots for the EC data are sketched in Figure 8.

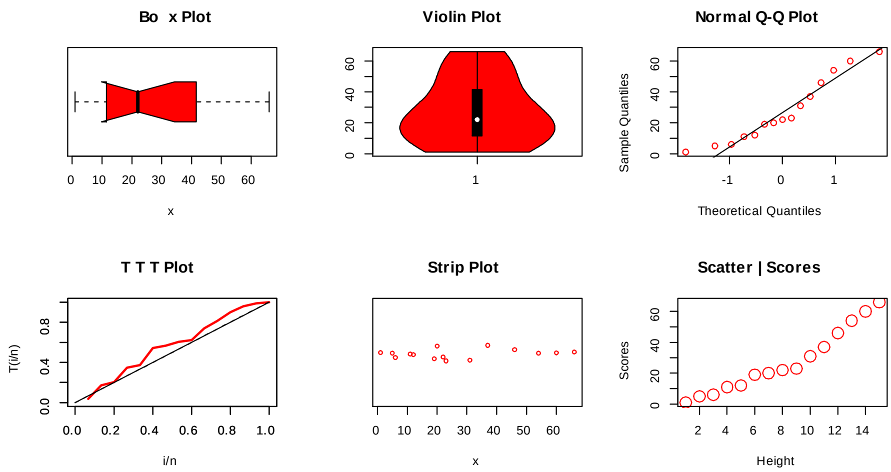

Figure 8. Non-parametric visualization plots for EC data.

The MLE with their SE, the CI for the parameter(s), and the goodness of fit (GOF) test for these data are reported in Tables 6 and 7.

Table 6. The MLE, SE, and CI for EC data.

Model	β			α		
	MLE	**SE**	**CI**	**MLE**	**SE**	**CI**
DMEx	0.931	0.012	$[0.910, 0.956]$	–	–	–
DR	0.999	2.6×10^{-4}	$[0.998, 0.999]$	–	–	–
DIR	1.8×10^{-7}	0.055	$[0, 0.107]$	–	–	–
DBH	0.999	0.008	$[0.984, 1.014]$	–	–	–
DPa	0.720	0.061	$[0.600, 0.839]$	–	–	–
Poi	27.533	1.355	$[24.878, 30.189]$	–	–	–
DINH	0.578	0.193	$[0.199, 0.957]$	29.072	20.384	$[0, 69.024]$
DB-XII	0.975	0.051	$[0.874, 1]$	13.367	27.785	$[0, 67.824]$

Table 7. The GOF test for EC data.

Statistic	DMEx	DR	DIR	DBH	DPa	Poi	DINH	DB-XII
$-L$	64.7898	66.394	89.096	91.368	77.402	151.206	67.879	75.724
AIC	131.5796	134.788	180.192	184.737	156.805	304.413	139.758	155.448
CAIC	131.8873	135.096	180.499	185.045	157.112	304.721	140.758	156.448
BIC	132.2877	135.496	180.899	185.445	157.513	305.121	141.174	156.864
HQIC	131.5721	134.781	180.184	184.729	156.797	304.405	139.743	155.433
KS	0.1144	0.216	0.698	0.791	0.405	0.381	0.207	0.388
p-value	0.9766	0.433	<0.0001	<0.0001	0.009	0.025	0.481	0.015

As can be noted, based on a significance level of 0.05, both the DMEx and DR models work quite well for modeling the EC data, but the DMEx distribution is the best. Figures 9 and 10 show the empirical CDFs as well as the probability-probability (P-P) plots, or "parametric plots", for the EC data, which prove the empirical results mentioned in Table 7.

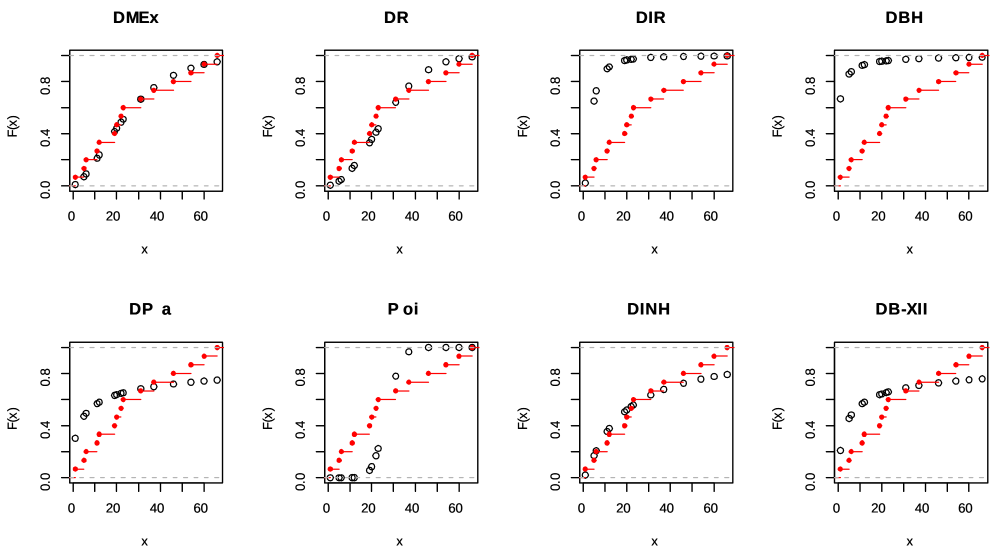

Figure 9. Empirical CDF visualization plots for EC data.

Table 8 lists various estimators for the EC data, and it is noted that the MLE and ME techniques work quite well for modeling these data.

Table 8. Different estimators for EC data.

Technique	β	KS	p-Value
ME	0.931	0.114	0.989
PE	0.811	0.588	0.627×10^{-4}

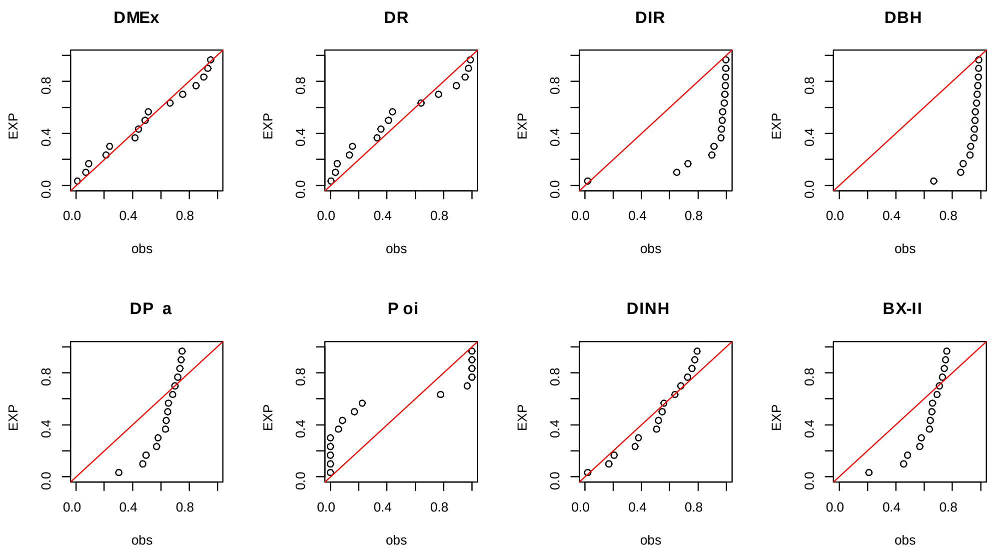

Figure 10. P-P visualization plots for EC data.

Table 9 reports some numerical accounts of the empirical descriptive statistics (EDS).

Table 9. The EDS for EC data.

Approach	$E(X)$	$Var(X)$	$D(X)$	$Sk(X)$	$Ku(X)$
MLE	27.536	393.082	14.275	1.414	5.999
PE	9.0508	45.6912	5.048	1.410	5.989
ME	27.533	393.017	14.274	1.414	5.999

Both theoretical and empirical "MLE and ME" scales are approximately equal. Thus, the ME approach works quite well beside the MLE method for estimating the unknown parameter. The performance of the PE approach is inferior compared to other techniques for some measures. The EC data are suffering from over-dispersed phenomena. Moreover, the EC data are skewed to the right and leptokurtic.

5.2. Data Set II: Leukemia Remission

These data involve leukemia remission (LR) times (in weeks) for 20 patients (see Damien and Walker [18]) according to the discretization concept. For these data, the ODS equals 19.55, 216.05, 0.637, and −0.739. Non-parametric plots for the LR data are displayed in Figure 11.

The MLE with their SE, the CI for the parameter(s), and the GOF test for the LR data are listed in Tables 10 and 11.

Table 10. The MLE, SE, and CI for LR data.

Model	β			α		
	MLE	SE	CI	MLE	SE	CI
DMEx	0.905	0.014	$[0.877, 0.933]$	–	–	–
DR	0.998	0.0004	$[0.998, 0.999]$	–	–	–
DIR	7.82×10^{-7}	–	–	–	–	–
DBH	0.998	0.009	$[0.981, 1.017]$	–	–	–
DPa	0.696	0.056	$[0.585, 0.806]$	–	–	–
Poi	19.550	0.989	$[17.612, 21.493]$	–	–	–
DINH	0.737	0.268	$[0.212, 1.262]$	14.798	9.997	$[0, 34.392]$
DB-XII	0.998	0.004	$[0.99, 1]$	182.367	94.801	$[0, 277.7.001]$

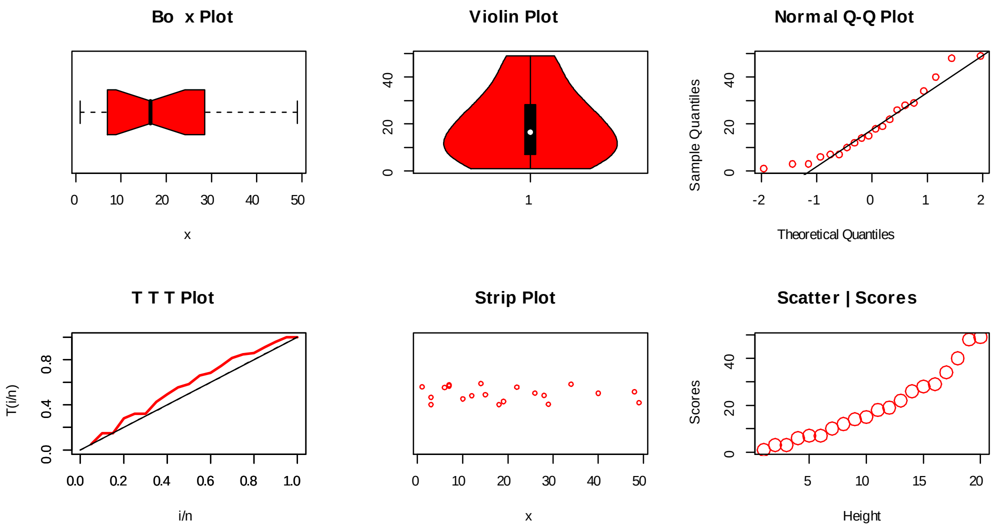

Figure 11. Non-parametric visualization plots for LR data.

Table 11. The GOF test for LR data.

Statistic	DMEx	DR	DIR	DBH	DPa	Poi	DINH	DB-XII
$-L$	79.071	81.175	101.987	110.283	95.448	152.718	82.818	92.602
AIC	160.141	164.351	205.975	222.565	192.896	307.436	169.635	189.203
CAIC	160.364	164.572	206.197	222.787	193.118	307.658	170.341	189.909
BIC	161.137	165.346	206.973	223.561	193.892	308.432	171.627	191.195
HQIC	160.336	164.544	206.169	222.759	193.090	307.630	170.024	189.592
KS	0.109	0.199	–	0.751	0.392	0.352	0.189	0.369
p-value	0.970	0.401	–	<0.001	0.004	0.014	0.467	0.008

Based on a significance level of 0.05, both the DMEx and DR distributions work quite well for analyzing the LR data, but the DMEx model is the best. Figures 12 and 13 show the empirical CDFs as well as the P-P plots for the LR data, which prove the empirical results reported in Table 11.

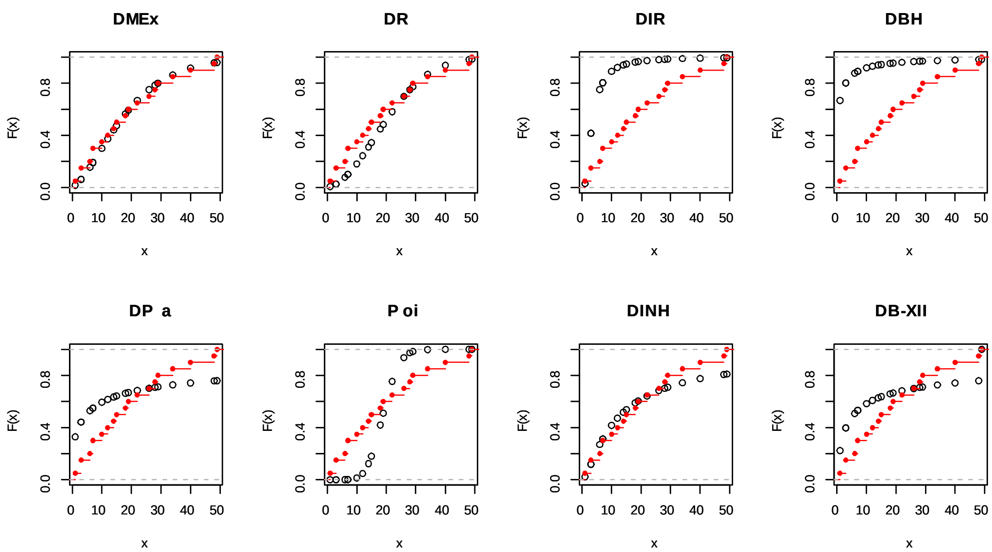

Figure 12. Empirical CDF visualization plots for LR data.

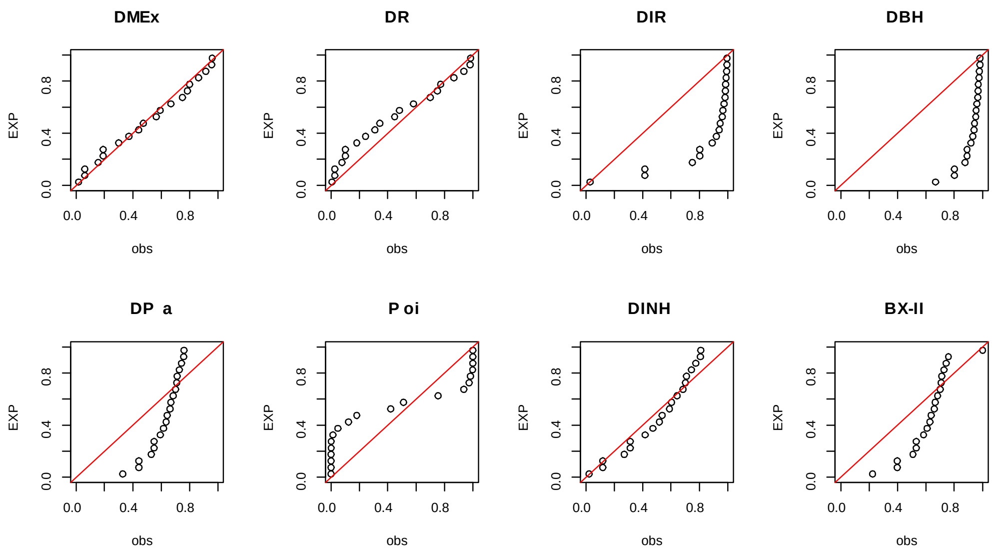

Figure 13. P-P visualization plots for LR data.

Table 12 reports different estimators for the LR data, and it is noted that the MLE and ME methods work quite well for modeling these data.

Table 12. Different estimators for LR data.

Technique	β	KS	p-Value
ME	0.905	0.100	0.841
PE	0.837	0.350	0.0147

Table 13 lists some numerical accounts of the EDS for the LR data.

Table 13. The EDS for LR data.

Approach	$E(X)$	$Var(X)$	$D(X)$	$Sk(X)$	$Ku(X)$
MLE	19.551	201.109	10.286	1.413	5.998
PE	10.756	63.433	5.897	1.411	5.992
ME	19.550	201.084	10.286	1.413	5.998

The ME approach works quite well beside the MLE technique for estimating the unknown parameter. The LR data are suffering from over-dispersed phenomena. Further, the LR data are skewed to the right and leptokurtic.

5.3. Data Set III: Coronavirus in Punjab

The third set of data represents the number of deaths due to coronavirus in Punjab (COV-P) during the period from 24 March 2020 to 30 April 2020. The data are as follows: 1, 2, 3, 5, 5, 6, 9, 9, 11, 11, 11, 12, 15, 15, 16, 17, 18, 19, 21, 23, 24, 28, 34, 36, 37, 41, 42, 45, 51, 58, 65, 73, 81, 83, 91, 100, 103, 106. Non-parametric plots for the COV-P data are listed in Figure 14.

The MLE with their SE, the CI for the parameter(s), and the GOF test for the COV-P data are proposed in Tables 14 and 15.

Based on a significance level of 0.05, the DMEx distribution is the best among all tested models. Figures 15 and 16 show the empirical CDFs as well as the P-P plots for the COV-P data, which prove the empirical results listed in Table 15.

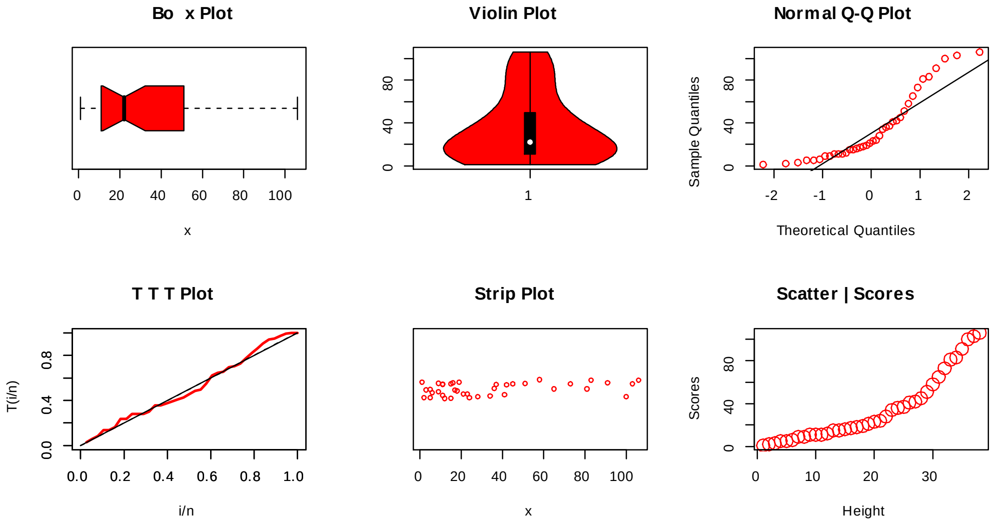

Figure 14. Non-parametric visualization plots for COV-P data.

Table 14. The MLE, SE, and CI for COV-P data.

Model	β			α		
	MLE	**SE**	**CI**	**MLE**	**SE**	**CI**
DMEx	0.9451	0.006	$[0.93, 0.957]$	–	–	–
DR	0.9996	0.00008	$[0.9994, 0.999]$	–	–	–
DIR	1.634×10^{-10}	–	–	–	–	–
DBH	0.999	0.004	$[0.993, 1.006]$	–	–	–
DPa	0.729	0.037	$[0.658, 0.803]$	–	–	–
Poi	34.921	0.959	$[33.04, 36.8]$	–	–	–
DINH	0.615	0.144	$[0.333, 0.896]$	29.319	13.995	$[1.889, 56.748]$
DB-XII	0.996	0.004	$[0.989, 1.003]$	79.588	82.339	$[0, 215.023]$

Table 15. The GOF test for COV-P data.

Statistic	DMEx	DR	DIR	DBH	DPa	Poi	DINH	DB-XII
$-L$	176.621	186.7	226.355	241.306	202.578	594.751	177.779	198.727
AIC	355.242	375.4	454.709	486.612	407.155	1191.5	359.558	401.454
CAIC	355.353	375.511	454.82	486.955	407.267	1191.61	359.901	401.797
BIC	356.879	377.038	456.347	489.888	408.793	1193.14	362.833	404.729
HQIC	355.824	375.983	455.292	487.778	407.738	1192.09	360.723	402.619
KS	0.1620	0.309	0.644	0.779	0.379	0.519	0.171	0.367
p-value	0.271	0.001	<0.001	<0.001	<0.001	<0.001	0.245	<0.001

Table 16 reports various estimators for the COV-P data, and it is noted that the MLE and ME techniques work quite well for discussing these data.

Table 16. Different estimators for COV-P data.

Technique	β	KS	p-Value
ME	0.945	0.162	0.271
PE	0.883	0.353	0.0002

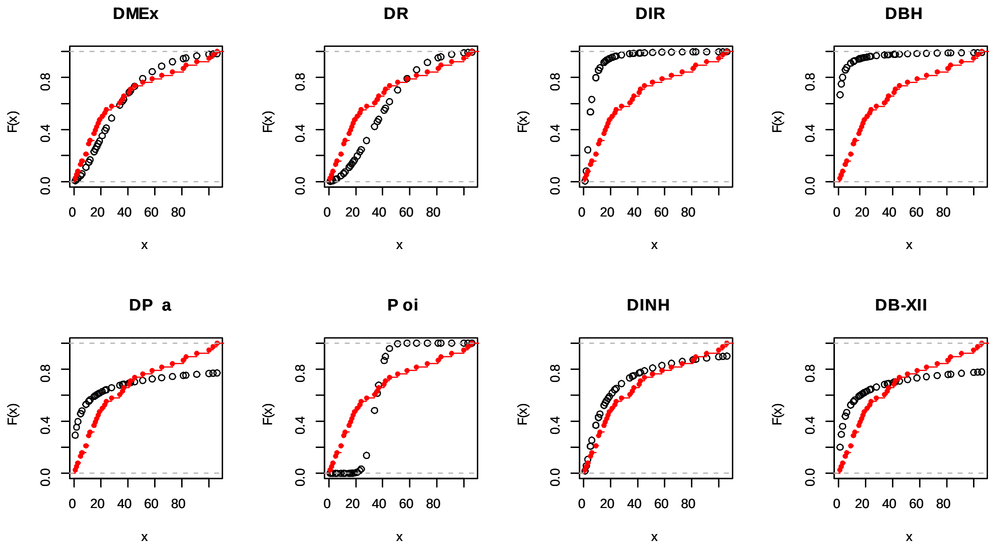

Figure 15. Empirical CDF visualization plots for COV-P data.

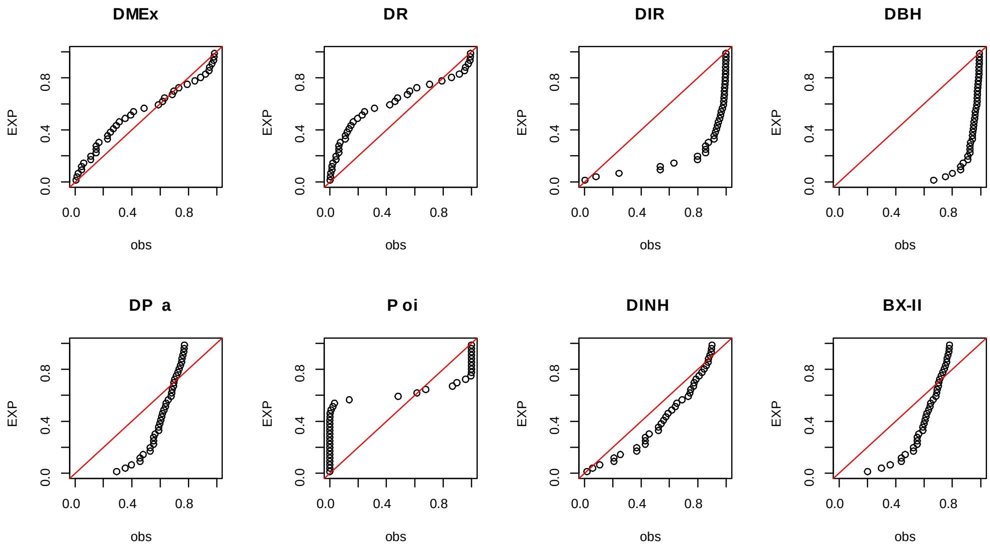

Figure 16. P-P visualization plots for COV-P data.

Table 17 introduces some numerical accounts of the EDS for the COV-P data.

Table 17. The EDS for COV-P data.

Approach	$E(X)$	$Var(X)$	$D(X)$	$Sk(X)$	$Ku(X)$
MLE	34.923	627.491	17.968	1.414	5.999
PE	15.561	129.054	8.2937	1.413	5.996
ME	34.921	627.409	17.967	1.414	5.999

The ME approach works quite well beside the MLE technique for estimating the unknown parameter. The COV-P data are suffering from over-dispersed phenomena. Moreover, the COV-P data are skewed to the right and leptokurtic. The profiles of the L functions for data sets I, II, and III are displayed in Figure 17, and it is noted that the estimator is a unique "unimodal function". Figure 18 shows the empirical CDFs for the different estimation approaches.

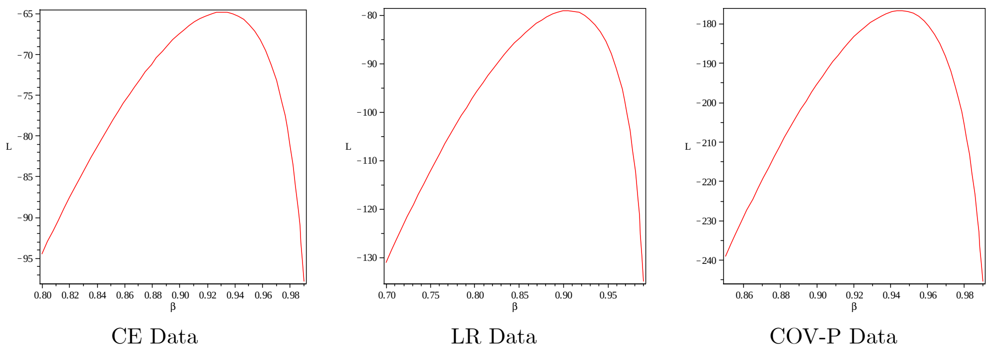

Figure 17. The profiles plots of L for data sets.

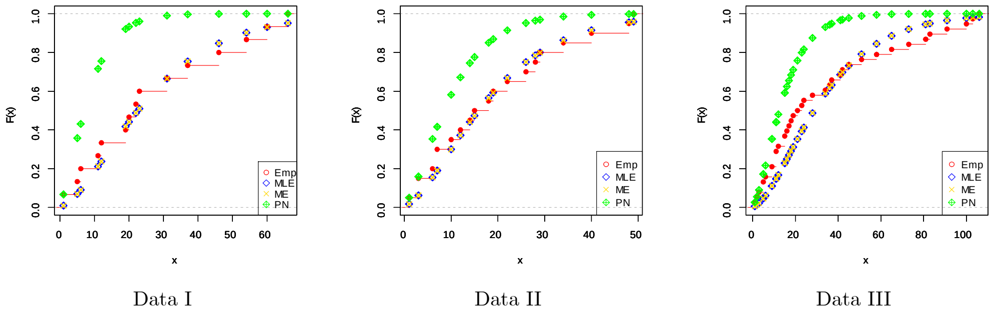

Figure 18. Empirical CDF visualization plots for different estimation methods.

6. Concluding Remarks and Future Work

This article focuses on a one-parameter discrete distribution created based on the survival discretization approach and called the DMEx distribution. The statistical properties of the DMEx model have been derived and expressed in closed forms. It was found that the DMEx model is proper for modeling right-skewed data sets of a leptokurtic shape. The presented discrete distribution can be used as a statistical tool to model different types of HRFs, including those that are decreasing and unimodally shaped. The DMEx parameter has been estimated utilizing various estimation approaches, i.e., MLE, ME, and PE. Simulation studies have been performed based on various sample sizes, and it was found that the MLE and ME techniques work quite effectively for estimating the DMEx parameter. Finally, three real data sets have been analyzed and discussed to illustrate the notability of the DMEx distribution, and it was found that DMEx outperforms all other competitive distributions in all aspects of the current study. In the future, the bivariate extension of the DMEx models will be proposed and discussed in detail. Furthermore, a regression model and INAR(1) process will be discussed alongside their applications.

Author Contributions: Conceptualization, M.S.E; data curation, M.E.-M. and H.S.S.; formal analysis, M.S.E. and B.A.; investigation, M.E.-M. and H.S.S.; methodology, M.E.-M. and M.S.E.; resources, H.S.S. and B.A.; software, M.E.-M., H.S.S., and M.S.E.; validation, M.S.E. and B.A.; writing— review and editing, M.S.E. All authors have read and agreed to the published version of the manuscript.

Funding: Project number (IF-PSAU-2022/01/22553).

Data Availability Statement: All data sets have been listed within the paper.

Acknowledgments: The authors extend their appreciation to the Deputyship for Research & Innovation of the Ministry of Education in Saudi Arabia for funding this research work through project number (IF-PSAU-2022/01/22553).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bhatti, F.A.; Hamedani, G.G.; Korkmaz, M.Ç.; Sheng, W.; Ali, A. On the Burr XII-moment exponential distribution. *PLoS ONE* **2021**, *16*, e0246935. [CrossRef] [PubMed]
2. Iqbal, Z.; Hasnain, S.A.; Salman, M.; Ahmad, M.; Hamedani, G. Generalized exponentiated moment exponential distribution. *Pak. J. Stat.* **2014**, *30*, 537–554.
3. Ahsan-ul-Haq, M. On Poisson moment exponential distribution with applications. *Ann. Data Sci.* **2022**, 1–22.. [CrossRef]
4. Abbas, S.; Jahngeer, A.; Shahbaz, S.H.; Afify, A.Z. Topp-Leone moment exponential distribution: Properties and applications. 2020. [CrossRef]
5. Akhter, Z.; MirMostafaee, S.M.T.K.; Ormoz, E. On the order statistics of exponentiated moment exponential distribution and associated inference. *J. Stat. Comput. Simul.* **2022**, *92*, 1322–1346. [CrossRef]
6. Hashmi, S.; Haq, M.A.U.; Usman, R.M.; Ozel, G. The Weibull-moment exponential distribution: Properties, characterizations and applications. *J. Reliab. Stat. Stud.* **2019**, *12*, 1–22.
7. Kumar, D.; Dey, T.; Dey, S. Statistical inference of exponentiated moment exponential distribution based on lower record values. *Commun. Math. Stat.* **2017**, *5*, 231–260. [CrossRef]
8. Iriarte, Y.A.; Astorga, J.M.; Venegas, O.; Gómez, H.W. Slashed moment exponential distribution. *J. Stat. Appl.* **2017**, *16*, 354–365. [CrossRef]
9. 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]
10. 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]
11. Yousof, H.M.; Chesneau, C.; Hamedani, G.; Ibrahim, M. A new discrete distribution: Properties, characterizations, modeling real count data, Bayesian and non-Bayesian estimations. *Statistica* **2021**, *81*, 135–162. [CrossRef]
12. Altun, E.; El-Morshedy, M.; Eliwa, M.S. A study on discrete Bilal distribution with properties and applications on integervalued autoregressive process. *Revstat-Stat. J.* **2022**, *20*, 501-528.
13. El-Morshedy, M.; Altun, E.; Eliwa, M.S. A new statistical approach to model the counts of novel coronavirus cases. *Math. Sci.* **2022**, *16*, 37–50. [CrossRef] [PubMed]
14. Almetwally, E.M.; Ibrahim, G.M. Discrete alpha power inverse Lomax distribution with application of COVID-19 data. *J. Appl. Math.* **2020**, *9*, 11–22.
15. Tyagi, A.; Choudhary, N.; Singh, B. A new discrete distribution: Theory and applications to discrete failure lifetime and count data. *J. Appl. Probab. Stat.* **2020**, *15*, 117–143.
16. Bakouch, H.S.; Jazi, M.A.; Nadarajah, S. A new discrete distribution. *Statistics* **2014**, *48*, 200–240. [CrossRef]
17. Johnston, G. *Statistical Models and Methods for Lifetime Data*; Wiley: New York, NY, USA, 2003.
18. Damien, P.; Walker, S. A Bayesian non-parametric comparison of two treatments. *Scand. J. Stat.* **2002**, *29*, 51–56. [CrossRef]