mathematics

MDPI

Article

A Discrete Exponential Generalized-G Family of Distributions: Properties with Bayesian and Non-Bayesian Estimators to Model Medical, Engineering and Agriculture Data

Mohamed S. Eliwa [1,2,4,*], Mahmoud El-Morshedy [3,4] and Haitham M. Yousof [5]

[1] Department of Statistics and Operation Research, College of Science, Qassim University, Buraydah 51482, Saudi Arabia
[2] Section of Mathematics, International Telematic University Uninettuno, I-00186 Rome, Italy
[3] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[4] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
[5] Department of Statistics, Mathematics and Insurance, Benha University, Benha 13518, Egypt
* Correspondence: m.eliwa@qu.edu.sa

Abstract: This paper introduces a new flexible probability tool for modeling extreme and zero-inflated count data under different shapes of hazard rates. Many relevant mathematical and statistical properties are derived and analyzed. The new tool can be used to discuss several kinds of data, such as "asymmetric and left skewed", "asymmetric and right skewed", "symmetric", "symmetric and bimodal", "uniformed", and "right skewed with a heavy tail", among other useful shapes. The failure rate of the new class can vary and can take the forms of "increasing-constant", "constant", "monotonically dropping", "bathtub", "monotonically increasing", or "J-shaped". Eight classical estimation techniques—including Cramér–von Mises, ordinary least squares, L-moments, maximum likelihood, Kolmogorov, bootstrapping, and weighted least squares—are considered, described, and applied. Additionally, Bayesian estimation under the squared error loss function is also derived and discussed. Comprehensive comparison between approaches is performed for both simulated and real-life data. Finally, four real datasets are analyzed to prove the flexibility, applicability, and notability of the new class.

Keywords: survival discretization; Gibbs sampler; Metropolis–Hastings technique; L-moment structure; bootstrapping approach; Kolmogorov method; Bayesian analysis; Markov chain Monte Carlo; extreme and zero-inflated count data

MSC: 62E99; 62E15

Citation: 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. https://doi.org/10.3390/math10183348

Academic Editors: Stelios Psarakis and David Carfi

Received: 2 August 2022
Accepted: 9 September 2022
Published: 15 September 2022

1. Introduction

The discretization of current continuous models has recently drawn significant interest. This is due to the fact that the data must often be recorded on a discrete scale rather than a continuous analog. Examples include the number of daily COVID-19 deaths, or the number of renal cysts caused by steroid use, among others. In order to discuss such data, new discrete models have been proposed and investigated in the statistical and mathematical literature. For instance, the discrete Weibull (DW) (Nakagawa and Osaki [1]), discrete Rayleigh (DR) (Roy [2]), discrete inverse Weibull (DIW) (Jazi et al. [3]), discrete exponential (DE) (Gomez-Déniz [4]), discrete inverse Rayleigh (DR) (Hesterberg [5]), discrete exponentiated Weibull (EDW) (Nekoukhou and Bidram [6]), discrete Lindley-II (DLy-II) (Hussain et al. [7]), discrete Lomax (DLx) and discrete Burr XII (DBXII) (Para and Jan [8]), discrete generalized Burr–Hatke distribution (for more details, see El-Morshedy et al. [9]), discrete exponentiated Lindley (EDLy) (see El-Morshedy et al. [10]), discrete generalized

Burr–Hatke distribution (see Yousof et al. [11]), and the discrete inverse Burr (DIB) distribution (see Chesneau et al. [12]), among others.

On the other hand, another statistical approach has recently been followed up to define new discrete G families of probability distributions. The statistical approach depends on generating new discrete G families of probability distributions based on some existing continuous families. For example, following Bourguignon et al. [13], Aboraya et al. [14] defined and studied the discrete Rayleigh G (DR-G) family of distributions, while Ibrahim et al. [15] proposed a discrete analog of the Weibull G family. Following Steutel and van Harn [16], Eliwa et al. [17] introduced and discussed the discrete Gompertz G (DGz-G) family of distributions. Recently, Yousof et al. [18] presented a new G family of continuous distributions called the exponential generalized G (EG-G) family. The CDF of the EG-G family can be expressed as follows:

$$F_{\alpha,\gamma}(z) = 1 - exp[-O(z;\gamma,\boldsymbol{\xi})] \mid_{(z\in R \text{ and } \alpha,\gamma>0)}, \quad (1)$$

where

$$O(z;\gamma,\boldsymbol{\xi}) = \dot{G}_{\boldsymbol{\xi}}^{-\gamma}(z) - 1|_{z\in R,\gamma>0},$$

is the generalized odds ratio argument, $\dot{G}_{\boldsymbol{\xi}}(z) = 1 - G_{\boldsymbol{\xi}}(z)$ refers to the survival function (SF) of any baseline model with the parameter vector $\boldsymbol{\xi}$, α is a scale parameter, and γ is an additional shape parameter. In this work and following Yousof et al. [18] and Steutel and van Harn [16], we define and study a new discrete analog of the EG-G family called *the discrete exponential generalized G* (DEG-G) family. We are motivated to introduce the DEG-G family for the following reasons:

- Generating new probability mass functions that can be "asymmetric and left skewed", "asymmetric and right skewed", "symmetric", "symmetric and bimodal", "uniformed", or "right skewed with a heavy tail", among other useful shapes. The wide flexibility of the probability mass function (PMF) for any new model allows us to employ the new model for analyzing many different environmental datasets.
- Presenting some new special models with different types of hazard rate functions (HRFs), such as "increasing-constant", "constant", "monotonically decreasing", "bathtub", "monotonically increasing", "decreasing–increasing-decreasing", and "J-shaped". The more forms of failure rates, the greater the elasticity of the distribution. These shapes facilitate the work of many practitioners, who may use the new distribution in statistical modeling and mathematical analysis. For this specific purpose, we give the problem of checking the failure rate function a great deal of attention.
- The degrees of the skew coefficient, kurtosis coefficient, failure rate function, and diversity in the PMF and failure rate functions all play a role in the flexibility of the new distribution. Additionally, the probability distribution's usability and effectiveness in statistical modeling are crucial in this regard. Examining the novel PMF, we discovered that it was quite adaptable in this and other areas. This is what motivated us to investigate this probability distribution thoroughly.
- Proposing new discrete models for modeling "over-dispersed", "equal-dispersed", and "under-dispersed" real data. As shown in this paper, the new discrete family has shown a remarkable superiority in modeling these types of data, whether symmetric or asymmetric, and containing outliers or not containing outliers.
- Introducing new discrete models for analyzing extreme and zero-inflated count data.
- Comparison of the estimation methods with both simulated and real-life data for recommending the best method in each case.
- A zero-inflated probability distribution, or distribution that permits many zero-valued observations, is the foundation of a statistical model known as a zero-inflated model in statistics. For instance, individuals who have not purchased insurance against the risk, and are therefore unable to make a claim, would cause the quantity of insurance claims within a community for a particular type of risk to be zero-inflated. Often, the zero-inflated Poisson regression model is used for modeling and predicting the

zero-inflated count data; however, in this paper, we are motivated to use the DEG-G family for this purpose.

- In statistical modeling of the bathtub hazard rate count data, the DEG-G family under the Weibull baseline model provides adequate results; hence, the DEG-G family under the Weibull baseline is recommended for modeling the bathtub hazard rate count data (see Section 6.1). Moreover, the same baseline model is also suitable for modeling the monotonically increasing failure rate count data with adequate fitting (see Section 6.2).
- In the case of zero-inflated medical data with a decreasing failure rate and some outliers, the new family is an appropriate choice to deal with this type of data (see Section 6.3).
- In case of zero-inflated agricultural data with a decreasing–increasing–decreasing failure rate and some outliers, the new class is an appropriate choice for modeling this kind of data (see Section 6.4).
- In fact, we empirically demonstrate that the proposed family of distributions fits four real datasets more accurately than 16 other extended relevant distributions with 3–4 parameters (see Section 7).
- Through simulation experiments and relying on the new class, many of the classical estimation techniques and Bayesian approaches are tested and evaluated, and important conclusions are reached in this regard, including the following:

The maximum likelihood estimation approach is still the most efficient and most consistent of the rest of the classical methods; however, most of the other methods perform well, except for the Kolmogorov estimation method.

i. Generally, the Bayesian technique and maximum likelihood estimation method can be recommended for statistical modeling and applications.
ii. The Kolmogorov estimation method provides the worst results for all real datasets; this problem still needs more investigation for understanding of its main reasons.

Many helpful statistical properties, such as the probability-generating function, central and ordinary moment, moment-generating function, cumulant-generating function, and dispersion index (Disp-Ix), are calculated and statistically examined in this article once the new generator is defined. Some special discrete members, based on Weibull (W), inverse Weibull (IW), Lomax (Lx), Burr X (BX), inverse Burr X (IBX), log-logistic (LL), Rayleigh (R), inverse Rayleigh (IR), exponential (E), inverse Lindley (ILi), inverse Lomax (ILx), inverse log-logistic (ILL), inverse exponential (IE), and Lindley (Li) distributions, are listed in Table 1. Different classical (non-Bayesian) methods of estimation, including the Cramér–von Mises estimation (CVME), maximum likelihood estimation (MLE), ordinary least squares estimation (OLSE), bootstrapping (Bootst), L-moments (L-mom), Kolmogorov estimates (KE), weighted least squares estimation (WLSE), and Anderson–Darling left-tail from the second order (AD2LE), are considered. For more details about these methods, see the works of Chesneau et al. [12], Yousof et al. [18], Aboraya et al. [14], and Ibrahim et al. [15]. The Bayesian estimation under the squared error loss function is also considered. The well-known Markov chain Monte Carlo (MCMC) simulations are performed to compare the classical and Bayesian methods. The applicability of the DEG-G family is explained and discussed using four real-life datasets. The DEG-G family under the Weibull model case provides a more adequate fit than many competitive models, due to the consistent Akaike information criterion (CAICR), Akaike information criterion (AICR), chi-squared (χ^2_V), Kolmogorov–Smirnov (K–S), and corresponding p-value (P.V). For more detail about these statistics, see the works of El-Morshedy et al. [9,10], Aboraya et al. [14], and Eliwa et al. [17].

Table 1. Some new sub-models.

Baseline Model	$O(z;\gamma,\xi)$	Sub Model
LL	$(1+z)^{\gamma}-1\big\vert_{\gamma>0}$	DEG LL
W	$exp\left(\gamma z^{\theta}\right)-1\Big\vert_{\gamma,\theta>0}$	DEGW
IW	$\left[1-exp\left(-z^{-\theta}\right)\right]^{\gamma}-1\Big\vert_{\gamma,\theta>0}$	DEGIW
E	$exp(\gamma z)-1\vert_{\gamma>0}$	DEGE
IE	$\left[1-exp(-z^{-1})\right]^{\gamma}-1\Big\vert_{\gamma,\theta>0}$	DEGIE
Lx	$(1+z)^{\gamma\beta}-1\Big\vert_{\gamma,\beta>0}$	DEGLx
R	$exp(\gamma z)^{2}-1\Big\vert_{\gamma>0}$	DEGR
IR	$\left[1-exp(-z^{-2})\right]^{\gamma}-1\Big\vert_{\gamma>0}$	DEGR

2. The DEG-G Class

Starting with (1) and utilizing the discretization approach, the CDF of the DEG-G family can be formulated as follows:

$$F_{\boldsymbol{\mathcal{W}}}(z)=1-p^{O(z+1;\gamma,\xi)}\,|_{(p\in I=(0,1)\text{ and }z\in N_{(0)})} \tag{2}$$

where $\boldsymbol{\mathcal{W}}=(p,\gamma,\xi)$, $exp(-\alpha)=p$, $N_{(0)}=N\cup\{0\}$ and

$$O(z+1;\gamma,\xi)=\left[1-G_{\xi}(z+1)\right]^{-\gamma}-1=\dot{G}_{\xi}^{-\gamma}(z+1)\mid_{(z\in N_{(0)})}.$$

The corresponding SF to (2) can be derived as follows:

$$\overline{F}_{\boldsymbol{\mathcal{W}}}(z)=p^{O(z+1;\gamma,\xi)}\,|_{(p\in I\text{ and }z\in N_{(0)})}. \tag{3}$$

According to Kemp [19] and (3), the PMF of the DEG-G family can be expressed as follows:

$$f_{\boldsymbol{\mathcal{W}}}(z)=\overline{F}_{\boldsymbol{\mathcal{W}}}(z-1)-\overline{F}_{\boldsymbol{\mathcal{W}}}(z), \tag{4}$$

where

$$f_{\boldsymbol{\mathcal{W}}}(z)=p^{O(z;\gamma,\xi)}-p^{O(z+1;\gamma,\xi)}\,|_{(p\in I\text{ and }z\in \mathbf{N}_{(0)})}.$$

Based on (3) and (4), the hazard rate function (HRF) can then be proposed as follows: $H_{\boldsymbol{\mathcal{W}}}(z)=f_{\boldsymbol{\mathcal{W}}}(z)/\overline{F}_{\boldsymbol{\mathcal{W}}}(z-1)=\frac{1}{p^{O(z;\gamma,\xi)}}\left[p^{O(z;\gamma,\xi)}-p^{O(z+1;\gamma,\xi)}\right]$.

In Table 1, some members of the DEG-G family are provided. The new PMF in (4) is most tractable when the CDF of the baseline member $G_{\xi}(z)$ has a simple analytic expression.

For the W model, we have $O(z;\gamma,\xi)=\left[exp\left(\gamma z^{\theta}\right)-1\right]^{1}|_{\gamma,\theta>0}$. Then, based on (3), the PMF of the DEGW model can be expressed as follows:

$$f_{\boldsymbol{\mathcal{W}}}(z)=p^{exp(\gamma z^{\theta})-1}-p^{exp[\gamma(z+1)^{\theta}]-1}\,|_{(z\in N_{(0)},p\in I\text{ and }\gamma,\theta>0)}. \tag{5}$$

In Figures 1 and 2, some PMF and HRF plots of the DEGW model are sketched under some selected parameter values. Based on Figure 1, it can be seen that the PMF of the DEGW can be "asymmetric and left skewed", "asymmetric and right-skewed", "symmetric", "uniform", or "right-skewed with heavy tail", among other useful PMF shapes. Moreover, it can be used as a probability tool to discuss zero-inflated data. According to Figure 2, we can conclude that the HRF of the DEGW can be "increasing-constant", "constant", "monotonically decreasing ", "bathtub or decreasing-constant-increasing ", "monotonically increasing", or "J-shaped".

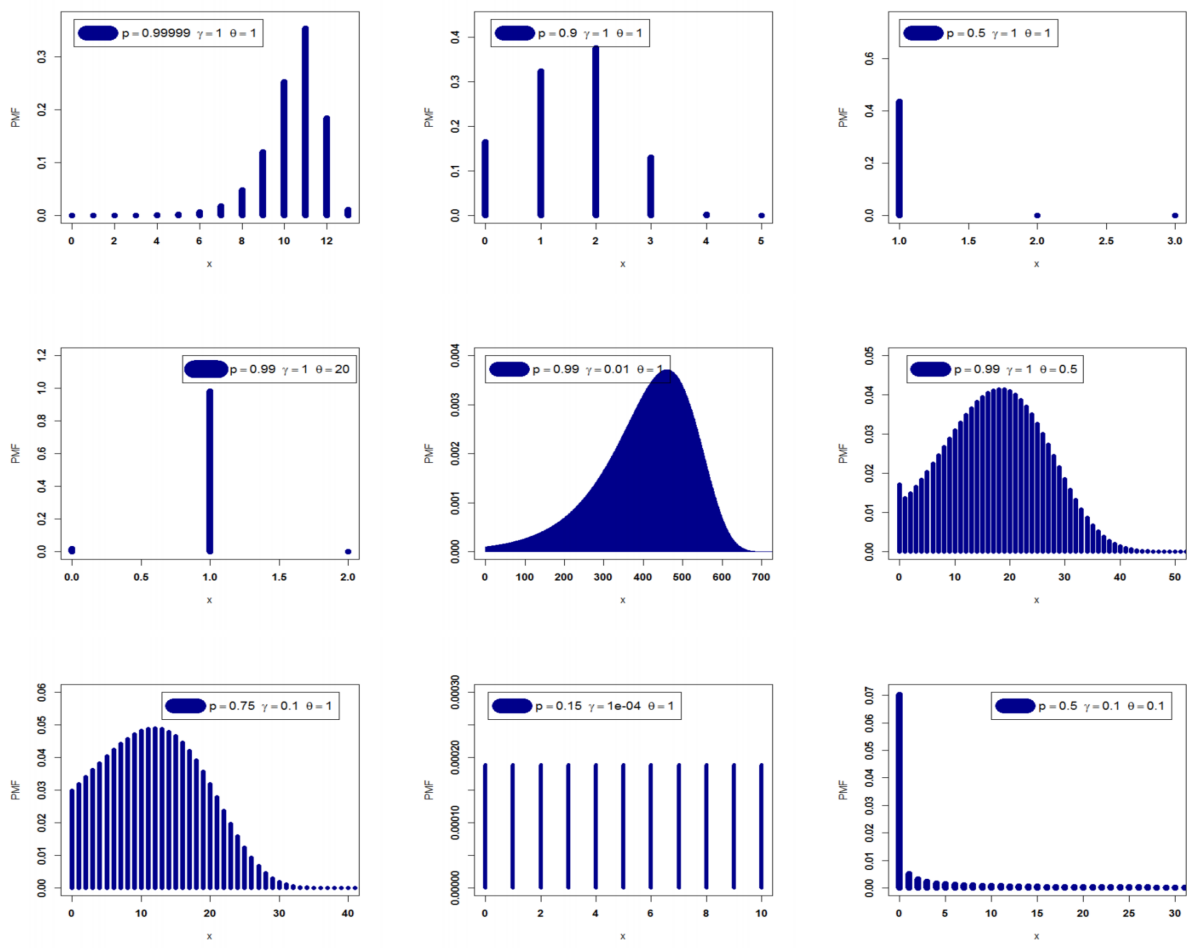

Figure 1. The PMF of the DEGW model.

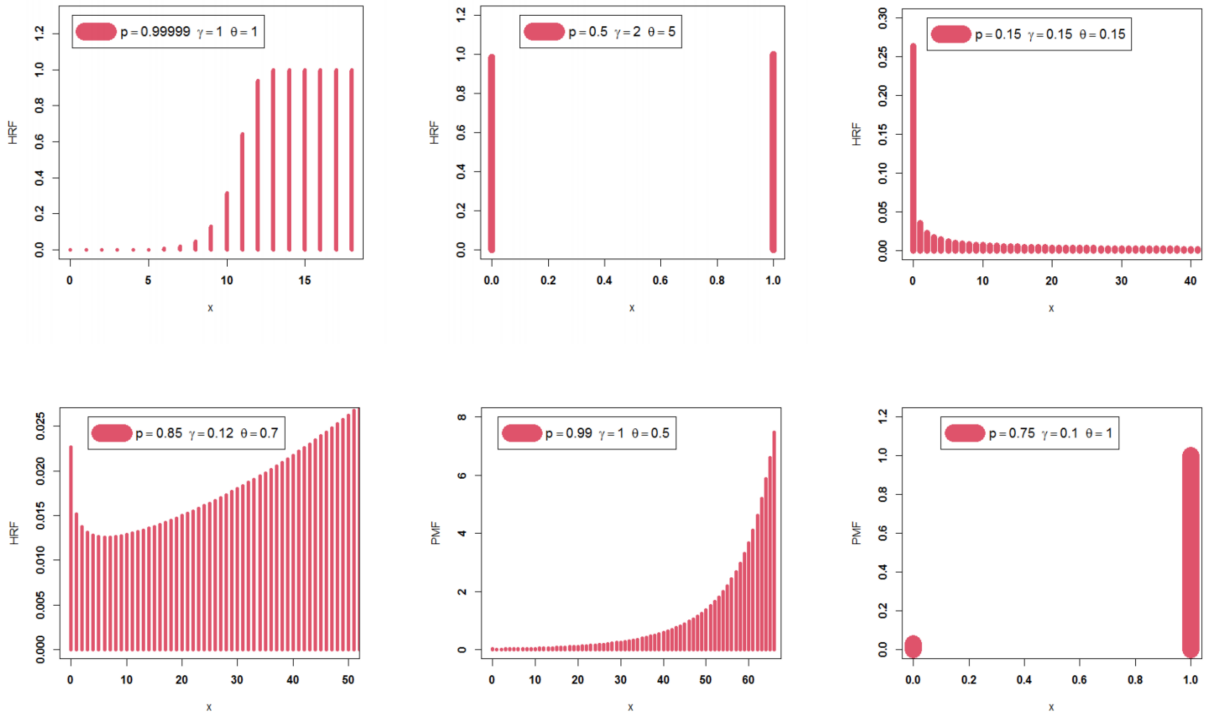

Figure 2. The HRF of the DEGW model.

3. Main Properties

3.1. Moments

Theorem 1. *Z is a non-negative random variable (RV), where Z ∼DEG-G ($\mathcal{W}$) family, then the r^{th} moment of the RV Z can be expressed as follows:*

$$\mu'_{r,Z} = E(Z^r) = \sum_{z=1}^{+\infty} \left[z^r - (z-1)^r\right] p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=1,2,3,\ldots)}.$$

Proof. Since

$$\mu'_{r,Z} = E(Z^r) = \sum_{z=0}^{+\infty} z^r f_{\mathcal{W}}(z).$$

then,

$$\mu'_{r,Z} = E(Z^r) = \sum_{z=0}^{+\infty} z^r \left[p^{O(z;\gamma,\xi)} - p^{O(z+1;\gamma,\xi)}\right] = \sum_{z=1}^{+\infty} \left[z^r - (z-1)^r\right] \overline{F}_{\mathcal{W}}(z-1).$$

Thus,

$$\mu'_{r,Z} = \sum_{z=1}^{+\infty} \left[z^r - (z-1)^r\right] p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=1,2,3,\ldots)}. \quad (6)$$

Using (5), the mean $(\mu_{1,Z})$, $\mu'_{2,Z}$, $\mu'_{3,Z}$, $\mu'_{4,Z}$, and variance $(V(Z))$, can be respectively written as follows:

$$\mu_{1,Z} = \mu'_{1,Z} = E(Z) = \sum_{z=1}^{+\infty} p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=1)},$$

$$\mu'_{2,Z} = E\left(Z^2\right) = \sum_{z=1}^{+\infty} (2z-1) p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=2)},$$

$$\mu'_{3,Z} = E\left(Z^3\right) = \sum_{z=1}^{+\infty} \left[z^3 - (z-1)^3\right] p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=3,\ldots)},$$

$$\mu'_{4,Z} = E\left(Z^4\right) = \sum_{z=1}^{+\infty} \left[z^4 - (z-1)^4\right] p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=4)}$$

and

$$V(Z) = \sum_{z=1}^{+\infty} (2z-1) p^{O(z;\gamma,\xi)} - \left(\sum_{z=1}^{+\infty} p^{O(z;\gamma,\xi)}\right)^2 |_{(z\in N_{(0)}, p\in I \text{ and } r=2)}. \square$$

Table 2 lists some numerical results for $\mu'_{1,Z}$, $\mu'_{2,Z}$, $\mu'_{3,Z}$, and $\mu'_{4,Z}$ under the DEGW model. It should be noted that the first four moments can be numerically evaluated, although they have no closed forms. All numerical results in Table 2 were derived using the R program, and the infinity problem was overcome by assuming a very large value instead of it (10^8), since values beyond this value can be ignored because they are too small. Therefore, the results of $\mu'_{1,Z}$, $\mu'_{2,Z}$, $\mu'_{3,Z}$, and $\mu'_{4,Z}$ are approximated.

Table 2. $\mu'_{1,Z}$, $\mu'_{2,Z}$, $\mu'_{3,Z}$, and $\mu'_{4,Z}$ of the DEGW distribution.

$\mathcal{W}$	$\approx\mu_{1,Z}$	$\approx\mu'_{2,Z}$	$\approx\mu'_{3,Z}$	$\approx\mu'_{4,Z}$
(0.10, 1.0, 1.0)	0.0191306	0.0191314	0.019133	0.0191363
(0.50, 1.0, 1.0)	0.3158439	0.3397145	0.387467	0.4830028
(0.99, 1.0, 1.0)	3.5744810	14.246690	60.28199	266.46120
(0.10, 0.5, 1.5)	0.2253027	0.2268436	0.2299255	0.2360891

Table 2. *Cont.*

$\mathcal{W}$	$\approx\mu_{1,Z}$	$\approx\mu'_{2,Z}$	$\approx\mu'_{3,Z}$	$\approx\mu'_{4,Z}$
(0.50, 0.5, 1.5)	0.7535885	0.9854351	1.4502100	2.3830030
(0.99, 0.5, 1.5)	3.4971860	13.014480	50.312590	200.27770
(0.75, 0.1, 1.0)	0.9713452	0.9736408	0.9782321	0.9874147
(0.75, 0.5, 1.0)	0.8297544	0.8297544	0.8297544	0.8297544
(0.75, 1.0, 1.0)	0.6099862	0.6099862	0.6099862	0.6099862
(0.75, 1.5, 1.0)	0.3672841	0.3672841	0.3672841	0.3672841
(0.1, 0.1, 1.0)	2.7581690	14.269410	94.86910	745.4975
(0.1, 0.1, 2.0)	1.1419800	1.9259440	3.704110	7.893829
(0.1, 0.1, 3.0)	0.8444187	0.9634029	1.2013710	1.677308
(0.1, 0.1, 4.0)	0.7850381	0.7852609	0.7857066	0.786598
(0.1, 0.1, 5.0)	0.7849267	0.7849267	0.7849267	0.7849267

3.2. Central Moment and Dispersion Index

The r^{th} central moment of the RV Z, i.e., $\mu_{r,Z}$, can be formulated as follows:

$$\mu_{r,Z} = E(Z - \mu'_{r,Z})^r = \sum_{\omega=0}^{r} (-\mu'_{1,Z})^{\omega} \binom{r}{\omega} \mu'_{r-\omega} |_{(|_{z\in N_{(0)}, p\in I \text{ and } r=1,2,3,\ldots})}.$$

Hence, the $V(Z)$ can be expressed as follows:

$$E(Z - \mu'_{r,Z})^2 = \mu_{1,Z} \sum_{\omega=0}^{2} (-\mu'_{1,Z})^{\omega} \binom{2}{\omega} \mu'_{2-\omega,Z} |_{(|_{z\in N_{(0)}, p\in I \text{ and } r=2})},$$

or

$$\mu_{2,Z} = V(Z) = \mu'_{2,Z} - \mu^2_{1,Z} = \sum_{z=1}^{+\infty} (2z-1) p^{O(z;\gamma,\xi)} - \left(\sum_{z=1}^{+\infty} p^{O(z;\gamma,\xi)} \right)^2 |_{(z\in N_{(0)}, p\in I \text{ and } r=2)}.$$

The dispersion index (Disp-Ix) of the DEG-G family can derived as follows:

$$\text{Disp-Ix}(Z) = \frac{\sum_{z=1}^{+\infty} (2z-1) p^{O(z;\gamma,\xi)}}{\sum_{z=1}^{+\infty} p^{O(z;\gamma,\xi)}} - \sum_{z=1}^{+\infty} p^{O(z;\gamma,\xi)} |_{(z\in N_{(0)} \text{ and } p\in I)}. \quad (7)$$

Some numerical results for the Disp-Ix(Z) are presented in Table 3, with useful comments. The Disp-Ix, also known as the coefficient of dispersion, relative variance, or variance-to-mean ratio (VMR), is a normalized measure of the dispersion of a probability distribution that is used in probability theory and statistics to determine whether a set of observed occurrences is clustered or dispersed in comparison to a common statistical model. It is described as the variance-to-mean ratio. In order to establish whether the observed real dataset can be modeled using a Poisson process, the Disp-Ix is used to measure whether a particular collection of observations is clustered or dispersed compared to a particular statistical model. When the Disp-Ix for any real dataset is less (greater) than 1, the dataset is referred to as under (over)-dispersed phenomena. Table 3 displays a numerical analysis and its associated computations for the Disp-Ix. The kurtosis, i.e., $K(Z)$, and skewness, i.e., $S(Z)$, of the RV Z can be obtained from the common relationships. Table 3 reports some numerical results for $E(Z)$, $V(Z)$, $S(Z)$, and $K(Z)$ of the DEGW distribution. Based on Table 3, it can be seen that $E(Z)$ increases as p increases, and decreases as γ and θ increase. The $S(Z) \in (-1.053185, \infty)$. The $K(Z)$ ranges from 1.02442 to ∞. The Disp-Ix$(Z) \in (0, 1)$, or " > 1", or "=1", like the standard well-known Poisson distribution (see Poisson [20]). Thus, the DEGW distribution could be useful in modeling "under-dispersed", "equi-dispersed", or "over-dispersed" count data.

Table 3. Numerical results for $E(Z)$, $V(Z)$, $S(Z)$, and $K(Z)$ of the DEGW distribution.

p	γ	θ	$E(Z)$	$V(Z)$	$S(Z)$	$K(Z)$	Disp-Ix(Z)
0.99999	1	1	10.43584	1.726577	−1.053185	5.105638	0.1654468
0.9999			8.134137	1.719510	−1.025913	4.907183	0.2113943
0.999			5.837383	1.673887	−0.929210	4.331043	0.2867530
0.99			3.574481	1.469777	−0.645338	3.230067	0.4111861
0.75			0.773245	0.510106	0.430022	2.279438	0.6596955
0.5			0.315844	0.239957	1.093988	2.899762	0.7597332
0.1			0.019131	0.018765	7.021297	50.30325	0.9809121
0.99999	10	5	0.802314	0.158607	−1.518194	3.304912	0.1976864
0.9999			0.110509	0.098297	2.484605	7.173261	0.8894908
0.999			2.6895×10^{-10}	2.6895×10^{-10}	60976.45	$\approx\infty$	1
0.99			7.2967×10^{-97}	7.2967×10^{-97}	$\approx\infty$	$\approx\infty$	1
0.5	0.001	2.5	12.741380	18.070030	−0.3936527	2.605803	1.4182170
	0.01		4.771435	2.934613	−0.3801066	2.62509	0.615038
	0.1		1.5940780	0.5300682	−0.2989435	2.847207	0.3325235
	0.25		0.9368591	0.2902798	−0.0505357	3.385981	0.3098437
	0.5		0.6378616	0.2310265	−0.5732478	1.329823	0.3621891
	0.65		0.5301451	0.2490913	−0.1208001	1.014593	0.4698549
	0.75		0.4610516	0.2484830	0.1562686	1.02442	0.5389484
	1.5		0.08951734	0.08150398	2.875646	9.269339	0.9104827
	2.5		0.00043026	0.00043009	48.17843	2322.161	0.9995697
	5		4.21×10^{-45}	4.21×10^{-45}	$\approx\infty$	$\approx\infty$	1
0.15	0.5	0.1	471.2272	47392362	51.07037	4085.929	100572.2
		0.5	0.5872831	1.464822	3.011574	15.06166	2.494234
		1	0.3318428	0.3039623	1.477674	4.494878	0.9159829
		5	0.2920874	0.2067724	0.9144595	1.836236	0.7079126
		20	0.2920874	0.2067724	0.9144595	1.836236	0.7079126
		150	0.2920874	0.2067724	0.9144595	1.836236	0.7079126

3.3. Generating Functions

Theorem 2. *Assume that Z is the non-negative RV, where $Z \sim$ DEG-G($\mathcal{W}$) class. Then, the moment-generating function (MGF) of the RV Z can be derived as follows:*

$$M_{r,Z}(t)|_{(z\in N_{(0)}, p\in I \text{ and } r=1,2,3,\ldots)} = 1 + \sum_{z=1}^{+\infty} \{exp(tz) - exp[t(z-1)]\} p^{O(z;\gamma,\xi)}. \quad (8)$$

Proof. The MGF of the non-negative RV Z can be derived as follows:

$$M_{r,Z}(t) = \sum_{z=0}^{+\infty} exp(tz)\left[p^{O(z;\gamma,\xi)} - p^{O(z+1;\gamma,\xi)}\right],$$

thus,

$$M_{r,Z}(t) = 1 + \sum_{z=1}^{+\infty} \{exp(tz) - exp[t(z-1)]\} p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=1,2,3,\ldots)}.$$

The first r derivatives of (6), with respect to t at $t = 0$, yield the first r moments around the origin, i.e.,

$$\mu'_{r,Z}|_{(t=0 \text{ and } r=1,2,3,\ldots)} = E(Z^r) = \frac{d^r}{dt^r} M_{r,Z}(t),$$

where

$$\mu'_{1,Z} = E(Z) = \frac{d}{dt} M_{r,Z}(t)|_{t=0} = \sum_{z=1}^{+\infty} p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=1)},$$

$$\mu'_{2,Z} = E\left(Z^2\right) = \frac{d^2}{dt^2} M_{r,Z}(t)|_{(t=0)} = \sum_{z=1}^{+\infty} (2z-1) p^{O(z;\gamma,\xi)}|_{(z\in N_{(0)}, p\in I \text{ and } r=2)},$$

$$\mu'_{3,Z} = E\left(Z^3\right) = \frac{d^3}{dt^3} M_{r,Z}(t)|_{(t=0)} = \sum_{z=1}^{+\infty} [3z(z-1)+1] p^{O(z;\gamma,\boldsymbol{\xi})}|_{(z\in N_{(0)}, p\in I \text{ and } r=3)},$$

and

$$\mu'_{4,Z} = E\left(Z^4\right) = \frac{d^4}{dt^4} M_{r,Z}(t)|_{(t=0)} = \sum_{z=1}^{+\infty} \left[z^4 - (z-1)^4\right] p^{O(z;\gamma,\boldsymbol{\xi})}|_{(z\in N_{(0)}, p\in I \text{ and } r=4)}. \square$$

The cumulant-generating function (CGF) is the logarithm of the MGF. Thus, the r^{th} cumulant, i.e., $C_{r,Z}$, can be derived as $C_{r,Z}|_{(t=0, \text{ and } r=1,2,3,\ldots)} = \frac{d^r}{dt^r} \log[M_{r,Z}(t)]$. In this context, we can highlight some important mathematical results: The 1st cumulant is the mean ($C_{1,Z} = \mu_{1,Z}$). The 2nd cumulant is known as the variance ($C_{2,Z} = \mu'_{2,Z} - \mu'^2_{1,Z} = \mu_{2,Z}$). The 3rd cumulant is the same as the 3rd central moment $C_{3,Z} = \mu'_{3,Z} - 3\mu'_{2,Z}\mu'_{1,Z} + 2\mu'^3_{1,Z} = \mu_{3,Z}$. However, the 4th and higher-order cumulants are not equal to central moments. The cumulants can be also expressed as follows:

$$C_{r,Z}|_{r\geq 1} = \mu'_{r,Z} - \sum_{\omega=0}^{r-1} \binom{r-1}{\omega-1} \mu'_{r-\omega,Z} C_{\omega,Z}.$$

It is possible to derive the probability-generating function (PGF) as follows:

$$P_Z(s) = 1 + \sum_{z=1}^{+\infty} \left(1 - \frac{1}{s}\right) s^z p^{O(z;\gamma,\boldsymbol{\xi})}|_{(z\in N_{(0)}, p\in I \text{ and } r=1,2,3,\ldots)}.$$

4. Estimation and Inference

In this section, we are concerned with the different estimation methods, including classical methods and Bayesian methods. The classical methods are many and varied, some of which depend on the theory of maximization, and some of which depend on the theory of minimization. In any case, the classical methods on the whole differ from Bayes' method in their origin and methodology of estimation, as will be explained in detail in theory and practice. The two subsections of this section cover Bayesian and non-Bayesian estimation techniques. Eight non-Bayesian estimation techniques, including the MLE, CVME, OLSE, WLSE, L-mom., KE, Bootst, and AD2LE methods, are taken into consideration in the first subsection. Then, the Bayesian estimation approach under the well-known squared error loss function (SELF) is taken into consideration in the second subsection.

4.1. Classical Estimation Techniques

4.1.1. The Maximum Likelihood Estimation Method

Maximum likelihood estimation (MLE) is a statistical technique for estimating the parameters of a probability distribution that has been assumed given some observed data. This is accomplished by maximizing a likelihood function to make the observed data as probable as possible given the assumed statistical model. The maximum likelihood estimate is the location in the parameter space where the likelihood function is maximized. Maximum likelihood is a popular approach for making statistical inferences, since its rationale is clear and adaptable.

The derivative test for figuring out maxima can be used if the likelihood function is differentiable. The ordinary least squares estimator, for example, maximizes the likelihood of the linear regression model, allowing the first-order requirements of the likelihood function to be explicitly solved in some circumstances. However, in many cases, it is essential to use numerical techniques to determine the probability function's maximum. MLE is typically comparable to maximum a posteriori (MAP) estimates under a uniform prior distribution of the parameters from the standpoint of Bayesian inference. MLE is a specific example of an extremum estimator in frequentist inference, with likelihood as the

objective function. If we assume a random sample $Z_1, Z_2, \ldots, Z_n$ from the presented class, then the log-likelihood function for the vector $\boldsymbol{\mathcal{W}}$ can be given as follows:

$$l(\boldsymbol{\mathcal{W}}) = \sum_{t=0}^{n} ln\Big[p^{O(z_{i,n};\gamma,\boldsymbol{\xi})} - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}\Big]\big|_{(p\in I \text{ and } z_{i,n}\in N_{(0)})}.$$

The $l(\boldsymbol{\mathcal{W}})$ can then be maximized via statistical programs such as "R", or by solving the nonlinear system obtained from $l(\boldsymbol{\mathcal{W}})$ by differentiation. Then, the score vectors

$$U(\boldsymbol{\mathcal{W}}) = \big(\partial l(\boldsymbol{\mathcal{W}})/\partial p, \partial l(\boldsymbol{\mathcal{W}})/\partial \gamma, \partial l(\boldsymbol{\mathcal{W}})/\partial \boldsymbol{\xi}_j\big)^T\big|_{j=1,2,\ldots,p},$$

can be easily derived as follows:

$$\partial l(\boldsymbol{\mathcal{W}})/\partial p = \sum_{t=0}^{n} \frac{O(z_{i,n};\gamma,\boldsymbol{\xi})p^{[O(z_{i,n};\gamma,\boldsymbol{\xi})]-1} - O(z_{i,n}+1;\gamma,\boldsymbol{\xi})p^{[O(z_{i,n}+1;\gamma,\boldsymbol{\xi})]-1}}{p^{O(z_{i,n};\gamma,\boldsymbol{\xi})} - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}},$$

$$\partial l(\boldsymbol{\mathcal{W}})/\partial \gamma = \sum_{t=0}^{n} \frac{\frac{\partial O(z_{i,n};\gamma,\boldsymbol{\xi})}{\partial \gamma}p^{O(z_{i,n};\gamma,\boldsymbol{\xi})}ln(p) - \frac{\partial O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}{\partial \gamma}p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}ln(p)}{p^{O(z_{i,n};\gamma,\boldsymbol{\xi})} - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}},$$

and

$$\partial l(\boldsymbol{\mathcal{W}})/\partial \boldsymbol{\xi}_j = \sum_{t=0}^{n} \frac{\frac{\partial O(z_{i,n};\gamma,\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_j}p^{O(z_{i,n};\gamma,\boldsymbol{\xi})}ln(p) - \frac{\partial O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_j}p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}ln(p)}{p^{O(z_{i,n};\gamma,\boldsymbol{\xi})} - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})}}\Big|_{j=1,2,\ldots,p}.$$

Setting

$$0 = \frac{\partial l(\boldsymbol{\mathcal{W}})}{\partial p}, 0 = \frac{\partial l(\boldsymbol{\mathcal{W}})}{\partial \gamma}, 0 = \frac{\partial l(\boldsymbol{\mathcal{W}})}{\partial \boldsymbol{\xi}_j},$$

and simultaneously solving them produces the MLEs for the DEG-G family parameters. For numerically addressing such problems, the Newton–Raphson algorithm is used.

4.1.2. The Cramér–von Mises Estimation Approach

Consider a random sample $Z_1, Z_2, \ldots, Z_n$ from the proposed generator. Then, the CVME of the parameter vector $\boldsymbol{\mathcal{W}}$ ($C_{(\boldsymbol{\mathcal{W}})}$) can be obtained by minimizing

$$C_{(\boldsymbol{\mathcal{W}})} = \frac{1}{12}n^{-1} + \sum_{i=1}^{n}\Big[F_{\boldsymbol{\mathcal{W}}}(z_{i,n}) - \omega_{(i,n)}^{[1]}\Big]^2\big|_{(p\in I \text{ and } z_{i,n}\in N_{(0)})},$$

with respect to (w.r.t) p, γ, and $\boldsymbol{\xi}$, respectively, where $\omega_{(i,n)}^{[1]} = \frac{1}{2n}(2i-1)$ and

$$C_{(\boldsymbol{\mathcal{W}})} = \sum_{i=1}^{n}\Big[1 - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})} - \omega_{(i,n)}^{[1]}\Big]^2.$$

The three nonlinear equations below are then solved to yield the CVME of the parameters p, γ, and $\boldsymbol{\xi}$:

$$0 = \sum_{i=1}^{n}\Big(1 - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})} - \omega_{(i,n)}^{[1]}\Big)\varsigma_{(p)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

$$0 = \sum_{i=1}^{n}\Big(1 - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})} - \omega_{(i,n)}^{[1]}\Big)\varsigma_{(\gamma)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

and

$$0 = \sum_{i=1}^{n}\Big(1 - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})} - \omega_{(i,n)}^{[1]}\Big)\varsigma_{(\boldsymbol{\xi}_j)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

where

$$\varsigma_{(p)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}) = \partial F_{\boldsymbol{\mathcal{W}}}(z_{i,n})/\partial p, \tag{9}$$

$$\varsigma_{(\gamma)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}) = \partial F_{\boldsymbol{\mathcal{W}}}(z_{i,n})/\partial \gamma, \tag{10}$$

and

$$\varsigma_{(\xi_j)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}) = \partial F_{\boldsymbol{\mathcal{W}}}(z_{i,n})/\partial \xi_j, \tag{11}$$

are the first derivatives of the CDF of DEG-G distribution w.r.t $p,\ \gamma$, and ξ_j, respectively.

4.1.3. The Ordinary Least Squares Technique

Geometrically, this is defined as the total of the squared distances between each data point in the set and its corresponding point on the regression surface, which are measured parallel to the axis of the dependent variable. The lower the differences, the better the model fits the data. Particularly in the case of a basic linear regression, when there is only one regressor on the right-hand side of the regression equation, the resultant estimator can be stated by a straightforward formula. If $F_{\boldsymbol{\mathcal{W}}}(z_{i,n})$ denotes the CDF of the DEG-G family and $Z_1 < Z_2 < \cdots < Z_n$ represents the n-ordered random sample, then the OLSEs$_{(\boldsymbol{\mathcal{W}})}$ can be obtained upon minimizing

$$\text{OLSEs}_{(\boldsymbol{\mathcal{W}})} = \sum_{i=1}^{n}\Big[1-\omega_{(i,n)}^{[2]} - p^{O(z_{i,n}+1;\gamma,\xi)}\Big]^2,$$

with respect to $p,\ \gamma$, and ξ, respectively, where $\omega_{(i,n)}^{[2]} = \frac{i}{n+1}$. The OLSEs are obtained by solving the following nonlinear equations:

$$0 = \sum_{i=1}^{n}\Big[1-p^{O(z_{i,n}+1;\gamma,\xi)} - \omega_{(i,n)}^{[2]}\Big]\varsigma_{(p)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

$$0 = \sum_{i=1}^{n}\Big[1-p^{O(z_{i,n}+1;\gamma,\xi)} - \omega_{(i,n)}^{[2]}\Big]\varsigma_{(\gamma)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

and

$$0 = \sum_{i=1}^{n}\Big[1-p^{O(z_{i,n}+1;\gamma,\xi)} - \omega_{(i,n)}^{[2]}\Big]\varsigma_{(\xi_j)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

with respect to $p,\ \gamma$, and ξ, respectively, where $\varsigma_{(p)}(z_{i,n}+1,\boldsymbol{\mathcal{W}})$, $\varsigma_{(\gamma)}(z_{i,n}+1,\boldsymbol{\mathcal{W}})$, and $\varsigma_{(\xi_j)}(z_{i,n}+1,\boldsymbol{\mathcal{W}})$ are defined in (9), (10) and (11), respectively.

4.1.4. The Weighted Least Squares Estimation Method

Ordinary least squares and linear regression can be generalized into weighted least squares (WLS), also known as weighted linear regression (WLR), which incorporates knowledge of the variance of the observations into the regression. Another variation of generalized least squares is WLS. If $F_{\boldsymbol{\mathcal{W}}}(z_{i,n})$ denotes the CDF of the DEG-G class, and we assume that $Z_1 < Z_2 < \cdots < Z_n$ is the n-ordered random sample, then the WLSE can be derived by minimizing the function $W_{(\boldsymbol{\mathcal{W}})}$ with respect to $p,\ \gamma$, and ξ_j, where

$$W_{(\boldsymbol{\mathcal{W}})} = \sum_{i=1}^{n} d_{(i,n)}^{[3]}\Big[F_{\boldsymbol{\mathcal{W}}}(z_{i,n}) - \omega_{(i,n)}^{[2]}\Big]^2,$$

and $d_{(i,n)}^{[3]} = \Big[(1+n)^2(2+n)\Big]/[i(1+n-i)]$. Furthermore, the WLSEs can be reported by solving

$$0 = \sum_{i=1}^{n} d_{(i,n)}^{[3]}\Big[1-p^{O(z_{i,n}+1;\gamma,\xi)} - \omega_{(i,n)}^{[2]}\Big]\varsigma_{(p)}(z_{i,n}+1,\boldsymbol{\mathcal{W}}),$$

$$0 = \sum_{i=1}^{n} d_{(i,n)}^{[3]} \left[1 - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})} - \omega_{(i,n)}^{[2]}\right] \varsigma_{(\gamma)}(z_{i,n}+1, \boldsymbol{\mathcal{W}}),$$

and

$$0 = \sum_{i=1}^{n} d_{(i,n)}^{[3]} \left[1 - p^{O(z_{i,n}+1;\gamma,\boldsymbol{\xi})} - \omega_{(i,n)}^{[2]}\right] \varsigma_{(\boldsymbol{\xi}_j)}(z_{i,n}+1, \boldsymbol{\mathcal{W}}),$$

with respect to p, γ, and $\boldsymbol{\xi}$, respectively, where $\varsigma_{(p)}(z_{i,n}+1, \boldsymbol{\mathcal{W}}), \varsigma_{(\gamma)}(z_{i,n}+1, \boldsymbol{\mathcal{W}})$, and $\varsigma_{(\boldsymbol{\xi}_j)}(z_{i,n}+1, \boldsymbol{\mathcal{W}})$ are defined in (9), (10) and (11), respectively.

4.1.5. L-Moments Estimation Approach

For a random sample taken from a certain population, the sample's L-mom can be established and utilized as estimators of the population's L-mom. The L-mom for the population can be obtained from

$$L_{(r)} = \frac{1}{r} \sum_{m=0}^{r-1} (-1)^m \binom{r-1}{m} E(Z_{r-m:m})|_{(r\geq 1)}.$$

The first three L-mom can be expressed as follows:

$$L_{(1)}(p, \gamma, \boldsymbol{\xi}) = E(z_{1:1}) = \mu'_1 = L_{(1)},$$

$$L_{(2)}(p, \gamma, \boldsymbol{\xi}) = \frac{1}{2} E(z_{2:2} - z_{1:2}) = \frac{1}{2}(\mu'_{2:2} - \mu'_{1:2}) = L_{(2)},$$

and

$$L_{(3)}(p, \gamma, \boldsymbol{\xi}) = \frac{1}{3} E(z_{3:3} - 2z_{2:3} + z_{1:3}) = \frac{1}{3}(\mu'_{3:3} - 2\mu'_{2:3} + \mu'_{1:3}) = L_{(3)},$$

where $L_i|_{(i=1.2.3)}$ is the L-mom for the sample. Then, the L-mom estimators of the parameters p, γ, and $\boldsymbol{\xi}$ can be obtained by solving the following three equations numerically

$$L_{(1)}(\hat{p}, \hat{\gamma}, \hat{\boldsymbol{\xi}}) = L_{(1)}, L_{(2)}(\hat{p}, \hat{\gamma}, \hat{\boldsymbol{\xi}}) = L_{(2)},$$

and

$$L_{(3)}(\hat{p}, \hat{\gamma}, \hat{\boldsymbol{\xi}}) = L_{(3)}.$$

4.1.6. The Kolmogorov Estimation Method

The Kolmogorov estimates (KEs) can be obtained by minimizing the following function:

$$K = K(p, \gamma, \boldsymbol{\xi}_j) = \overset{1\leq i\leq n}{max} \left\{ \frac{1}{n} i - F_{\boldsymbol{\mathcal{W}}}(z_{i,n}), F_{\boldsymbol{\mathcal{W}}}(z_{i,n}) - \frac{1}{n}(i-1) \right\}.$$

For estimating each parameter, the KEs are obtained by comparing $\left[\frac{1}{n} i - F_{\boldsymbol{\mathcal{W}}}(z_{i,n})\right]|_{1\leq i\leq n}$ and $F_{\boldsymbol{\mathcal{W}}}(z_{i,n}) - \frac{1}{n}(i-1)$ and selecting the maximum. However, for $1 \leq i \leq n$, we minimize the whole function $K(p, \gamma, \boldsymbol{\xi}_j)$. For more detail about the KE approach, see the work of Aguilar et al. [21].

4.1.7. Bootstrapping Technique

Bootst, which is a type of test or metric that mimics the sampling process by using random sampling with replacement, belongs to the larger category of resampling techniques. With Bootst, sample estimates are given accuracy ratings such as bias, variance, confidence intervals, prediction error, etc. Using random sampling techniques, this strategy enables estimation of the sample distribution for nearly any statistic. The observed data's empirical distribution function is a common option for an approximate distribution. A few resamples with replacement of the observed dataset can be constructed in the case where a set of observations can be believed to come from an independent and identically distributed

population (and of equal size to the observed dataset). A potent statistical procedure, the Bootst approach is particularly helpful with small sample size. The assumption of a "normal" or "t" distribution cannot normally be utilized to cope with sample sizes smaller than 40 (see Efron [22] and Hesterberg [5]). Techniques for Bootst perform very well with samples that contain fewer than 40 observations. This is because Bootst requires resampling. These methods make no assumptions regarding the distribution of our data. With the increased accessibility of computational resources, Bootst has grown in popularity. This is due to the necessity of using a computer for Bootst to be useful. In the application section, we examine this in more detail.

4.1.8. The Anderson–Darling Left-Tail from the Second Order Approach

The AD2LEs of p, γ, and ξ_j can be obtained by minimizing

$$AD2LE(\xi) = \frac{1}{n}\sum_{i=1}^{n}\frac{2i-1}{F_{\mathcal{W}}(z_{i,n})} + 2\sum_{i=1}^{n}\log[F_{\mathcal{W}}(z_{i,n})].$$

Thus, the AD2LEs can be derived by solving the following nonlinear equations:

$$\partial[AD2LE(\xi)]/\partial p = 0, \partial\left[AD2LE\left(\xi_j\right)\right]/\partial\gamma = 0$$

and

$$\partial[AD2LE(\xi)]/\partial\xi_j = 0.$$

4.2. Bayesian Estimation

Before we can even begin to discuss how a Bayesian approach might estimate a population parameter, we must first recognize one important distinction between frequentist and Bayesian statisticians. The distinction is whether a statistician views a parameter as a random variable or as an unknowable constant. A Bayes estimator, also known as a Bayes action, is an estimator or decision rule used in estimation theory and decision theory that minimizes the posterior expected value of a loss function (i.e., the posterior expected loss). In other words, it optimizes the utility function's posterior expectation. Maximum a posteriori estimation is a different approach to constructing an estimator in the context of Bayesian statistics. We can assume the following prior distributions for the parameters p, γ, and ξ_j, where

$$p_{1;(\varsigma_1,\varpi_1)}(p) \sim \text{beta}(\varsigma_1,\varpi_1), p_{2;(\varsigma_2,\varpi_2)}(\gamma) \sim \text{Gamma}(\varsigma_2,\varpi_2)$$

and

$$p_{3;(\varsigma_3,\varpi_3)}\left(\xi_j\right) \sim \text{Uniform}(\varsigma_3,\varpi_3).$$

We can also assume that the parameters are independently distributed. The joint prior distribution can be written as follows:

$$\pi_{(\varsigma_i,\varpi_i)}\left(p,\gamma,\xi_j\right) = \frac{p^{\varsigma_1}(1-p)^{\varpi_1}\varpi_2^{\varsigma_2}}{(\varpi_3-\varsigma_3)B(\varsigma_1,\varpi_1)\Gamma(\varsigma_2)}\gamma^{\,\varsigma_2-1}exp[-(\gamma\varpi_2)].$$

The posterior distribution $p\left(p,\gamma,\xi_j|\underline{X}\right)$ of the parameters is defined as follows:

$$p\left(p,\gamma,\xi_j|\underline{Z}\right) \propto \text{likelihood}(\boldsymbol{\mathcal{W}}_j|\underline{Z}) \times \pi_{(\varsigma_i,\varpi_i)}\left(p,\gamma,\xi_j\right),$$

where the likelihood $(\boldsymbol{\mathcal{W}}_j|\underline{Z}) = \prod_{i=1}^{n} f_{\mathcal{W}}(z_i)$ and $\pi_{(\varsigma_i,\varpi_i)}\left(p,\gamma,\xi_j\right)$ is the joint prior distribution. Under SELF, the Bayesian estimators of p, γ, and ξ_j are the means of their marginal posteriors. Those marginal posteriors' formulae cannot be utilized to derive the Bayesian estimates. Hence, the numerical approximation is required. Markov chain Monte Carlo (MCMC) methods are a class of algorithms used in statistics for probability distribution sampling. One can obtain a sample of the desired distribution by building a Markov chain

with the desired distribution as its equilibrium distribution and recording states from the chain. The distribution of the sample closely resembles the real target distribution as the number of steps increases. For building chains, a few algorithms are available—notably the Metropolis–Hastings algorithm. In this work, we suggest using the Gibbs sampler and Metropolis–Hastings (M–H) algorithm—two MCMC approaches. Since the conditional posteriors of the parameters p, γ, and ξ_j cannot be obtained in any standard forms, it is advisable to pull samples from the joint posterior of the parameters using a hybrid MCMC. The full conditional posteriors of p, γ, and ξ_j can be easily calculated. The simulation algorithm can be summarized in the following steps:

(1) Assume the initial values for p, γ, and ξ_j at the $\imath^{(\text{th})}$ stage.
(2) Consider the elementary values for p, γ, and ξ_j at the $\imath^{\text{th}}$ stage.
(3) The M–H approach is utilized to derive

$$p_{(\imath)} \sim p_1\left(p_{(\imath)}, \underline{X}\right) | p_{(\imath-1)}, \gamma_{(\imath-1)}, \xi_{j(\imath-1)},$$
$$\gamma_{(\imath)} \sim p_2\left(\gamma_{(\imath)}, \underline{X}\right) | p_{(\imath)}, \gamma_{(\imath-1)}, \xi_{j(\imath-1)},$$

and

$$\xi_{j(\imath)} \sim p_2\left(\gamma_{(\imath)}, \underline{X}\right) | p_{(\imath)}, \gamma_{(\imath)}, \xi_{j(\imath-1)}.$$

(1) To obtain the samples of size from the relevant posteriors of interest, repeat steps 2–3 for M = 100,000 times.
(2) Obtain the Bayesian estimates of p, γ, and ξ_j using the following formulae:

$$\hat{p}_{\text{Bayes}} = \frac{1}{M - M_0} \sum_{h=1+M_0}^{M} p^{[h]},$$

$$\hat{\gamma}_{\text{Bayes}} = \frac{1}{M - M_0} \sum_{h=1+M_0}^{M} \gamma^{[h]} \text{ and } \hat{\xi}_{j_{\text{Bayes}}} = \frac{1}{M - M_0} \sum_{h=1+M_0}^{M} \xi_j{}^{[h]},$$

respectively, where $M_0 \approx 50,000$ is the burn-in period of the generated MCMC.

5. Simulations: Comparing Classical and Bayesian Estimation Methods

For comparing the classical and Bayesian methods, MCMC simulation studies were performed. The results are presented in Table 4 ($p = 0.1, \gamma = 0.2, \theta = 0.9 \,|\, n = 50, 100, 200, 300$), Table 5 ($p = 0.3, \gamma = 0.9, \theta = 0.2 \,|\, n = 50, 100, 200, 300$), Table 6 ($p = 0.5, \gamma = 0.3, \theta = 0.6 \,|\, n = 50, 100, 200, 300$), and Table 7 ($p = 0.9, \gamma = 0.9, \theta = 1.2 \,|\, n = 50, 100, 200, 300$). The numerical assessments were performed based on the mean squared errors (MSEs). First, we generated $N = 1000$ samples of the DEGW model. Based on Tables 4–7, it should be noted that the performance of all estimation methods improves when $n \to +\infty$. The value with "*" is the best estimation in its row for all estimation methods. Generally, it should be noted that the MLE and the Bayesian methods are recommended for statistical modeling and applications; this assessment, as shown in Tables 4–7, is mainly reliant on a comprehensive simulation study, and the simulation, as is well known, precedes the application on real data. Additionally, despite the diversity of classical methods and their abundance, the MLE method is still the most efficient and most consistent of the rest of the classical methods; however, most of the other methods perform well. In this section, we use simulation studies to evaluate different estimation methods, and not to compare these methods; this does not preclude the use of simulation for comparisons between different estimation methods, but the real data are often used to compare the different estimation methods, and this is what prompts us to present four applications for this specific purpose. This is in addition to four other applications to compare the competing models.

Table 4. Results of the MSEs where $p = 0.1$, $\gamma = 0.2$, $\theta = 0.9$.

n		MLE	OLS	WLS	CVM	Bayesian	L-mom	KE	Bootst	AD2LE
50	p_0	0.00111	0.00116	0.00178	0.00126	0.00069 *	0.00140	0.00120	0.00155	0.00178
	θ_0	0.00160 *	0.02066	0.01452	0.01760	0.00430	0.00211	0.00885	0.00258	0.02066
	γ_0	0.00113	0.00080	0.00112	0.00071	0.00042 *	0.00144	0.00066	0.00130	0.00095
100	p_0	0.00058	0.00068	0.00106	0.00072	0.00052 *	0.00082	0.00066	0.00096	0.00094
	θ_0	0.00067 *	0.00886	0.00508	0.00816	0.00415	0.00094	0.00456	0.00255	0.01008
	γ_0	0.00058	0.00040	0.00063	0.00038	0.00036 *	0.00083	0.00033	0.00076	0.00051
200	p_0	0.00026	0.00036	0.00065	0.00037	0.00022 *	0.00037	0.00032	0.00032	0.00050
	θ_0	0.00030 *	0.00401	0.00186	0.00385	0.00254	0.00039	0.00221	0.00065	0.00493
	γ_0	0.00026	0.00020	0.00037	0.00020	0.00014 *	0.00037	0.00017	0.00028	0.00028
300	p_0	0.00019 *	0.00025	0.00043	0.00025	0.00021	0.00024	0.00022	0.00023	0.00033
	θ_0	0.00020 *	0.00248	0.00118	0.00243	0.00078	0.00024	0.00154	0.00040	0.00315
	γ_0	0.00020	0.00013	0.00025	0.00013	0.00009 *	0.00024	0.00011	0.00017	0.00018

The value with "*" is the best estimation in its row for all estimation methods.

Table 5. Results of the MSEs where $p = 0.5$, $\gamma = 0.3$, $\theta = 0.6$.

n		MLE	OLS	WLS	CVM	Bayesian	L-mom	KE	Bootst	AD2LE
50	p_0	0.00336 *	0.00403	0.00400	0.00391	0.00503	0.01825	0.00382	0.00679	0.00396
	θ_0	0.00468	0.03063	0.03013	0.01105	0.00613	0.01639	0.00448 *	0.00448 *	0.01126
	γ_0	0.00380 *	0.01173	0.01155	0.01050	0.02408	0.02421	0.01026	0.01313	0.01065
100	p_0	0.00166 *	0.00194	0.00193	0.00190	0.00198	0.01294	0.00173	0.00204	0.00192
	θ_0	0.00178	0.00688	0.00524	0.00530	0.00099 *	0.01224	0.00140	0.00212	0.00548
	γ_0	0.00171 *	0.00543	0.00537	0.00513	0.00392	0.01420	0.00438	0.00243	0.00520
200	p_0	0.00081 *	0.00101	0.00102	0.00100	0.00162	0.00930	0.00090	0.00135	0.00101
	θ_0	0.00083	0.00287	0.00175	0.00248	0.00039 *	0.00919	0.00067	0.00083	0.00259
	γ_0	0.00082 *	0.00274	0.00275	0.00268	0.00297	0.00944	0.00221	0.00220	0.00271
300	p_0	0.00056 *	0.00066	0.00067	0.00065	0.00068	0.00649	0.00057	0.00134	0.00065
	θ_0	0.00057	0.00159	0.00080	0.00143	0.00035 *	0.00649	0.00049	0.00065	0.00151
	γ_0	0.00058 *	0.00177	0.00180	0.00173	0.00148	0.00650	0.00137	0.00214	0.00175

The value with "*" is the best estimation in its row for all estimation methods.

Table 6. Results of the MSEs where $p = 0.5$, $\gamma = 0.3$, $\theta = 0.6$.

n		MLE	OLS	WLS	CVM	Bayesian	L-mom	KE	Bootst	AD2LE
50	p_0	0.00241 *	0.00333	0.00498	0.00312	0.01159	0.00347	0.00308	0.00320	0.00480
	θ_0	0.00242	0.00289	0.00189 *	0.00250	0.00442	0.00339	0.00299	0.00327	0.00499
	γ_0	0.00240	0.00119	0.00225	0.00105 *	0.00316	0.00337	0.00106	0.00322	0.00202
100	p_0	0.00129	0.00175	0.00314	0.00170	0.00339	0.00181	0.00160	0.00111 *	0.00265
	θ_0	0.00130	0.00139	0.00074	0.00130	0.00042 *	0.00180	0.00149	0.00218	0.00248
	γ_0	0.00129	0.00059	0.00135	0.00056	0.00038 *	0.00180	0.00053	0.00116	0.00108
200	p_0	0.00058 *	0.00084	0.00195	0.00082	0.00063	0.00083	0.00079	0.00067	0.00132
	θ_0	0.00059	0.00062	0.00032	0.00059	0.00019 *	0.00082	0.00071	0.00056	0.00111
	γ_0	0.00058	0.00027	0.00082	0.00026	0.00014 *	0.00082	0.00026	0.00056	0.00053
300	p_0	0.00040	0.00049	0.00131	0.00049	0.00039 *	0.00055	0.00048	0.00043	0.00078
	θ_0	0.00040	0.00037	0.00018	0.00036	0.00014 *	0.00055	0.00045	0.00043	0.00065
	γ_0	0.00040	0.00016	0.00055	0.00016	0.00009 *	0.00055	0.00015	0.00043	0.00031

The value with "*" is the best estimation in its row for all estimation methods.

Table 7. Results of the MSEs where $p = 0.9$, $\gamma = 0.9$, $\theta = 1.2$.

n		MLE	OLS	WLS	CVM	Bayesian	L-mom	KE	Bootst	AD2LE
50	p_0	0.00026	0.00028	0.00060	0.00026	0.00022 *	0.00023	0.00030	0.00268	0.00046
	θ_0	0.00041	0.00625	0.00578	0.00547	0.00271	0.00030 *	0.00428	0.01436	0.00725
	γ_0	0.00034	0.00342	0.01019	0.00318	0.00252	0.00026 *	0.00463	0.01624	0.00783
100	p_0	0.00012 *	0.00015	0.00035	0.00014	0.00012 *	0.00012 *	0.00016	0.00021	0.00023
	θ_0	0.00012 *	0.00337	0.00270	0.00308	0.00269	0.00014	0.00264	0.00053	0.00435
	γ_0	0.00012 *	0.00189	0.00653	0.00181	0.00161	0.00013	0.00246	0.00097	0.00375
200	p_0	0.00006 *	0.00007	0.00020	0.00007	0.00010	0.00006 *	0.00007	0.00006	0.00011
	θ_0	0.00006 *	0.00155	0.00107	0.00147	0.00105	0.00006 *	0.00132	0.00019	0.00222
	γ_0	0.00006 *	0.00091	0.00412	0.00090	0.00096	0.00006 *	0.00114	0.00018	0.00171
300	p_0	0.00004 *	0.00004 *	0.00014	0.00004 *	0.00009	0.00004 *	0.00005	0.00006	0.00007
	θ_0	0.00004 *	0.00084	0.00059	0.00082	0.00103	0.00004 *	0.00084	0.00007	0.00138
	γ_0	0.00004 *	0.00055	0.00325	0.00055	0.00093	0.00004 *	0.00072	0.00012	0.00110

The value with "*" is the best estimation in its row for all estimation methods.

6. Comparing Various Estimation Techniques Via Real Data

6.1. Dataset I: Failure Times

This dataset represents the failure times of 50 devices, in weeks. The data observations are available in the work of Bodhisuwan and Sangpoom [23], and were recently analyzed

by Eliwa et al. [17], Aboraya et al. [14], and Ibrahim et al. [15]. Table 8 lists the estimates, K–S, and P.V statistics for failure time data. Based on Table 8, it can be seen that the Bayesian method is the best, with K–S $= 0.14712$ and P.V $= 0.22927$, followed by the MLE method, with K–S = 0.163038 and P.V = 0.15266. The KE method provides undesirable results or unexpected results (K–S = 0.51000 and P.V < 0.0001), and this may be due to the nature of the data used, or to any other random reasons. In any case, these results need further study and analysis, one way or another. Figure 3 gives the Kaplan–Meier (estimated survival function) plots using failure time data for the nine estimation methods. The graphical results in Figure 3 confirm and support the numerical results shown in Table 8.

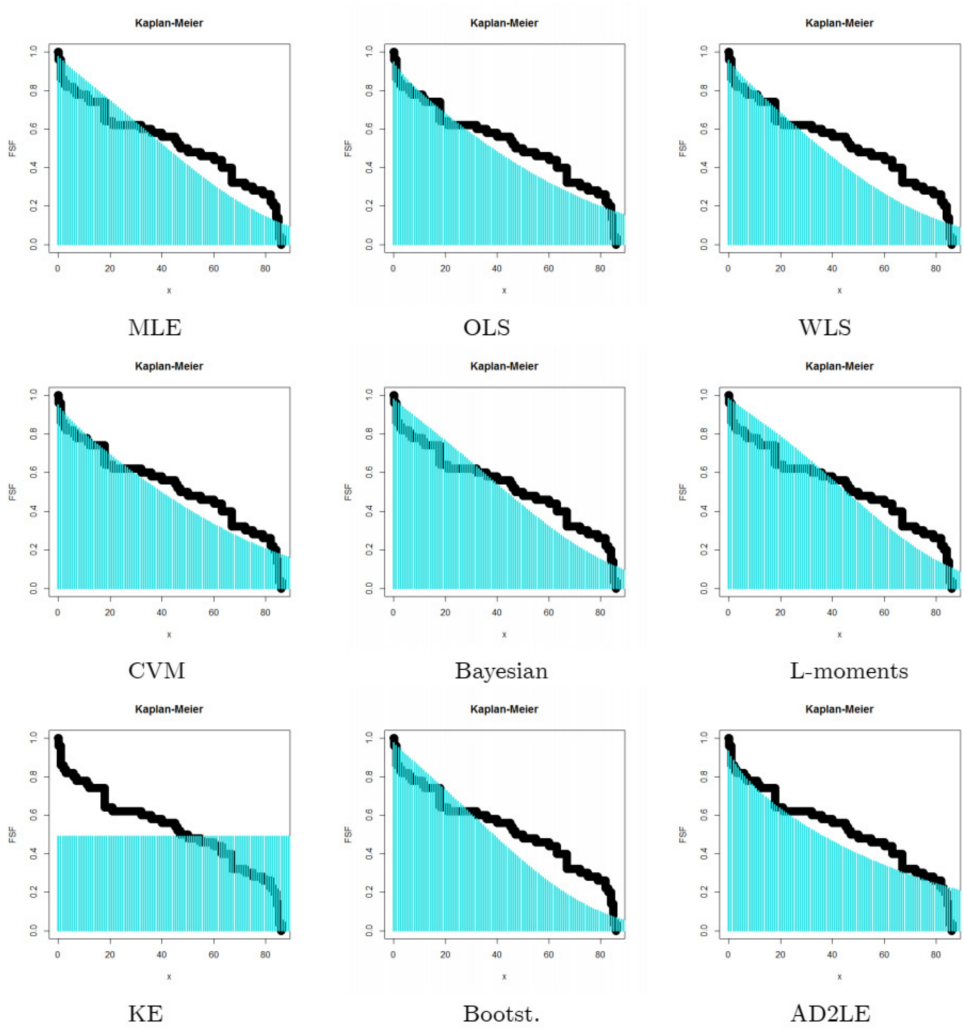

Figure 3. Kaplan–Meier plots based on dataset I.

Table 8. Comparing methods using dataset I.

Method	p	γ	θ	K.S	P.V
MLE	0.8450441207	0.1248812582	0.6840750167	0.163038	0.15266
OLS	0.9115300723	0.4505403055	0.4249120961	0.16907	0.11468
WLS	0.9273026987	0.4467200744	0.4572149891	0.20003	0.03659
CVM	0.9183256058	0.4514888661	0.4288501393	0.17538	0.09231
* Bayesian	0.8344024241	0.0987677765	0.7294684266	0.14712	0.22927
L-mom	0.8883682738	0.1303778446	0.7019449219	0.16801	0.11884
KE	0.0173097899	0.1619946274	0.0000203909	0.51000	<0.0001
Bootst	0.8465409851	0.1275235069	0.6948444678	0.21355	0.02092
AD2LE	0.6215684527	0.1341556153	0.5297912271	0.22093	0.01518

The value with "*" is the best estimation in its row for all estimation methods.

6.2. Dataset II: Failure Times of 15 Electronic Components

This lifetime data gives the failure times for 15 electronic components in an acceleration lifetime test (see Lawless et al. [24]). Table 9 gives the estimates, K–S, and P.V statistics for second failure time data. Based on Table 9, it can be noted that the AD2LE method is the best, with K–S = 0.09885 and P.V = 0.99855, followed by the Bayesian method, with K–S = 0.09937 and P.V = 0.99843. Figure 4 gives the Kaplan–Meier plots using second failure time data for the nine estimation methods. The graphical results in Figure 4 support the results in Table 9. Again, the KE method provided undesirable results or unexpected results (K–S = 0.53331 and P.V = 0.00039), and this may be due to the nature of the data used, or to any other random reasons. In any case, these results need further study and analysis, one way or another.

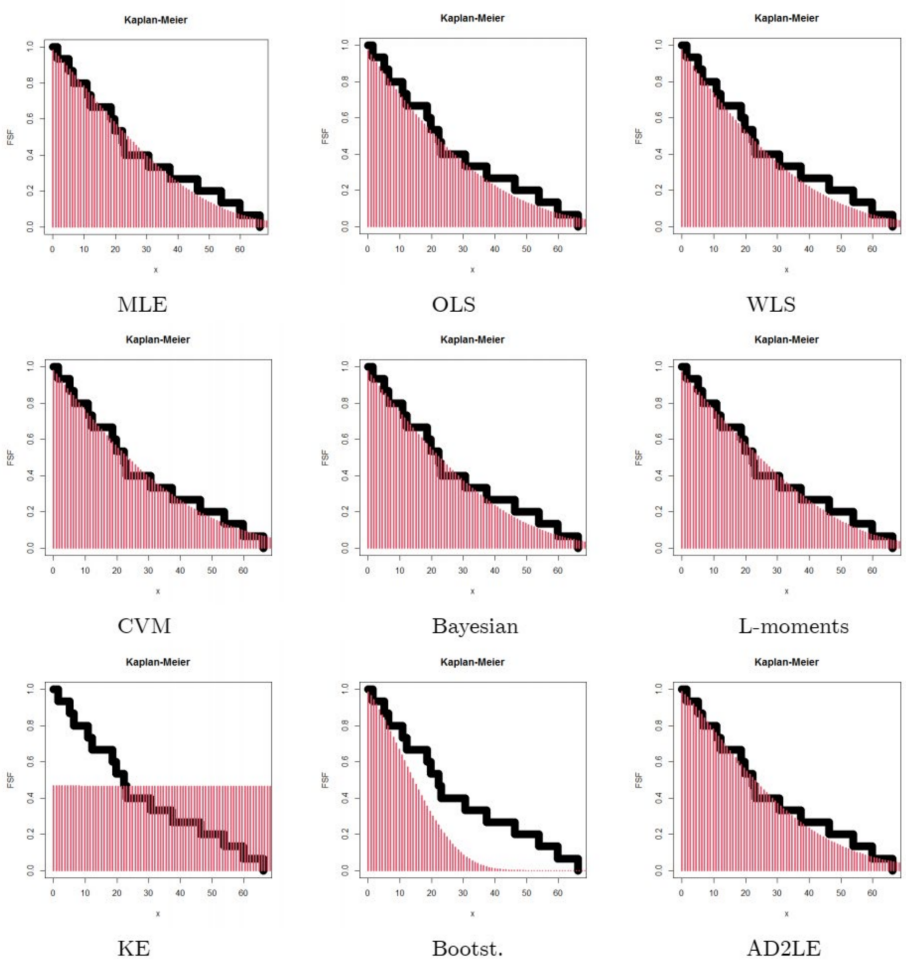

Figure 4. Kaplan–Meier plots under dataset II.

Table 9. Comparing methods using dataset II.

Method	p	γ	θ	K.S	P.V
MLE	0.2738836678	0.012957314	1.0860441145	0.11998	0.98219
OLS	0.1985257104	0.0168180005	0.9854823405	0.11603	0.98761
WLS	0.2398746891	0.0180291436	0.9944931716	0.11229	0.99153
CVM	0.1327154712	0.0103680076	1.0482619245	0.10679	0.99553
Bayesian	0.3038462428	0.0199252587	0.9928028081	0.09937	0.99843
L-mom	0.5901252129	0.0522561583	0.8582513340	0.12438	0.97446
KE	0.0000033106	0.0586353219	0.0001265448	0.53331	0.00039
Bootst	0.2244536022	0.0101742641	1.3246956780	0.31898	0.09447
* AD2LE	0.0070230912	0.0024977558	1.2480067543	0.09885	0.99855

The value with "*" is the best estimation in its row for all estimation methods.

6.3. Dataset III: Counts of Kidney Cysts

The data on kidney cyst counts reflect the number of cysts in lymphogenic kidneys caused by corticosteroids, which are linked to the expression of recognized cytogenic molecules and Indian hedgehog (see the works of Chan et al. [25], Eliwa et al. [17], Aboraya et al. [14], and Ibrahim et al. [15]). Table 10 gives the estimates, K–S, and P.V statistics for the kidney dataset. Based on Table 10, we can observe that the AD2LE method is the best, with K–S = 0.09885 and P.V = 0.99855, followed by the CVM method, with K–S = 0.28412 and P.V = 0.86757. Figure 5 gives the Kaplan–Meier plots using kidney data for all methods. The graphical results in Figure 5 confirm the results of Table 10. Moreover, the KE method provided undesirable results or unexpected results (K–S = 134.915 and P.V < 0.0001), and this may be due to the nature of the data used, or to any other random reasons. In any case, these results need further study and analysis, one way or another.

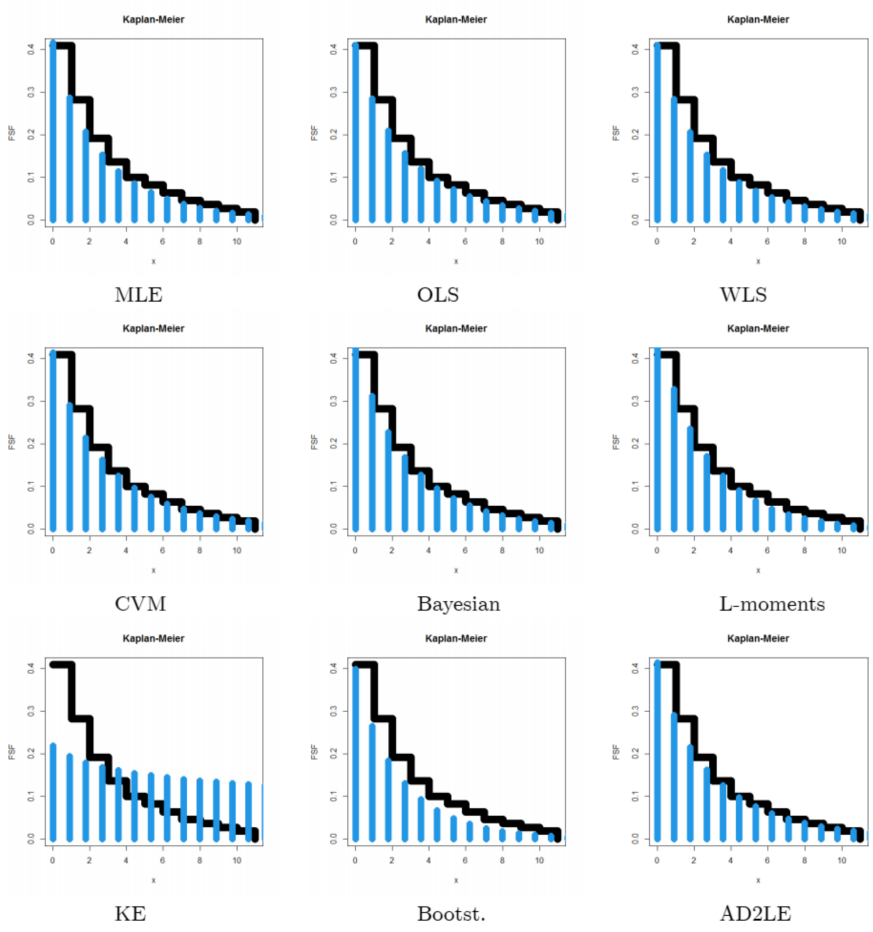

Figure 5. Kaplan–Meier plots according to dataset III.

Table 10. Comparing methods using dataset III.

Method	p	γ	θ	K.S	P.V
MLE	0.4459431905	0.7342772117	0.3800575939	0.35698	0.83653
OLS	0.3572219176	0.623814996	0.3861454942	0.33875	0.84419
WLS	0.3351236932	0.5963314357	0.3977530334	0.38168	0.82627
CVM	0.3559783727	0.616701488	0.3850433614	0.28412	0.86757
Bayesian	0.4334316130	0.6761278712	0.4033948113	0.61746	0.73438
L-mom	0.6508195705	1.0129827319	0.3676423109	1.69547	0.42838
KE	0.09858974996	0.5052793233	0.0933711548	134.915	<0.0001
Bootst	0.4502907161	0.7658835657	0.3892316532	1.15192	0.56216
* AD2LE	0.3560565879	0.6169232134	0.383446173	0.28323	0.86796

The value with "*" is the best estimation in its row for all estimation methods.

6.4. Dataset IV: Number of European Corn Borer Larvae

These data represent the number of European corn borer larvae in the field (see the works of Bebbington et al. [26], Eliwa et al. [17], and Aboraya et al. [14]). Table 11 gives the estimates, K–S, and P.V statistics for dataset IV. Based on Table 11, the L-mom method is the best, with K–S = 1.34090 and P.V = 0.51148, followed by the CVM method, with K–S = 2.21687 and P.V = 0.36774. Figure 6 gives the Kaplan–Meier plots using corn borer larvae data for all methods. The plots in Figure 6 confirm the results in Table 11. The KE and Bootst methods provided undesirable results or unexpected results (K–S = 8.69×10^6 and P.V < 0.0001), and this may be due to the nature of the data used, or to any other random reasons. In any case, these results need further study and analysis, one way or another.

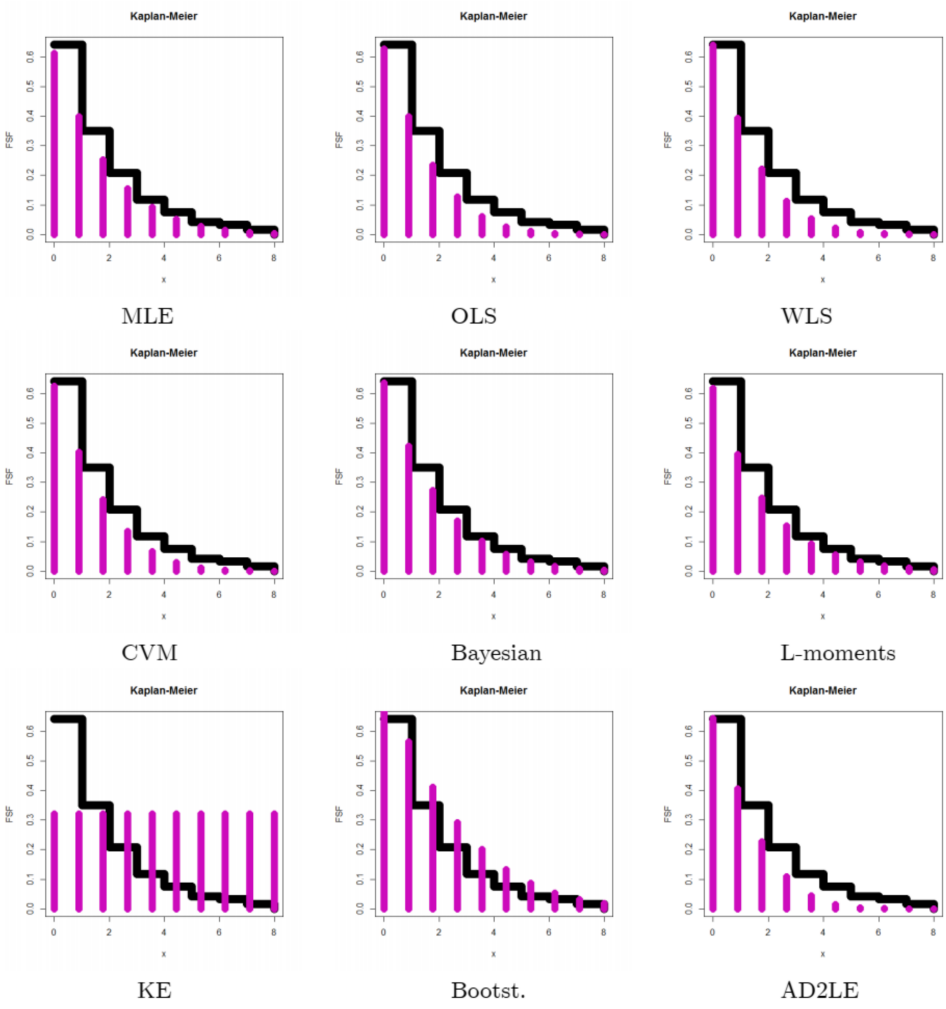

Figure 6. Kaplan–Meier plots for dataset IV.

Table 11. Comparing methods for dataset IV.

Method	p	γ	θ	K.S	P.V
MLE	0.0421915321	0.1441660755	0.8986850359	2.07337	0.35463
OLS	0.207451888	0.2600860974	0.9026929140	2.06132	0.35677
WLS	0.0328021927	0.1224019888	1.0741417897	2.57368	0.27614
CVM	0.2560193547	0.2955407600	0.8642332727	2.00076	0.36774
Bayesian	0.0388192223	0.1299953255	0.9335014673	2.21687	0.33007
* L-mom	0.0000162586	0.0427720724	1.0054904684	1.34090	0.51148
KE	0.0268390486	0.2732464266	7.22789×10^{-6}	8.69×10^6	<0.0001
Bootst	0.0305812278	0.0787811852	1.0371758772	117.888	<0.0001
AD2LE	0.4094146417	0.4016848541	0.8725794897	4.33318	0.11457

The value with "*" is the best estimation in its row for all estimation methods.

7. Competitive Models: Comparative Study and Interpretation

We used four real data applications to demonstrate the adaptability, usefulness, and significance of the DEGW distributions. The fitted distributions were analyzed and compared using the log-likelihood function, AICR, CAICR, χ^2_V with degree of freedom (d.f) P.V, and K–S and its P.V. Table 12 shows the competitive models and their abbreviations.

Table 12. The competitive models.

Discrete Model	Abbreviations
Exponential	DE
Inverse Rayleigh	DIR
Weibull	DW
Lindley-II	DLy-II
Rayleigh	DR
Inverse Weibull	DIW
Generalized exponential-II	DGE-II
Burr type XII	DBXII
Lindley	DLi
Log-logistic	DLL
Lomax	DLx
Poisson	Poisson
Exponentiated Lindley	EDLy
Pareto	DPa
Exponentiated Weibull	EDW
Negative binomial (see Dougherty [27])	NB

7.1. Dataset I: Failure Time Data of 50 Devices

Using the data of Bebbington et al. [26], we compared the fits of the DEGW distribution with some competitive discrete models, such as EDW, DW, DIW, DLy-II, EDLy, DLL, and DPa. The failure time data are shown in Figure 7 together with the quantile–quantile (Q-Q) plot (middle panel), boxplot (left panel), and total time on test (TTT) plot (right panel). Table 13 displays the MLEs and associated standard errors (St.Ers). Table 14 displays the goodness-of-fit test statistics. The MATHCAD application was used to generate the results for Table 13, Table 14, and all other comparable results in the following subsections. Based on Table 14, the DEGW provides the best fits against all competitive models, with $-l$ = 233.467, AICR = 472.933, CAICR = 473.455, K–S = 0.16304, and P.V = 0.15266. Figure 8 gives the fitted HRF (FHRF), fitted SF (FSF) (also called the Kaplan–Meier SF), and probability–probability (P–P) plots for failure time data. Based on Table 13, we have $E(Z) = 17.7667$, $V(Z) = 811.8649$, $S(Z) = 1.359128$, $K(Z) = 3.271034$, and Disp-Ix$(Z) = 45.6957$.

Table 13. The MLEs (and their corresponding St.Ers) for dataset I.

Model	p	γ	θ
DEGW	0.84504	0.12488	0.68408
	(0.1316)	(0.1337)	(0.1787)
EDW	0.98914	1.13934	0.78444
	(0.1644)	(3.2274)	(3.0535)
DW	0.98126	1.02342	
	(0.0114)	(0.1322)	
DIW	0.01832	0.58244	
	(0.0131)	(0.063)	
DLy-II	0.96934	0.0585	
	(0.00504)	(0.0274)	
EDLy	0.97222	0.48020	
	(0.0053)	(0.0873)	

Table 13. *Cont.*

Model	p	γ	θ
DLLc	1.00010 (0.3213)	0.43941 (0.0623)	
DPa	0.73922 (0.03212)		

Table 14. The goodness-of-fit test statistics for comparing the competitive models for dataset I.

	DEGW	EDW	DW	DIW	DLy-II	EDLy	DLLc	DPa
$-l$	233.47	240.24	241.65	261.94	240.67	240.38	294.99	275.99
AICR	472.93	486.76	487.25	527.89	485.23	484.69	593.84	553.74
CAICR	473.46	487.29	487.59	528.18	485.44	484.88	594.0	553.84
K–S	0.1630	0.1957	0.1872	0.2587	0.1868	0.1954	0.5354	0.3354
P.V	0.1527	0.0457	0.0619	0.0036	0.06499	0.045	<0.001	<0.0013

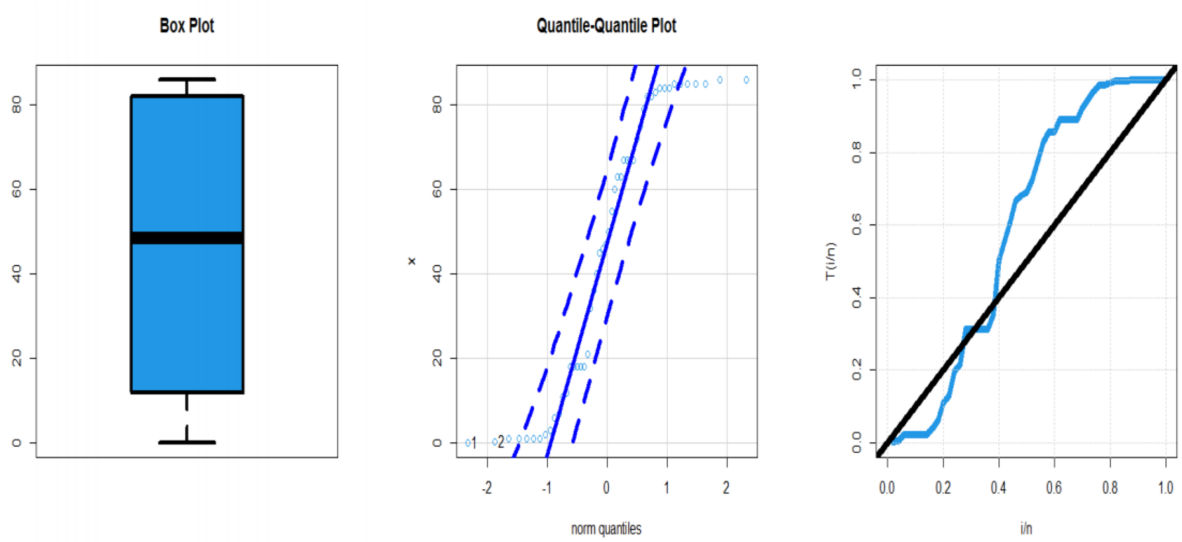

Figure 7. Box, Q-Q, and TTT plots for dataset I.

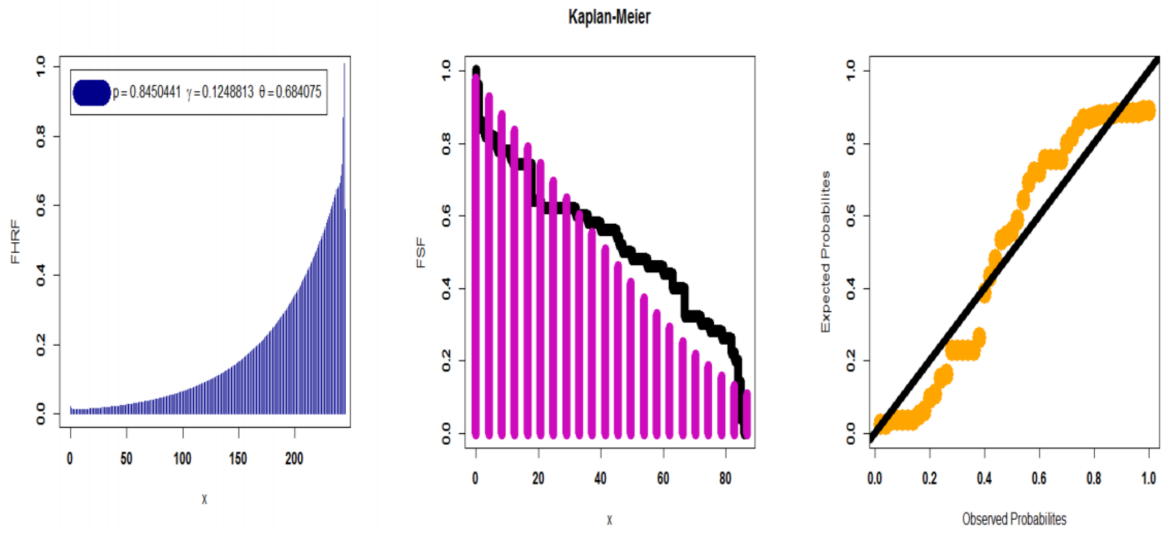

Figure 8. The FHRF, ESF, and P–P plots for dataset I.

7.2. Dataset II: Failure Times of 15 Electronical Components

We compared the DEGW distributions' fits to some models, including the DGE-II, DEx, DR, DIW, DIR, DLx, DPa, and DBXII, using data pertaining to electronic components. Table 15 exhibits both the MLEs and St.Ers. The test statistics are presented in Table 16. Based on Table 16, the DEGW provides the best fit compared to all discrete competitive models, with $-l$ = 63.791, AICR = 133.581, CAICR = 135.763, K–S = 0.11998, and P.V = 0.98219. Figure 9 gives the Q–Q plot (middle panel), boxplot (left panel), and TTT plot (right panel) for the data of the second failure times. Figure 10 gives the FHRF, ESF, and P–P plots for second failure times. Based on Table 15, we have $E(Z) = 4.873615$, $V(Z) = 133.6553$, $S(Z) = 2.890644$, $K(Z) = 11.67092$, and Disp-Ix$(Z) = 27.42426$.

Table 15. The MLEs (and their corresponding St.Ers) for dataset II.

Model	p	γ	θ
DEGW	0.27388	0.01296	1.08604
	(0.83537)	(0.0426)	(0.48154)
DGE-II	0.9561	1.4912	
	(0.0133)	(0.535)	
DIW	2.3×10^{-4}	0.8752	
	(7.8×10^{-4})	(0.164)	
DLx	0.0123	104.506	
	(0.039)	(84.409)	
DBXII	0.9753	13.367	
	(0.051)	(27.785)	
DR	0.9992		
	(2.58×10^{-4})		
DIR	1.832×10^{-7}		
	(0.055)		
DPa	0.7201		
	(0.061)		

Table 16. The goodness-of-fit test statistics for comparing the competitive models for dataset II.

	DEGW	DE	DGE-II	DR	DIR	DIW	DLo	DB-XII
$-l$	63.7911	65.002	64.423	66.390	89.0964	68.703	65.864	75.724
AICR	133.581	134.02	134.88	134.83	180.191	141.413	135.728	155.45
CAICR	135.763	136.34	135.81	136.11	180.501	142.412	136.728	156.45
K–S	0.11998	0.1777	0.1291	0.2161	0.6982	0.20923	0.20524	0.3889
P.V	0.98219	0.6734	0.9373	0.4333	<0.0001	0.4821	0.4912	0.0159

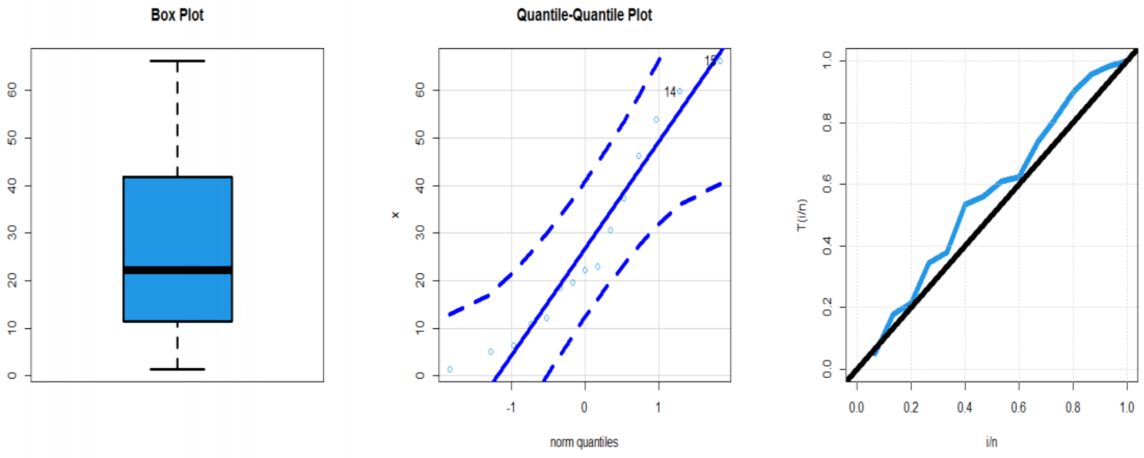

Figure 9. Boxplot, Q-Q plot, and TTT plot for dataset II.

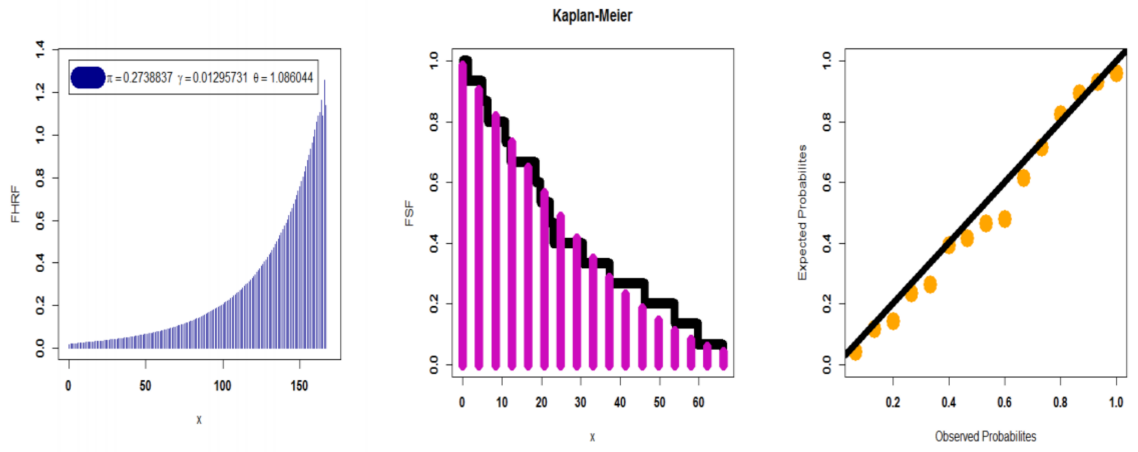

Figure 10. The FHRF, ESF, and P–P plots for dataset II.

7.3. Dataset III: Counts of Kidney Data

We compared the DEGW distribution's fits to some competing models, including the DW, DIW, DR, DE, DLi, DLy-II, DLx, and Poisson models, for this set of data. Table 17 shows the MLEs and St.Ers. The goodness-of-fit statistics are provided in Table 18. Based on Table 18, the DEGW provides the best fits against all competitive models, with $-l$ = 167.047, AICR = 340.094, CAICR = 340.321, χ^2_V = 0.35698, and P.V=0.83653. Figure 11 provides the boxplot, Q-Q plot, and TTT plot for the kidney data. Figure 12 gives the fitted PDF (FPMF), fitted SF (FSF), fitted HRF (FHRF), and fitted CDF (FCDF) plots. Based on Table 17, we have $E(Z) = 1.432338$, $V(Z) = 4.86933$, $S(Z) = 2.018928$, $K(Z) = 7.255506$, and Disp-Ix$(Z) = 3.399567$.

Table 17. The MLEs (and their corresponding St.Ers) for dataset III.

Model	p	γ	θ
DEGW	0.44594	0.73428	0.38006
	(0.9079)	(1.3168)	(0.2637)
DW	0.7503	0.43143	
	(0.084)	(0.3402)	
DIW	0.5813	1.0492	
	(0.0483)	(0.1463)	
DLy-II	0.5814	0.0011	
	(0.0455)	(0.058)	
DLx	0.1505	1.8303	
	(0.0980)	(0.9513)	
DR	0.90133		
	(0.0093)		
DE	0.5814		
	(0.0304)		
DLi	0.4363		
	(0.0263)		
Poi	1.3903		
	(0.1124)		

Table 18. The goodness-of-fit test statistics for comparing the competitive models for dataset III.

Z	OF	DEGW	DW	DIW	DR	DEx	DLi	DLy-II	DLx	Poi
0	65	64.163	59.01	63.91	11.00	46.09	40.25	46.03	61.89	27.42
1	14	15.626	19.84	20.70	26.83	26.78	29.83	26.77	21.01	38.08
2	10	9.184	10.78	8.050	29.55	15.56	18.36	15.57	9.650	26.47
3	6	6.038	6.260	4.230	22.23	9.040	10.35	9.050	5.240	12.26

Table 18. *Cont.*

Z	OF	DEGW	DW	DIW	DR	DEx	DLi	DLy-II	DLx	Poi
4	4	4.169	4.190	2.60	12.49	5.250	5.530	5.270	3.170	4.260
5	2	2.955	2.010	1.750	5.420	3.050	2.860	3.060	2.060	1.180
6	2	2.127	1.990	1.260	1.850	1.770	1.440	1.780	1.420	0.270
7	2	1.545	1.320	0.950	0.520	1.030	0.710	1.040	1.020	0.050
8	1	1.128	0.990	0.740	0.110	0.600	0.350	0.600	0.760	0.010
9	1	0.827	0.860	0.590	0.020	0.350	0.170	0.350	0.580	0.000
10	1	0.606	0.760	0.480	0.000	0.200	0.080	0.200	0.460	0.000
11	2	0.445	1.990	4.740	0.000	0.280	0.070	0.280	2.740	0.000
$-l$		167.047	170.14	172.93	277.78	178.77	189.1	178.8	170.48	246.21
AICR		340.094	344.28	349.87	557.56	359.53	380.2	361.5	344.96	494.42
CAICR		340.321	344.39	349.98	557.59	359.57	380.3	361.6	345.07	494.46
χ^2_V		0.35698	3.125	6.463	321.07	22.88	43.48	22.89	3.316	294.10
d.f		2	3	3	4	4	4	3	3	4
P.V		0.83653	0.373	0.091	<0.0001	0.0001	<0.0001	<0.0001	0.345	<0.0001

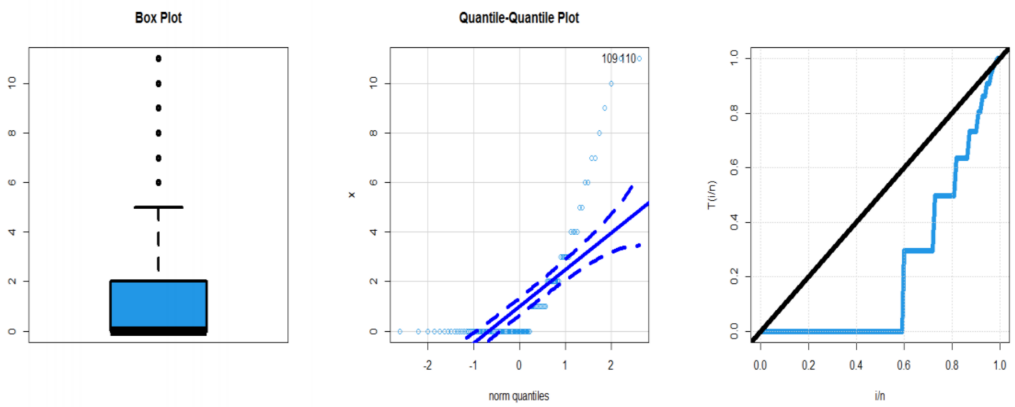

Figure 11. Boxplot, Q-Q plot, and TTT plot for dataset III.

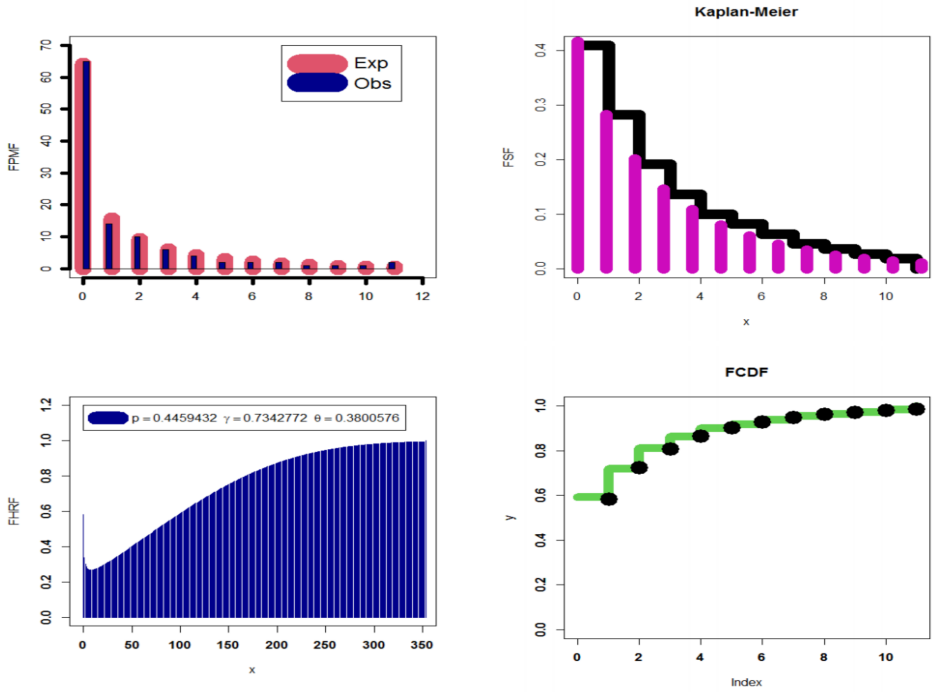

Figure 12. The FPMF, FSF, FHRF, and FCDF plots for dataset III.

7.4. Dataset IV: European Corn Borer Data

In this section, we compare the DEGW distributions' fits to those of other rival models, including the DGIW, DIW, DBXII, DIR, DR, NB, DPa, and Poisson model. Table 19 displays the MLEs and St.Ers. The goodness-of-fit statistics are shown in Table 20. Based on Table 20, the DEGW provides the best fits against all competitive models, with $-l = 200.956$, AICR = 407.912, CAICR = 408.119, χ^2_V = 2.07337, and P.V $= 0.35463$. Figure 13 gives the boxplot, Q-Q plot, and TTT plot. Figure 14 gives the FPMF, FSF, FHRF, and FCDF plots for corn borer larvae data. Based on Table 19, we have $E(z) = 1.44422$, $V(z) = 2.811083$, $S(z) = 1.329544$, $K(z)$ =4.446335, and Disp-Ix$(Z) = 1.946437$.

Table 19. The MLEs (and their corresponding St.Ers) for dataset IV.

Model	p	γ	θ	λ
DEGW	0.04222	0.14417	0.89869	
	(0.12268)	(0.13283)	(0.1634)	
DGW	0.04503	2.539324	2.1593	0.47933
	(0.4293)	(4.70345)	(2.6988)	(0.4655)
DIW	0.34523	1.54132		
	(0.0433)	(0.1564)		
DBXII	0.51933	2.35811		
	(0.0513)	(0.3663)		
NB	0.87013	9.95623		
	(0.0364)	(0.09623)		
DIR	0.31923			
	(0.0422)			
DR	0.86721			
	(0.01244)			
DPa	0.3299			
	(0.0344)			
Poi	1.48344			
	(0.0254)			

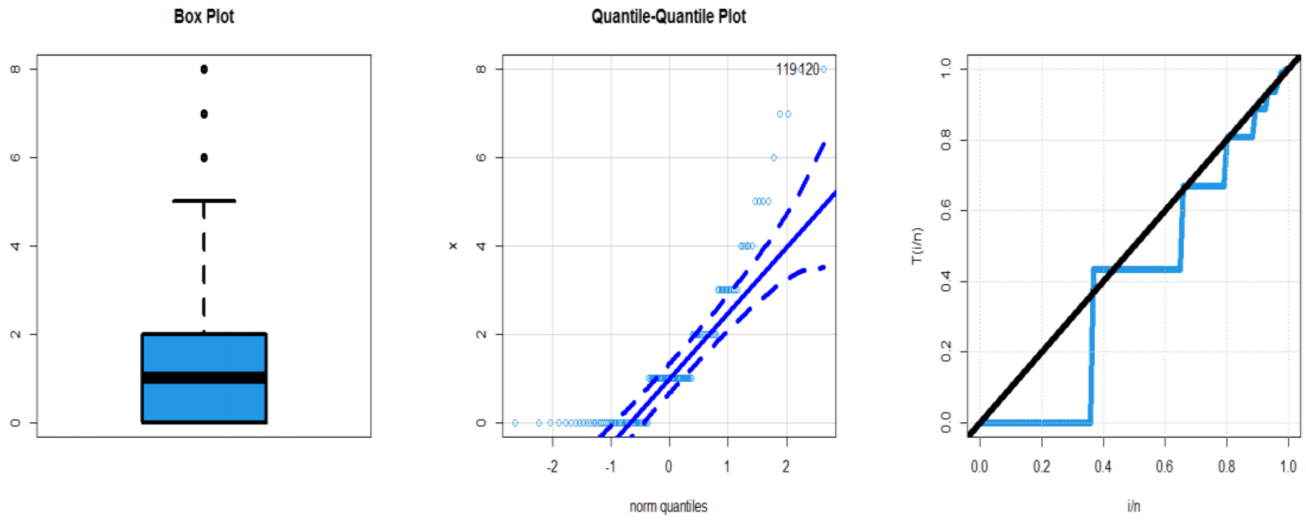

Figure 13. Boxplot, Q-Q plot, and TTT plot for dataset IV.

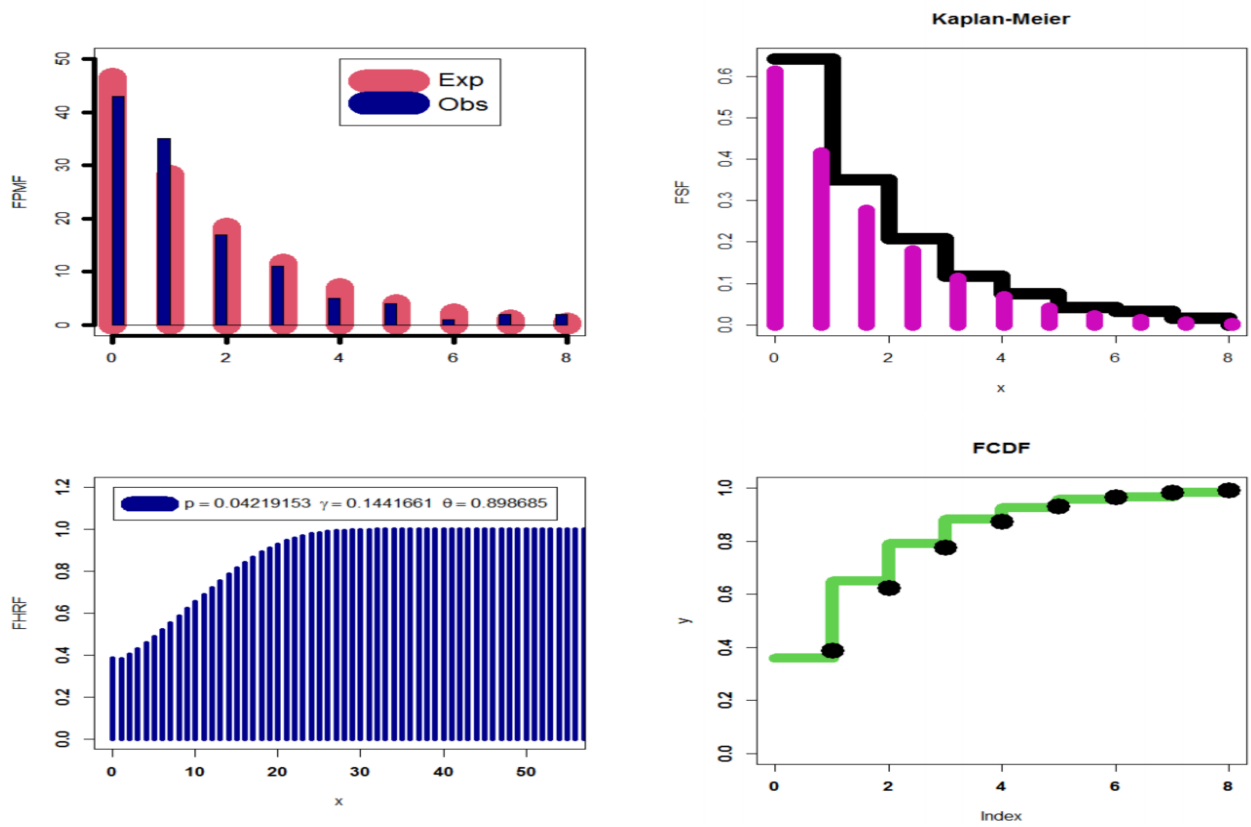

Figure 14. The FPMF, FSF, FHRF, and FCDF plots for dataset IV.

Table 20. The goodness-of-fit test statistics for comparing the competitive models for dataset IV.

Z	OF	DEGW	DIW	DBXII	DIR	DR	NB	DPa	Poi	DEGW
0	43	46.551	41.37	43.84	38.28	15.92	30.12	64.45	27.23	46.551
1	35	28.238	41.85	39.61	51.90	36.17	38.87	20.15	40.38	28.238
2	17	18.318	15.42	15.62	15.51	34.58	27.61	9.690	29.95	18.318
3	11	11.588	7.170	7.200	6.040	21.03	14.26	5.650	14.81	11.588
4	5	7.030	3.940	3.910	2.910	8.890	5.990	3.680	5.490	7.030
5	4	4.052	2.420	2.370	1.610	2.700	2.170	2.580	1.630	4.052
6	1	2.202	1.610	1.560	0.980	0.600	0.700	1.900	0.400	2.202
7	2	1.120	1.130	1.090	0.640	0.090	0.210	1.460	0.090	1.120
8	2	0.530	5.090	4.800	2.140	0.020	0.060	10.44	0.020	0.530
$-l$		200.956	204.810	204.293	208.440	235.23	211.52	220.63	219.19	200.956
AICR		407.912	413.621	412.587	418.881	472.45	427.05	443.24	440.38	407.912
CAICR		408.119	413.723	412.689	418.915	472.49	427.14	443.27	440.41	408.119
χ^2_V		2.07337	5.511	4.664	14.274	70.688	20.367	32.462	38.478	2.07337
d.f		2	3	3	4	4	3	4	4	2
P.V		340.321	344.39	349.98	557.59	359.57	380.3	361.6	345.07	494.46

8. Concluding Remarks

The discrete exponential generalized G (DEG-G) family is a new discrete variation of the exponential family that we suggest in this study. Numerous pertinent DEG-G family features—including the dispersion index, central and ordinary moments, cumulant-generating function, probability-generating function, and moment-generating function—were developed and studied, with numerical illustrations. After the new family was proposed, the DEGW model was introduced and studied in detail. The skewness $\in (-1.053185, \infty)$. The spread of its kurtosis was from 1.02442 to ∞. The dispersion index $\in (0, 1)$, or "> 1", or "=1". In order to simulate "under-dispersed", "equi-dispersed", or

"over-dispersed" count data, the DEGW distribution may be helpful. In addition to being "symmetric", "symmetric and bimodal", "uniform", or "right skewed with long tail", the probability mass function of the DEGW distribution can also be "asymmetric and right-skewed", "asymmetric and left-skewed", "symmetric", or "symmetric and bimodal". The DEGW distribution's failure rate can take one of five different forms: "constant," "growing-constant," "bathtub," "monotonically increasing", or "J-shapeed". The DEGW parameters were estimated via various techniques to determine the best estimator for each data. A thorough comparison of the various methodologies was conducted for both simulated and real-life data. Finally, four real-life datasets were analyzed, and the following results can be concluded:

- In modeling the asymmetric failure time count data (the data of 50 devices), it can be seen that the Bayesian method is the best method, with K–S = 0.14712 and P.V = 0.22927, followed by the MLE method, with K–S = 0.163038 and P.V = 0.15266. However, for this dataset, the discrete exponential generalized G family provides the best fit under the Weibull baseline, with $-l$ = 233.467, AICR = 472.933, CAICR = 473.455, K–S = 0.16304, and P.V = 0.15266.
- In modeling the asymmetric failure time count data (the data of 15 electronic components), the Anderson–Darling (left-tail second-order) method is the best method, with K–S = 0.09885 and P.V = 0.99855, followed by the Bayesian method, with K–S = 0.09937 and P.V = 0.99843. However, for this dataset, the discrete exponential generalized G family provides the best fit under the Weibull baseline, with $-l$ = 63.791, AICR = 133.581, CAICR = 135.763, K–S = 0.11998, and P.V = 0.98219.
- In modeling the asymmetric counts of kidney data, the Anderson–Darling (left-tail second-order) method is the best, with K–S = 0.09885 and P.V = 0.99855, followed by the Cramér–von Mises estimation method, with K–S = 0.28412 and P.V = 0.86757. However, for this dataset, the discrete exponential generalized G family provides the best fit under the Weibull baseline, with $-l$ = 167.047, AICR = 340.094, CAICR = 340.321, χ^2_V = 0.35698, and P.V = 0.83653.
- In modeling the asymmetric European corn borer larvae data, the L-moment method is the best, with K–S = 1.34090 and P.V = 0.51148, followed by the Cramér–von Mises estimation method, with K–S = 2.21687 and P.V = 0.36774. However, for this dataset, the discrete exponential generalized G family provides the best fit under the Weibull baseline, with $-l$ = 200.956, AICR = 407.912, CAICR = 408.119, χ^2_V = 2.07337, and P.V = 0.35463.

Discrete distributions still need more studies and applications, especially with regard to the statistical testing of hypotheses and validation, whether in the case of complete data or in the case of censored data. In this regard, the reader may find a guide in the works of Goual and Yousof [28], Yousof [29], Yadav et al. [30], Yadav et al. [31], and Mansour [32].

Author Contributions: M.S.E.: review and editing, validation, writing the original draft preparation, conceptualization, data curation, formal analysis, project administration, software; M.E.-M.: review and editing, methodology, conceptualization, software; H.M.Y.: review and editing, software, validation, writing the original draft preparation, conceptualization, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deputyship for Research and Innovation, Ministry of Education, Saudi Arabia, grant number QU-IF-2-5-3-25110.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Acknowledgments: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education, Saudi Arabia for funding this research work through the project number (QU-IF-2-5-3-25110). The authors also thank Qassim University for technical support.

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

Abbreviations

RV	Random variable
PMF	Probability mass function
CDF	Cumulative distribution function
EG-G	Exponential generalized G
DEG-G	Discrete exponential generalized G
exp-G	Exponential G
SF	Survival function
HRF	Hazard rate function
Disp-Ix	Dispersion index
MLEs	Maximum likelihood estimations
CVME	Cramér–von Mises estimation
OLSE	Ordinary least squares estimation
Bootst	Bootstrapping
KE	Kolmogorov estimation
WLSE	Weighted least squares estimation
AD2LE	Anderson–Darling method of left-tail second-order estimation
MSE	Mean square error
St.Ers	Standard errors
MCMC	Markov chain Monte Carlo
l	Log-likelihood
AICR	Akaike information criterion
CAICR	Consistent Akaike information criterion
BIC	Bayesian information criterion
HQIC	Hannan–Quinn information criterion
K–S	Kolmogorov–Smirnov
P.V	*p*-value
P–P	Probability–probability
TTT	Total time in test
Q-Q	Quantile–quantile
FHRF	Fitted hazard rate function
FSF	Fitted survival function

References

1. Nakagawa, T.; Osaki, S. The discrete Weibull distribution. *IEEE Trans. Reliab.* **1975**, *24*, 300–301. [CrossRef]
2. Roy, D. Discrete Rayleigh distribution. *IEEE Trans. Relib.* **2004**, *53*, 255–260. [CrossRef]
3. Jazi, A.M.; Lai, D.C.; Alamatsaz, H.M. Inverse Weibull distribution and estimation of its parameters. *Stat. Methodol.* **2010**, *7*, 121–132. [CrossRef]
4. Gomez-Déniz, E. Another generalization of the geometric distribution. *Test* **2010**, *19*, 399–415. [CrossRef]
5. Hesterberg, T. Bootstrap, Wiley Interdisciplinary Reviews: Computational Statistics. *WIREs Comp. Stat.* **2011**, *3*, 497–526. [CrossRef]
6. Nekoukhou, V.; Bidram, H. The exponentiated discrete Weibull distribution. *Stat. Oper. Res. Trans.* **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. Int. J.* **2016**, *4*, 1–15.
9. 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]
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]
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**, *81*, 135–162.
12. Chesneau, C.; Yousof, H.M.; Hamedani, G.; Ibrahim, M. A New One-parameter Discrete Distribution: The Discrete Inverse Burr Distribution: Characterizations, Properties Applications, Bayesian and Non-Bayesian Estimations. *Stat. Optim. Inf. Comput.* **2022**, *10*, 352–371. [CrossRef]
13. Bourguignon, M.; Silva, R.B.; Cordeiro, G.M. The Weibull-G family of probability distributions. *J. Data Sci.* **2014**, *12*, 53–68. [CrossRef]

14. Aboraya, M.M.; Yousof, H.; Hamedani, G.G.; Ibrahim, M. A new family of discrete distributions with mathematical properties, characterizations, Bayesian and non-Bayesian estimation methods. *Mathematics* **2020**, *8*, 1648. [CrossRef]
15. Ibrahim, M.; Ali, M.M.; Yousof, H.M. The discrete analog of the Weibull G family: Properties, different applications, Bayesian and non-Bayesian estimation methods. *Ann. Data Sci.* **2021**, 1–38. [CrossRef]
16. Steutel, F.W.; van Harn, K. *Infinite Divisibility of Probability Distributions on the Real Line*; Marcel Dekker: New York, NY, USA, 2004.
17. 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]
18. Yousof, H.M.; Majumder, M.; Jahanshahi, S.M.A.; Ali, M.M.; Hamedani, G.G. A new Weibull class of distributions: Theory, characterizations and applications. *J. Stat. Res. Iran JSRI* **2018**, *15*, 45–82. [CrossRef]
19. Kemp, A.W. Classes of discrete lifetime distributions. *Commun. Stat. Theor. Methods.* **2004**, *33*, 3069–3093. [CrossRef]
20. Poisson, S.D. *Probabilité des Jugements en Matiére Criminelle et en Matiére Civile, Précédées des RéGles Génerales du Calcul des Probabilitiés*; Bachelier: Paris, France, 1837; pp. 206–207.
21. Aguilar, G.A.; Moala, F.A.; Cordeiro, G.M. Zero-Truncated Poisson Exponentiated Gamma Distribution: Application and Estimation Methods. *J. Stat. Theory Pract.* **2019**, *13*, 1–20. [CrossRef]
22. Efron, B. The bootstrap and modern statistics. *J. Am. Stat. Assoc.* **2000**, *95*, 1293–1296. [CrossRef]
23. Bodhisuwan, W.; Sangpoom, S. The discrete weighted Lindley distribution. In Proceedings of the International Conference on Mathematics, Statistics, and Their Applications, Banda Aceh, Indonesia, 4–6 October 2016.
24. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*; Wiley: New York, NY, USA, 2003.
25. Chan, S.; Riley, P.R.; Price, K.L.; McElduff, F.; Winyard, P.J. Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as Indian hedgehog. *Am. J. Physiol.-Ren. Physiol.* **2009**, *298*, 346–356. [CrossRef] [PubMed]
26. Bebbington, M.; Lai, C.D.; Wellington, M.; Zitikis, R. The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure data. *Reliab. Eng. Syst. Saf.* **2012**, *106*, 37–44. [CrossRef]
27. Dougherty, E.R. *Probability and Statistics for the Engineering, Computing and Physical Sciences*; Prentice Hall: Englewood Cliffs, NJ, USA, 1992; pp. 149–152, ISBN 0-13-711995-X.
28. Goual, H.; Yousof, H.M. Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test. *J. Appl. Stat.* **2020**, *47*, 393–423. [CrossRef] [PubMed]
29. Yousof, H.M.; Ali, M.M.; Goual, H.; Ibrahim, M. A new reciprocal Rayleigh extension: Properties, copulas, different methods of estimation and modified right censored test for validation. *Stat. Transit. New Ser.* **2021**, *23*, 1–23. [CrossRef]
30. Yadav, A.S.; Goual, H.; Alotaibi, R.M.; Ali, M.M.; Yousof, H.M. Validation of the Topp-Leone-Lomax model via a modified Nikulin-Rao-Robson goodness-of-fit test with different methods of estimation. *Symmetry* **2020**, *12*, 57. [CrossRef]
31. Yadav, A.S.; Shukla, S.; Goual, H.; Saha, M.; Yousof, H.M. Validation of xgamma exponential model via Nikulin-Rao-Robson goodness-of- fit test under complete and censored sample with different methods of estimation. *Stat. Optim. Inf. Comput.* **2020**, *10*, 457–483. [CrossRef]
32. Mansour, M.; Rasekhi, M.; Ibrahim, M.; Aidi, K.; Yousof, H.M.; Elrazik, E.A. A New Parametric Life Distribution with Modified Bagdonavičius–Nikulin Goodness-of-Fit Test for Censored Validation, Properties, Applications, and Different Estimation Methods. *Entropy* **2020**, *22*, 592. [CrossRef]