Pak.j.stat.oper.res. Vol.20 No. 4 2024 pp 745-770 DOI: http://dx.doi.org/10.18187/pjsor.v20i4.4616

Pakistan Journal of Statistics and Operation Research

A New Discrete Generator with Mathematical Characterization, Properties, Count Statistical Modeling and Inference with Applications to Reliability, Medicine, Agriculture, and Biology Data

Haitham M. Yousof[1], Abdullah H. Al-Nefaie[2], Nadeem S. Butt[3], G.G. Hamedani[4], Hleil Alrweili[5*], Abdussalam Aljadani[6], Mahmoud M. Mansour[1,7], Mohamed S. Hamed[1,8] and Mohamed Ibrahim[2,9]

* Corresponding Author

[1]Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Benha University, Benha 13511, Egypt; haitham.yousof@fcom.bu.edu.eg
[2]Department of Quantitative Methods, college of Business, King Faisal University, Al Ahsa 31982, Saudi Arabia; aalnefaie@kfu.edu.sa
[3]Department of Family and Community Medicine, King Abdul Aziz University, Jeddah, Kingdom of Saudi Arabia; nshafique@kau.edu.sa
[4]Department of Mathematical and Statistical Sciences, Marquette University, USA; gholamhoss.hamedani@marquette.edu
[5]Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia; ∗Corresponding author: Hleil.Alrweili@nbu.edu.sa
[6]Department of Management, College of Business Administration in Yanbu, Taibah University, Al-Madinah, Al-Munawarah 41411, Kingdom of Saudi Arabia; ajadani@taibahu.edu.sa
[7]Department of Management Information Systems, Yanbu, Taibah University, Yanbu 46421, Saudi Arabia; mmmansour@taibahu.edu.sa
[8] Department of Business Administration, Gulf Colleges, KSA; mssh@gulf.edu.sa
[9]Department of Applied, Mathematical and Actuarial Statistics, Faculty of Commerce, Damietta University, Damietta, Egypt; mohamed_ibrahim@du.edu.eg & miahmed@kfu.edu.sa

Abstract

In this piece of work, we examine and present a completely new discrete family of distributions that we have created. Our investigation into the relevant mathematical properties and characterizations of the system makes use of both analytical and numerical methods. We focus on a particular member of this family so that we can study its theoretical foundations as well as its graphical and numerical representations. This new model contains a few different hazard rate functions, some of which are referred to as "increasing constant", "decreasing-constant-increasing (U)", "constant", "U-constant", "decreasing", and "J-shape" In a similar vein, the model's probability mass function provides a variety of forms, all of which are helpful and practical. These forms include "asymmetric left skewed," "right skewed with wide peak," "right skewed," "bimodal," "symmetric," and "right skewed," amongst others. Each of these forms is valuable and applicable in their own way. These forms might be discovered in the probability mass function that the model generates. In this investigation, in addition to the Bayesian estimating technique under the traditional loss function of squared errors, we investigate and make use of a total of eight estimate strategies that are not founded on Bayesian theory (classical methods). Simulations employing the Markov Chain Monte-Carlo method are run for comparing the Bayesian way of estimation with the more traditional approach of estimating values. According to the findings that we've compiled, the estimation strategy that is referred to as maximum likelihood yields the most accurate results across the board and for all different types of sample sizes. In addition, we evaluate and contrast the various methods of estimation by making use of six distinct real dataset sets; this indicates the versatility of the unique model that we have developed.

Key Words: Discrete Models; Bayesian estimation; Bootstrapping; Cramér-von-Mises; Discretization; Gibbs sampler; Metropolis-Hastings; Markov Chain Monte Carlo; Maximum Likelihood; Statistical Modeling.

1. Introduction

The discrete probabilistic distributions are fundamental tools in many areas of applied mathematics, engineering, computer science, and statistics. These distributions allow us to model a wide range of phenomena, such as the

outcomes of random experiments, the behavior of complex systems, and the uncertainty associated with dataset analysis. Creating new families of discrete probabilistic distributions that may accurately represent the traits of more complex and varied systems has attracted increasing interest in recent years. This interest stems from the realization that many real-life problems require more flexible and versatile models than those provided by classical probabilistic models such as the Bernoulli, the geometric, the Poisson, the binomial and the negative binomial distributions. The development of the new discretized probabilistic distributions is often motivated by the need to simplify or improve the estimation process of the underlying parameters, for more details and applications about the zero inflated dataset see Bahrami et al. (2012), Bahrami and Ganjali (2014) and Razie et al. (2016). In many applications of dataset analysis and modelling, this is a stage that plays a highly crucial role, and there are several reasons for this, including the following:

One method for achieving these aims is to discretize continuous variables, which will simplify the dataset, making it much simpler to interpret and much easier to analyze. The statistical technique of discretization makes it possible to find non-linear relationships between various variables. Discretizing either one of the variables or both of them can reveal a relationship that is simpler to define, even if just one of the variables is discretized, this may still be the case. The dimensionality of the dataset can be decreased with the assistance of discretization, which then results in the dataset being easier to see and evaluate. The discretization method can also help in handing outliers by dividing and hence dividing them into sone sub-intervals. One reason why new discrete families of probabilistic distributions are needed is that many phenomena exhibit non-standard behaviors that cannot be accurately modeled by existing distributions. For example, in social network analysis, the degree distribution of the nodes often follows a power law, which cannot be captured by classical distributions. In computational biology, the distribution of the number of mutations in a DNA sequence may exhibit long tails, which cannot be modeled by the Poisson distribution. Another reason why new discrete families of probabilistic distributions are needed is that many applications require distributions with specific properties or features. For instance, some applications may require distributions that are unimodal, while others may require distributions that are multimodal. Some applications may require distributions that are robust to outliers, while others may require distributions that can capture extreme events. Furthermore, new discrete families of probabilistic distributions can provide a more efficient and accurate way of modeling and analyzing dataset. For example, some new families of distributions may require fewer parameters to fit the dataset or have more efficient algorithms for inference and simulation. The new families of distribution can also lead to new insights and discoveries in various areas of science and engineering.

Recent research has led to the development of a novel strategy for the generation of a fresh set of discrete distribution families. The continuous Gompertz-G (Gz-G) generator was defined and studied by Alizadeh et al. (2017). The continuous Gz-G generator had the following survival function (SF):

$$S_{\beta_1,\beta_2}(y) = exp\left(\frac{\beta_1}{\beta_2}\left\{1-\left[1-G_{\underline{\boldsymbol{W}}}(y)\right]^{-\beta_2}\right\}\right)|_{y\in R,\beta_1,\beta_2>0}.$$

where $1-G_{\underline{\boldsymbol{W}}}(y)$ refers to the SF of the baseline model and $\underline{\boldsymbol{W}}$ refers to the parameter vector of the baseline model. Following Alizadeh et al. (2017) and Aboraya et al. (2020) defined and studied the discrete Rayleigh-G (DR-G) family of distributions based on the continuous Weibull-G family of Bourguignon et al. (2014). In this context, some well-known examples can be given of some of the distributions that were presented as discrete distributions, see, for example, Nakagawa and Osaki (1975) (for a simple discrete version of the Weibull (DW)); Roy (2004) (for a simple discrete version of the Rayleigh (DR)); Gomez-Deniz (2010) (for a simple discrete version of the exponential (DE)); Jazi et al. (2010) (for a simple discrete version of the inverted Weibull (DINW)); Gommez-Deniz and Calderin-Ojeda (2011) (for a simple discrete version of the Lindley (DLN)); Hussaine et al. (2016) (for a simple discrete version of the Lindley type II (DLN-II)); Para and Jan (2016) (for a simple discrete version of the loglogistic (DLL)); Para and Jan (2016) (for a simple discrete version of the Lomax (DLX)); Para and Jan (2016) (for a simple discrete version of the Burr XII (DBTXII)); Krishna and Pundir (2009) (for a simple discrete version of the Pareto (DPAR)); among others, and the contributions of scholars in this field are limited to these two main directions.

Based on the Weibull distribution and the generated odd ration argument $\nabla\left(y|\beta_1,\beta_2,\underline{\boldsymbol{W}}\right)$ where

$$\nabla\left(y|\beta_1,\beta_2,\underline{\boldsymbol{W}}\right) = \left[\dot{G}_{\underline{\boldsymbol{W}}}^{-\beta_2}(y)-1\right]^{\beta_1}|_{y\in R,\beta_1,\beta_2>0},$$

where the standard SF of any baseline model if referred by $\dot{G}_{\underline{\boldsymbol{W}}}(y) = 1-G_{\underline{\boldsymbol{W}}}(y)$. Recently, Yousof et al. (2018) investigated and analyzed a novel probabilistic continuous class called the Weibull generated G (WG-G) family can be expressed as

$$\mathcal{F}_{\alpha,\beta_1,\beta_2}(y) = 1-exp\left[-\alpha\nabla\left(y|\beta_1,\beta_2,\underline{\boldsymbol{W}}\right)\right] \quad |_{(y\in R \text{ and } \alpha,\beta_1,\beta_2>0)}, \tag{1}$$

where α represents the scale parameter of the model and β_1, β_2 represents two additional shape parameters. However other useful models can be chosen for discretization like Alizadeh et al. (2018) and Nadarajah et al. (2018), among others. In this work, following Yousof et al. (2018), we propose and study a new discrete analogue of the WG-G family called the discrete Weibull Generated-G (DWG-G) family. Some simple special members based on the standard Weibull (W), the standard inverted Weibull (INW), the standard Lomax (Lx), the standard inverted Lomax (INLx), the standard Burr X (BX), the standard inverted Burr X (INBX), the standard loglogistic (LL), the standard inverted loglogistic (INLL), the standard Rayleigh (R), the standard inverted Rayleigh (R), the standard exponential (E), the standard inverted exponential (INE), the standard Lindley (Li) and the standard inverted Lindley (INLi) distributions are discussed theoretically, graphically, and numerically.

Numerous important features are derived, including central moments, cumulant generating functions, probability generating functions, ordinary moments, moment generating functions, and dispersion index (DisIx). Cramer-von-Mises Estimation (CRVME), The ordinary version of the least-square estimation (ORLSE) method, estimation Bootstrapping method, maximum of the likelihood estimation (MLKE), estimation method of Kolmogorov (KE), the weighted version of least-square estimation (WLSE), method of L-moments, and Anderson Darling 2LD (Left-Tail Second Order) Methods are a few non-Bayesian estimation techniques that are taken into consideration. The DWG-G family provided a better fit than many well-known competing models according to some criterion like the Akaike information criterion (AIC), the Akaike information consistency criterion (CAIC), Chi-square ($\chi^2_{\tilde{W}}$) (P-value (p_v)), Kolmogorov-Smirnov (KG-SM) and its corresponding p_v. The following is a discussion of the considerations that went into our decision to launch the DGzR-G family:

I. Constructing a novel probability mass function (PMF) that is capable of taking on a variety of useful PMF shapes, including "asymmetric left skewed," "right skewed with wide peak," "right skewed," "bimodal PMF," "symmetric PMF," and "right skewed." Several innovative models are presented, each of which has a different hazard rate form, such as a "upside down failure rate," a "monotonically decreasing failure rate," a "bathtub failure rate," a "monotonically growing failure rate," a "decreasing-constant failure rate," and a "constant failure rate with one value.
II. In order to accurately describe "over-dispersed," "equal-dispersed," and "under-dispersed" real dataset, a number of novel discrete models have been proposed. It has been shown that the DWG-G family is more effective at modeling many various kinds of dataset, regardless of whether the dataset is symmetric or asymmetric or if they contain outliers. This is the case regardless of whether the dataset contains outliers. This is still the case even when the dataset is modeled using values at the extreme end of the scale.
III. Because the results that it produced in the statistical modeling of the dataset were so impressive, it is strongly advised that the creative family be utilized in the examination of the bathtub hazard rate count dataset. The results that it created were amazing. If the dataset is evaluated correctly, the same core principle can also be used to correctly explain dataset suggesting a gradually increasing number of failures; however, this requires the dataset to be interpreted correctly.
IV. The novel G class provided a viable alternative for reproducing zero-inflated dataset (agricultural real-life dataset) along with some outliers and a failure rate that was decreasing, increasing, and dropping over time.
V. Analyzing the differences between the estimate techniques used for the simulated dataset and the count, the zero-inflated dataset sets in order to recommend and suggest the most effective course of action for each specific scenario.
VI. The new family, which has a dropping failure rate and only a few outlier observations, can be recommended as a plausible statistical choice for managing zero-inflated real-life and count medical datasets. This is because the failure rate has been going down. This is due to the fact that the number of extreme cases has dropped.
VII. Actually, we demonstrate that the suggested G family of distributions gives a superior fit for six real dataset sets by making use of an experimental method in comparison many extended competing distributions and relevant versions that contain three and four parameters. These distributions are in competition with each other.
VIII. The new family can make discrete distributions with two modes. Bimodal distributions can be used to model things that are too complicated for a unimodal distribution to properly describe. In the area of genetics, for example, bimodal distributions can be used to model how genes with two different ways of being expressed are expressed. Bimodal distributions can help people understand the dataset they are looking at better. By figuring out what the two modes are, researchers can learn more about what causes the two peaks and how they might be linked. Bimodal distributions can also help find outliers or extreme numbers that might be hidden in a unimodal distribution. By finding the two modes, researchers can easily find numbers that aren't

in the normal range for each mode. Bimodal distributions can help you decide what to do when there are more than two possible outcomes. In economics, for example, a bimodal distribution can help people make decisions by showing how income or wealth is spread out and what the effects of policy changes might be on different income groups.

IX. The new family can produce discrete distributions with heavy tails. When modeling unusual events that are poorly captured by the traditional discrete distributions, heavy-tailed distributions are helpful. For instance, heavy-tailed distributions in finance can be used to simulate severe occurrences like market crashes or sharp price changes. Compared to existing distributions, heavy-tailed distributions are more resistant to outliers. This is due to the fact that they give extreme events a probability that is not zero, which can assist limit the impact of extreme outliers on the distribution as a whole. In comparison to current distributions, heavy-tailed distributions offer more precise estimates of the tail probability. This is significant for applications like risk management where precise tail probability estimations are essential.

2. The new family

With the use of the discretization principle, the recently developed CDF for the new family can be stated as follows

$$\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y) = 1 - \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{W}})} \quad |_{(\pi\in I=(0,1) \text{ and } y\in \mathrm{N}_0)}, \tag{2}$$

where $\underline{\boldsymbol{\Psi}} = \pi, \beta_1, \beta_2, \underline{\boldsymbol{W}}$, $\alpha = -\log(\pi)$ and $N_0 = N \cup \{0\}$. The following is one formulation of the relevant SF that might be possible:

$$S_{\underline{\boldsymbol{\Psi}}}(y) = \dot{\mathcal{F}}_{\underline{\boldsymbol{\Psi}}}(y) = \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{W}})} \quad |_{(\pi\in I \text{ and } y\in \mathrm{N}_0)}. \tag{3}$$

Then, Due to Kemp (2004), the PMF ($P_{\underline{\boldsymbol{\Psi}}}(y)$) of the DWG-G family is

$$P_{\underline{\boldsymbol{\Psi}}}(y) = \pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{W}})} - \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{W}})} \quad |_{(\pi\in I \text{ and } y\in \mathrm{N}_0)}. \tag{4}$$

For $\beta = 2$, the novel family reduces to discrete Rayleigh G (DRG) family. For $\beta_1 = 1$, the novel family reduces to discrete exponential G (DEG) family. In addition to this, the new hazard rate function (HRF) that corresponds to the DWG-G family can be stated as

$$H_{\underline{\boldsymbol{\Psi}}}(y) = P_{\underline{\boldsymbol{\Psi}}}(y)/\dot{\mathcal{F}}_{\underline{\boldsymbol{\Psi}}}(y-1) = \frac{\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{W}})} - \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{W}})}}{\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{W}})}}.$$

For many innovative models, ensuring identifiability is difficult, both theoretically and practically. To demonstrate this characteristic as effectively as feasible. For more information on the identifiability of model parameters, please see Tabrizi et al. (2020). Several examples of the DWG-G family are presented below for your review. The PMF (4) will be most tractable when the baseline model CDF $G_{\underline{\boldsymbol{W}}}(y)$ has a simple analytic expression. Here in Table 1, we provide some sub-models of this family corresponding to the baseline CDFs W, INW, Lx, INLx, BX, INBX, LL, INLL, R, IR, E, IE, Li and INLi distributions.

Table 1: A selection of new models established based on the new DWG-G family.

Model	$G_{\underline{\boldsymbol{W}}}(y)$	$\nabla_{\underline{\boldsymbol{W}}}^{\beta_1,\beta_2}(y)\|_{y\in N_0}$	The new model
W	$1-exp(-y^{\theta})$	$\left[exp(\beta_2 y^{\theta})-1\right]^{\beta_1}\Big\|_{\beta_1,\beta_2,\theta>0}$	DWGW
INW	$exp(-y^{\theta})$	$\left\{\left[1-exp(-y^{-\theta})\right]^{\beta_2}-1\right\}^{\beta_1}\Big\|_{\beta_1,\beta_2,\theta>0}$	DWGINW
R	$1-exp(-y^2)$	$[exp(\beta_2 y^2)-1]^{\beta_1}\Big\|_{\beta_1,\beta_2>0}$	DWGR
IR	$exp(-y^2)$	$\left\{[1-exp(-y^{-2})]^{\beta_2}-1\right\}^{\beta_1}\Big\|_{\beta_1,\beta_2>0}$	DWGIR
E	$1-exp(-y)$	$[exp(\beta_2 y)-1]^{\beta_1}\Big\|_{\beta_1,\beta_2>0}$	DWGE
IE	$exp(-y^{-1})$	$\left\{[1-exp(-y^{-1})]^{\beta_2}-1\right\}^{\beta_1}\Big\|_{\beta_1,\beta_2>0}$	DWGIR
Lx	$1-(1+y)^{-\theta}$	$\left[(1+y)^{\beta_2\theta}-1\right]^{\beta_1}\Big\|_{\beta_1,\beta_2,\theta>0}$	DWGLx
INLx	$(1+y^{-1})^{-\theta}$	$\left\{\left[1-(1+y^{-1})^{-\theta}\right]^{-\beta_2}-1\right\}^{\beta_1}\Big\|_{\beta_1,\beta_2,\theta>0}$	DWGINLx
BX	$[1-exp(-y^2)]^{\theta}$	$\left(\left\{1-[1-exp(-y^2)]^{\theta}\right\}^{-\beta_2}-1\right)\Big\|_{\beta_2,\theta>0}$	DWGBX
INBX	1-$[1-exp(-y^2)]^{\theta}$	$\left([1-exp(-y^{-2})]^{-\theta\beta_2}-1\right)\Big\|_{\beta_2,\theta>0}$	DWGINBX
LL	$1-(1+y)^{-1}$	$\left[(1+y)^{\beta_2}-1\right]\Big\|_{\beta_2>0}$	DWGLL
INLL	$(y^{-1}+1)^{-1}$	$\left\{[1-(1+y^{-1})^{-1}]^{-\beta_2}-1\right\}^{\beta_1}\Big\|_{\beta_1,\beta_2>0}$	DWGINLL
Li	$1-\frac{\theta+1+\theta y}{\theta+1}exp(-\theta y)$	$\left[\left(\frac{\theta+1+\theta y}{\theta+1}\right)^{-\beta_2}exp(\beta_2\theta y)-1\right]^{\beta_1}\Big\|_{\beta_1,\beta_2,\theta>0}$	DWGLi

ILi	$\left[1+\frac{\theta}{(\theta+1)y}\right]exp(-\theta y^{-1})$	$\left(\left\{1-\left[1+\frac{\theta}{(\theta+1)y}\right]exp(-\theta y^{-1})\right\}^{-\beta_2}-1\right)^{\beta_1}\Big\vert_{\beta_1,\beta_2,\theta>0}$	DWGILi

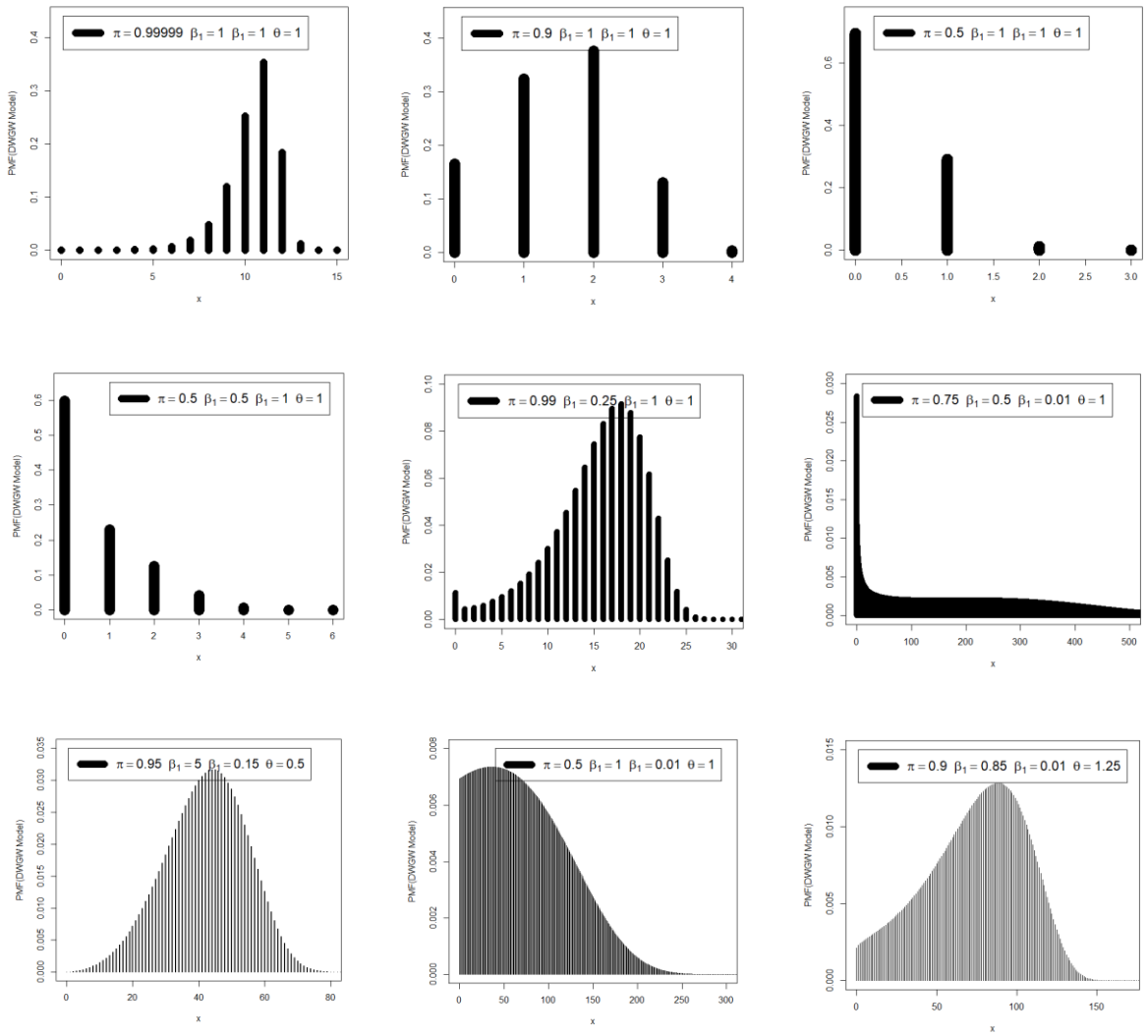
Consider the standard Weibull probabilistic distribution, then, we have

$$\nabla(y|\beta_1,\beta_2,\theta)=\left[exp\left(\beta_2 y^{\theta}\right)-1\right]^{\beta_1}|_{\beta_1,\beta_2,\theta>0}.$$

Then, depending on (3), the novel PMF of the new DWGW model is

$$P_{\underline{\Psi}}(y)=\pi^{\left[exp\left(\beta_2 y^{\theta}\right)-1\right]^{\beta_1}}-\pi^{\left\{exp\left[\beta_2(y+1)^{\theta}\right]-1\right\}^{\beta_1}}\ |_{(y\in N_0,\pi\in I \text{ and } \beta_1.\beta_2,\theta>0)}.$$

Clearly, when $\theta=2$, the DWGW probabilistic distribution reduces to the DWGR probabilistic distribution and when $\theta=1$ it reduces to the DWGE probabilistic distribution. The novel PMF of the DWGW probabilistic distribution is presented in Figure 1 for a few parameter values that were picked at random. In Figure 2, different graphs of the HRF of the DWGW probabilistic distribution are displayed for a few parameter values that have been selected.



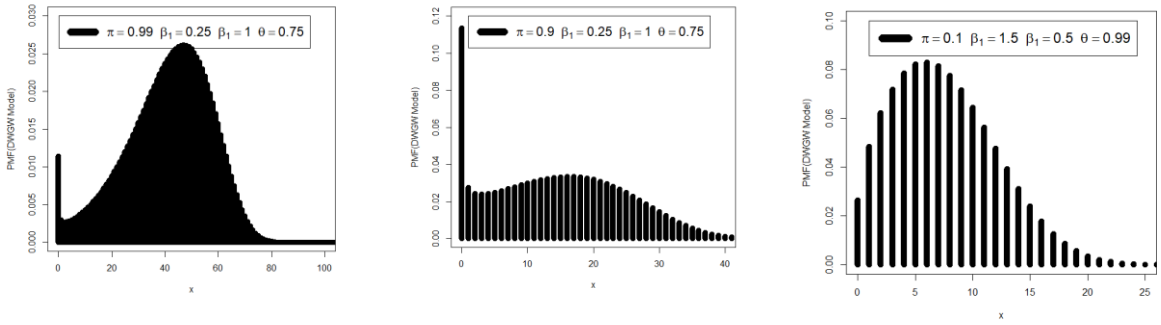
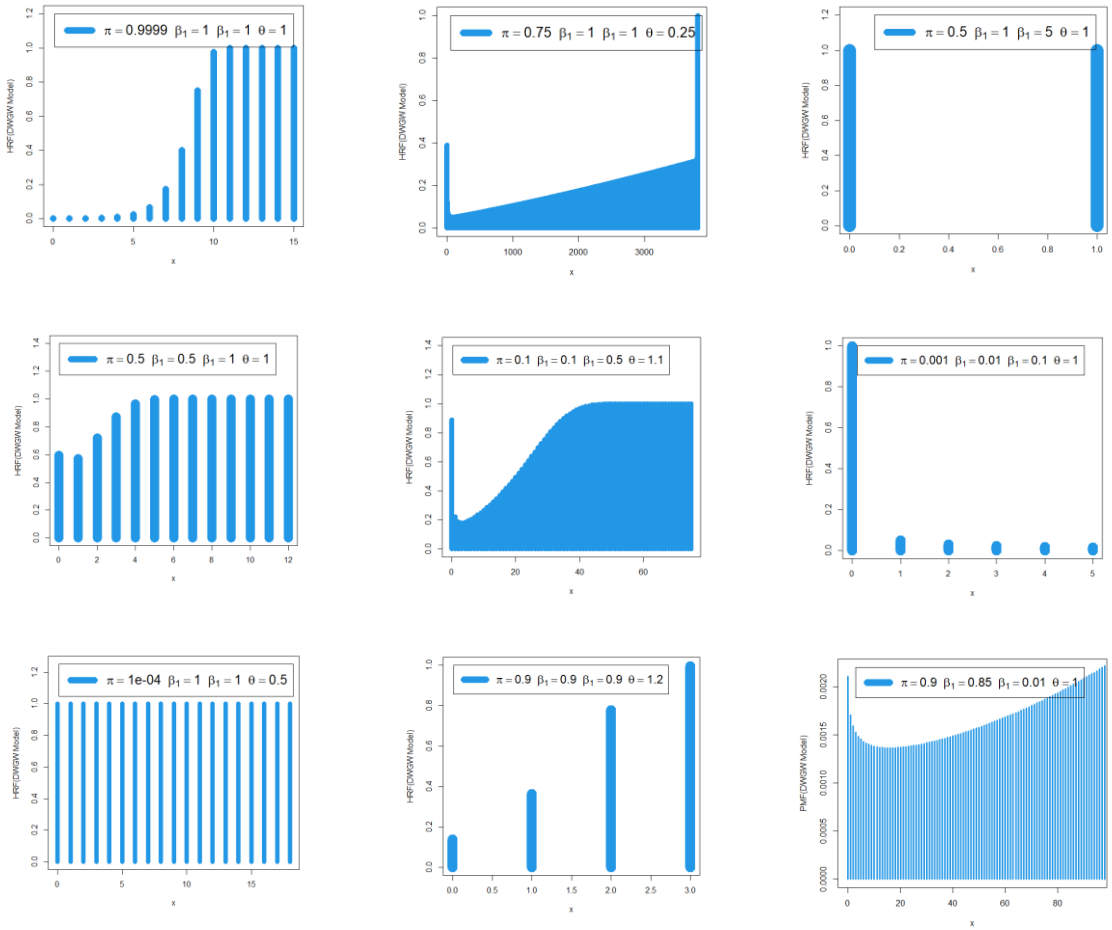


Figure 1: The PMF of the DWGW for a number of well chosen parameter values.



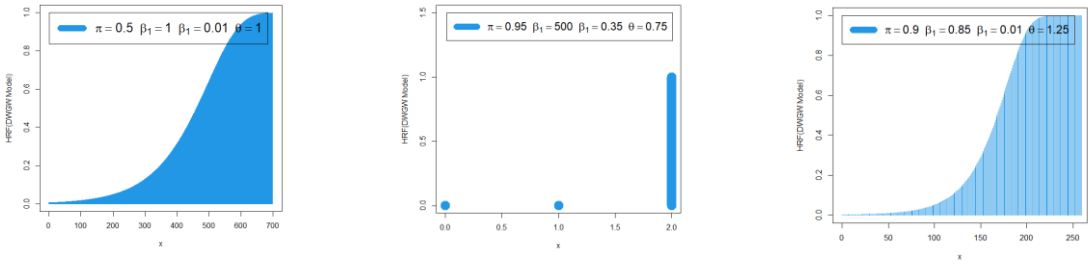


Figure 2: The HRF of the DWGW for a number of well-chosen parameter values.

From Figure 1, we can determine that the PMF of the DWGW probabilistic distribution can take on a number of advantageous PMF shapes. These shapes include "asymmetric left skewed," "right skewed with wide peak," "right skewed," "bimodal PMF," "symmetric PMF," and "right skewed." Other useful PMF shapes include "symmetric PMF" and "right skewed." We can see from Figure 2 that the HRF of the DWGW can be any one of the following: "increasing-constant," "decreasing-constant-increasing (U-HRF)," "constant," "U- constant," "decreasing," and "J-HRF."

3. Major mathematical properties

3.1 The ordinary moments

Theorem 1:

Let $\boldsymbol{y}$ be nonnegative random variable (NN-RV), and suppose that $\boldsymbol{y} \sim$ DWG-G ($\underline{\boldsymbol{\Psi}}$) family, then the r^{th} ordinary moment of $\boldsymbol{y}$ can be calculated from:

$$\mu'_{r,\boldsymbol{y}} = E(\boldsymbol{y}^r) = \sum_{y=1}^{\infty} \nabla\left(y|\beta_1,\beta_2,\underline{\boldsymbol{w}}\right)\left[y^r - (y-1)^r\right]|_{(y\in N_0,\pi\in I.)}.$$

Proof: Since

$$\mu'_{r,\boldsymbol{y}} = E(y^r) = \sum_{y=0}^{\infty} y^r\, P_{\underline{\boldsymbol{y}}}(y).$$

Then,

$$\mu'_{r,\boldsymbol{y}} = \sum_{y=0}^{\infty} y^r \begin{bmatrix} \pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} \\ -\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} \end{bmatrix} = \sum_{y=1}^{\infty} \begin{bmatrix} -(y-1)^r \\ +y^r \end{bmatrix} \dot{\mathcal{F}}_{\underline{\boldsymbol{y}}}(y-1) = \sum_{y=1}^{\infty} \pi^{\nabla_{\underline{\boldsymbol{w}}}^{\beta_1,\beta_2}(y)} \begin{bmatrix} -(y-1)^r \\ +y^r \end{bmatrix}. \quad (5)$$

Then, using (5), the mean, $\mu'_{1,\boldsymbol{y}}$ and the second moment can be respectively written as

$$E(\boldsymbol{y}) = \sum_{y=1}^{\infty} \pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}, \text{ and } \mu'_{2,\boldsymbol{y}} = E(\boldsymbol{y}^2) = \sum_{y=1}^{\infty} (2y-1)\, \pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})},$$

3.2 Central moment and DisIx

The r[th] central moment of $\boldsymbol{y}$, say μ_r, follows as

$$\mu_{r,\boldsymbol{y}} = E(y - \mu'_{r,\boldsymbol{y}})^r = \sum_{w=0}^{r} \left(-\mu'_{1,\boldsymbol{y}}\right)^w \binom{r}{w} \mu'_{(r,\boldsymbol{y})-w}.$$

Therefore, the variance (V($\boldsymbol{y}$)) can be obtained from

$$\mathrm{V}(\boldsymbol{y}) = E(\boldsymbol{y} - \mu'_{r,\boldsymbol{y}})^2 = \mu \sum_{w=0}^{2} (-\mu'_1)^w \binom{2}{w} \mu'_{2-w},$$

Or from

$$\mathrm{V}(\boldsymbol{y}) = \mu'_{2,\boldsymbol{y}} - \mu^2_{1,\boldsymbol{y}} = \mu_{2,\boldsymbol{y}} = \sum_{y=1}^{\infty} (2y-1)\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} - \left(\sum_{y=1}^{\infty} \pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right)^2.$$

The DisIx or the variance to mean ratio (VMR) of the DWG-G family can derived as

$$\text{DisIx}(y) = \frac{\sum_{y=1}^{\infty}(2y-1)\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\sum_{y=1}^{\infty}\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}} - \sum_{y=1}^{\infty}\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}.$$

3.3 The moment and cumulant generating function (MGF & CGF)

Theorem 2:

Let $\boldsymbol{Y}$ be NN-RV, where $\boldsymbol{Y} \sim$ DWG-G($\underline{\boldsymbol{\Psi}}$) family, then the MGF of $\boldsymbol{Y}$ can be obtained as

$$M_{\boldsymbol{Y}}(t) = 1 + \sum_{y=1}^{\infty}\{exp(ty) - exp[t(y-1)]\}\,\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}.$$

Proof: The MGF of our NN-RV $\boldsymbol{Y}$ can be derived from $M_{\boldsymbol{Y}}(t) = \sum_{y=0}^{\infty} P_{\underline{\boldsymbol{\Psi}}}(y)\, exp(ty)$. Then, using (3) we have

$$M_{\boldsymbol{Y}}(t) = \sum_{y=0}^{\infty} exp(ty)\begin{bmatrix}\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\\ -\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\end{bmatrix} = 1 + \sum_{y=1}^{\infty}\begin{Bmatrix} exp(ty)\\ -exp[t(y-1)]\end{Bmatrix}\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}. \quad (6)$$

Due to (6) we have

$$\mu'_{r,\boldsymbol{Y}} = E(y^r) = \frac{d^r}{dt^r} M_{\boldsymbol{Y}}(t)|_{(t=0)}.$$

The logarithm of the MGF can be written as the CGF. Therefore, the r[th] cumulant, denoted by $K_{r,\boldsymbol{Y}}$, can be determined from $K_{r,\boldsymbol{Y}} = \frac{d^r}{dt^r}\log\left[M_{\boldsymbol{Y}}(t)\right]|_{(t=0,\text{ and } r=1,2,3,\ldots)}$. The following is an example of how the probability generating function (PGNF) can be written:

$$P_{\boldsymbol{Y}}(s) = 1 + \sum_{y=1}^{\infty}(1-1/s)\, s^{y}\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}|_{(y\in N_0,\pi\in I \text{ and } r=1,2,3,\ldots)}.$$

3.4 Numerical analysis

The parameters E(X), V(X), skewness (S(X)), and kurtosis (K(X)) are used to represent the DWGW distribution in Table 2. According to Table 2, it can be seen that the value of E(X) increases when goes up, but it decreases when _1, _2, and go up at the same time. This is something that can be observed. The S(X)∈ (-0.364031,∞). Its K(Z) has a spread that varies in either direction from 1.001 all the way up to 6.0231026. In accordance with the conventional Poisson distribution (Poisson, 1837), the value of DisIx(X) is either less than 0.1 or larger than one or equal to one. Consequently, applying the DWGW distribution to model count dataset that is "under-dispersed" or "over-dispersed" might be beneficial to the modeling process.

Table 2: E(X), V(X), S(X) and K(X) of the DWGW distribution.

π	β_1	β_2	θ	E(X)	V(X)	S(X)	K(X)	DisIx(X)
0.0001	5	0.5	0.15	1.7×10^{-27}	1.70×10^{-27}	2.45×10^{-13}	∞	1
0.01				4.11×10^{-14}	4.11×10^{-14}	4953981	2.45×10^{13}	1
0.2				2.1×10^{-05}	2.1×10^{-05}	218.6625	47814.59	1
0.4				0.0021679	0.002165	21.44029	461.6198	0.998843
0.6				0.0331932	0.033058	5.455052	32.51006	0.995921
0.8				0.2646421	0.285540	2.145748	8.036714	1.078965
0.99				7.755066	39.26587	1.164480	4.585201	5.063254
0.999				171.56190	10599.05	0.708147	3.303535	61.77975
0.9999				4660.7490	3981772	0.231413	2.691724	854.3203
0.5	0.0001	1.5	1.5	3.496094	12.23946	0.002161	1.000709	3.500897
	0.01			3.098353	10.99523	0.220869	1.129161	3.548736
	0.5			0.203511	0.162279	1.477083	3.195794	0.7973982
	1			0.025766	0.0251017	5.986483	36.83798	0.9742345
	1.5			0.000224	0.0002235	66.85855	4471.066	0.9997764
	2.5			5.4×10^{-20}	5.41×10^{-20}	∞	1.861019	1
0.25	0.5	0.0001	5	10624.45	7105385	1.153516	3.857263	668.7767
		0.01		105.7495	710.6219	1.153313	3.857062	6.719860
		0.1		10.12474	7.187220	1.13423	3.838421	0.709867
		0.5		1.635884	0.331469	1.328594	3.759675	0.2026237
		0.75		0.861674	0.119197	2.094995	5.389769	0.1383322
		0.95		0.732785	0.195811	1.052123	2.106964	0.2672148
		1		0.6858454	0.215461	0.800750	1.641201	0.3141546
0.99	1.5	1.5	0.01	762.5819	38615236	34.02323	1300.836	506375
			0.5	5.020604	6.599397	0.072457	2.460788	1.314463
			1	1.369589	0.3598386	0.364031	2.329377	0.262735
			2	0.7799511	0.1716274	1.351509	2.826577	0.220049
			3	0.2215724	0.1724781	1.340835	2.797839	0.778428
			4	2.61×10^{-09}	2.61×10^{-09}	19616.26	38479774	1

4. Characterizations of the DWG-G Family

The difficulty that arises when attempting to describe a distribution is a key component of distribution theory that has attracted the focus of a significant number of academics working in the applied sciences. When conducting research in these areas, an investigator is frequently curious about identifying whether or not their model follows the suitable distribution.

Proposition 2.1. Let $\boldsymbol{Y}: \Omega \to N_0$ be a random variable. The PMF of $\boldsymbol{Y}$ is (4) if and only if

$$E\left\{\left[\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} + \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right] | \boldsymbol{Y} > \mathcal{U}\right\} = \pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}, \quad \mathcal{U} \in N_0. \tag{7}$$

If $\boldsymbol{Y}$ has PMF (4), then the left-hand side of (7) will be

$$\left[1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U})\right]^{-1} \sum_{y=\mathcal{U}+1}^{\infty}\left[\pi^{2\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} - \pi^{2\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right] = \pi^{-\nabla_{\underline{\boldsymbol{w}}}^{\beta_1,\beta_2}(\mathcal{U}+1)} \sum_{y=\mathcal{U}+1}^{\infty}\left[\pi^{2\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} - \pi^{2\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right]$$

$$= \pi^{-\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}\pi^{2\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})} = \pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}, \quad \mathcal{U} \in N^*.$$

In contrast, in the event that condition (7) is satisfied, then:

$$\sum_{y=\mathcal{U}+1}^{\infty}\left[\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} + \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right] P_{\underline{\boldsymbol{\psi}}}(y) = \left[1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U})\right]\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}$$

$$= \left[1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1) + P_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1)\right]\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}. \tag{8}$$

From (8), we also have

$$\sum_{y=\mathcal{U}+2}^{\infty}\left[\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})} + \pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right] P_{\underline{\boldsymbol{\psi}}}(y) = \left[1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1)\right]\pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}. \tag{9}$$

Now, subtracting (9) from (8), we arrive at

$$(1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1))\left[\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})} - \pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right] + \pi^{\nabla(1+K;\beta_1,\beta_2,\underline{\boldsymbol{w}})}P_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1)$$

$$= \left[\pi^{\nabla(1+K;\beta_1,\beta_2,\underline{\boldsymbol{w}})} + \pi^{\nabla(2+K;\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right]P_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1),$$

or

$$\left[1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1)\right]\left[\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})} - \pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right] = \pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}P_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1).$$

From the last equality, we have

$$h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(\mathcal{U}+1) = \frac{P_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1)}{1-\mathcal{F}_{\underline{\boldsymbol{\psi}}}(\mathcal{U}+1)} = \frac{\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}} - 1,$$

which is the hazard function of PMF (4).

Proposition 2.2. Let $\boldsymbol{Y}: \Omega \to N_0$ be a NN-RV. Then

$$h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(\mathcal{U}+1) - h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(\mathcal{U}) = \left(\frac{\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}\right) - \left(\frac{\pi^{\nabla(\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}\right), \tag{10}$$

$\boldsymbol{Y} \in N_0$, with the initial condition $h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(0) = \pi^{-\left\{[1-G_{\boldsymbol{W}}(1)]^{-\beta_2}-1\right\}^1} - 1$.

Proof. Clearly, if $\boldsymbol{Y}$ has PMF (4), then (10) holds. Now, if (10) holds, then

$$\sum_{\mathcal{U}=0}^{y-1}\left[h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(\mathcal{U}+1) - h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(\mathcal{U})\right] = \sum_{\mathcal{U}=0}^{y-1}\left\{\left(\frac{\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\pi^{\nabla(2+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}\right) - \left(\frac{\pi^{\nabla(\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\pi^{\nabla(1+\mathcal{U};\beta_1,\beta_2,\underline{\boldsymbol{w}})}}\right)\right\},$$

or

$$h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(y) - h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(0) = \left(\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}/\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right) - \frac{1}{\pi^{\nabla(1;\beta_1,\beta_2,\underline{\boldsymbol{w}})}},$$

also, depending on the following initial condition $h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(0) = \pi^{-\nabla_{\underline{\boldsymbol{w}}}^{\beta_1,\beta_2}(1)} - 1$, we have

$$h_{\mathcal{F}_{\underline{\boldsymbol{\psi}}}}(y) = \left(\frac{\pi^{\nabla(y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}}{\pi^{\nabla(1+y|\beta_1,\beta_2,\underline{\boldsymbol{w}})}}\right) - 1, \ \mathcal{U} \in N_0,$$

which refers to the failure function that is equivalent to the PMF (4).

5. Estimation and inference

5.1 Non-Bayesian estimation methods

MLKE method

Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n$ be a random sample (RS) which drown from the DWG-G distribution. The log-likelihood function for $\underline{\boldsymbol{\Psi}}$ is given by

$$\ell(\underline{\boldsymbol{\Psi}}) = \sum_{t=1}^{n} \ln\left[\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{w}})} - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{w}})}\right]|_{(\pi\in I \text{ and } y_t\in N_0)}$$

It may be optimized either via the use of statistical programs or by the resolution of the nonlinear system obtained from $\ell(\underline{\boldsymbol{\Psi}})$ by differentiation. The score vector, $U(\underline{\boldsymbol{\Psi}}) = \left(\partial\ell(\underline{\boldsymbol{\Psi}})/\partial\pi, \partial\ell(\underline{\boldsymbol{\Psi}})/\partial\beta_2, \partial\ell(\underline{\boldsymbol{\Psi}})/\partial\underline{\boldsymbol{\mathcal{W}}}_j\right)^T$, and

$$\partial\ell(\underline{\boldsymbol{\Psi}})/\partial\pi = \frac{\nabla_{\underline{\boldsymbol{\mathcal{W}}}}^{\beta_1,\beta_2}(y_t)\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})-1} - \nabla\left(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})-1}}{\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}},$$

$$\partial\ell(\underline{\boldsymbol{\Psi}})/\partial\beta_1 = \frac{\frac{\partial\nabla\left(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)}{\partial\beta_1}\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\,ln(\pi) - \frac{\partial\nabla\left(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)}{\partial\beta_1}\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\,ln(\pi)}{\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}},$$

$$\partial\ell(\underline{\boldsymbol{\Psi}})/\partial\beta_2 = \frac{\frac{\partial\nabla\left(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)}{\partial\beta_2}\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\,ln(\pi) - \frac{\partial\nabla\left(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)}{\partial\beta_2}\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\,ln(\pi)}{\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}},$$

and

$$\partial\ell(\underline{\boldsymbol{\Psi}})/\partial\underline{\boldsymbol{\mathcal{W}}}_j|j = 1,2,..,p = \frac{\frac{\partial\nabla\left(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)}{\partial\underline{\boldsymbol{\mathcal{W}}}_j}\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\,ln(\pi) - \frac{\partial\nabla\left(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}\right)}{\partial\underline{\boldsymbol{\mathcal{W}}}_j}\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\,ln(\pi)}{\pi^{\nabla(y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}}.$$

Setting $0 = \partial\ell(\underline{\boldsymbol{\Psi}})/\partial\pi = \partial\ell(\underline{\boldsymbol{\Psi}})/\partial\beta_1 = \partial\ell(\underline{\boldsymbol{\Psi}})/\partial\beta_2 = \partial\ell(\underline{\boldsymbol{\Psi}})/\partial\underline{\boldsymbol{\mathcal{W}}}_j$ and then by those equations simultaneously we obtain the MLKEs. In circumstances such as this one, the Newton-Raphson algorithms are the ones that are utilized to carry out the required numerical problem solving.

CRVME method

The CRVME of the major parameter vector ($\underline{\boldsymbol{\Psi}}$) can be calculated via minimizing the following expressions:

$$\mathrm{CRVME}_{(\underline{\boldsymbol{\Psi}})} = \frac{1}{12}n^{-1} + \sum_{t=1}^{n}\left[\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t) - \upsilon_{(t,n)}^{[1]}\right]^2 |_{(\pi\in I \text{ and } y_t\in N_0)},$$

where $\upsilon_{(t,n)}^{[1]} = \frac{2t-1}{2n}$ and then $\mathrm{CRVME}_{(\underline{\boldsymbol{\Psi}})} = \sum_{t=1}^{n}\left[1 - \pi^{\nabla_{\underline{\boldsymbol{\mathcal{W}}}}^{\beta_1,\beta_2}(y_t+1)} - \upsilon_{(t,n)}^{[1]}\right]^2$. Then, we have:

$$\sum_{t=1}^{n} D_{(\pi)}\left(y_t+1,\underline{\boldsymbol{\Psi}}\right)\begin{pmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[1]}\end{pmatrix} = 0, \sum_{t=1}^{n} D_{(\beta_1)}(y_t+1,\underline{\boldsymbol{\Psi}})\begin{pmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[1]}\end{pmatrix} = 0,$$

$\sum_{t=1}^{n} D_{(\beta_2)}(y_t+1,\underline{\boldsymbol{\Psi}})\begin{pmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[1]}\end{pmatrix} = 0$,and $\sum_{t=1}^{n} D_{(\underline{\boldsymbol{\mathcal{W}}}_j)}(y_t+1,\underline{\boldsymbol{\Psi}})\begin{pmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[1]}\end{pmatrix} = 0,$

Which should be solved numerically, where $D_{(\pi)}(y_t+1,\underline{\boldsymbol{\Psi}}) = \partial\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t)/\partial\pi$, $D_{(\beta_1)}(y_t+1,\underline{\boldsymbol{\Psi}}) = \partial\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t)/\partial\beta_1$, $D_{(\beta_2)}(y_t+1,\underline{\boldsymbol{\Psi}}) = \partial\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t)/\partial\beta_2$ and $D_{(\underline{\boldsymbol{\mathcal{W}}}_j)}(y_t+1,\underline{\boldsymbol{\Psi}}) = \partial\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t)/\partial\underline{\boldsymbol{\mathcal{W}}}_j$ are the first mathematical derivative for the CDF of DWG-G distribution WRT π, β_1, β_2 and $\underline{\boldsymbol{\mathcal{W}}}_j$ respectively.

ORLS method

Let $\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t)$ denotes the CDF of DWG-G model and let $y_{1|n} < y_{2|n} < \cdots < y_{n|n}$ be the n ordered RS. The ORLSEs are obtained upon minimizing $\mathrm{ORLSE}_{(\underline{\boldsymbol{\Psi}})} = \sum_{t=1}^{n}\left[\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t) - \upsilon_{(t,n)}^{[2]}\right]^2$, then, we have

$$\mathrm{ORLSE}_{(\underline{\boldsymbol{\Psi}})} = \sum_{t=1}^{n}\left[1 - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} - \upsilon_{(t,n)}^{[2]}\right]^2,$$

where $\upsilon_{(t,n)}^{[2]} = \frac{t}{n+1}$. In light of this, in order to obtain the OLSEs, one must first solve the following non-linear equations:

$$\sum_{t=1}^{n} D_{(\pi)}\left(y_t+1,\underline{\boldsymbol{\Psi}}\right)\begin{bmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[2]}\end{bmatrix} = 0, \sum_{t=1}^{n} D_{(\beta_1)}(y_t+1,\underline{\boldsymbol{\Psi}})\begin{bmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[2]}\end{bmatrix} = 0,$$

$\sum_{t=1}^{n} D_{(\beta_2)}(y_t+1,\underline{\boldsymbol{\Psi}})\begin{bmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[2]}\end{bmatrix} = 0$,and $\sum_{t=1}^{n} D_{(\underline{\boldsymbol{\mathcal{W}}}_j)}(y_t+1,\underline{\boldsymbol{\Psi}})\begin{bmatrix}1-\pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})}\\ -\upsilon_{(t,n)}^{[2]}\end{bmatrix} = 0,$

where $D_{(\pi)}(y_t+1,\underline{\boldsymbol{\Psi}})$, $D_{(\beta_1)}(y_t+1,\underline{\boldsymbol{\Psi}})$, $D_{(\beta_2)}(y_t+1,\underline{\boldsymbol{\Psi}})$ and $D_{(\underline{\boldsymbol{\mathcal{W}}}_j)}(y_t+1,\underline{\boldsymbol{\Psi}})$ defined above.

WLSE method

The WLSE are obtained by minimizing the function WLSE $\left(\underline{\boldsymbol{\Psi}}\right)$ WRT π, β_1, β_2 and $\underline{\boldsymbol{\mathcal{W}}}_j$

$$\text{WLSE}_{(\underline{\boldsymbol{\Psi}})} = \sum_{t=1}^{n} C_{(t,n)}^{[3]} \left[\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_t) - v_{(t,n)}^{[2]}\right]^2,$$

where $C_{(t,n)}^{[3]} = [(1+n)^2(2+n)]/[t(1+n-t)]$. The WLSEs are obtained by solving

$$0 = \sum_{t=1}^{n} C_{(t,n)}^{[3]} \begin{bmatrix} 1 - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} \\ -v_{(t,n)}^{[2]} \end{bmatrix} D_{(\pi)}(y_t+1,\underline{\boldsymbol{\Psi}}), 0 = \sum_{t=1}^{n} C_{(t,n)}^{[3]} \begin{bmatrix} 1 - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} \\ -v_{(t,n)}^{[2]} \end{bmatrix} D_{(\beta_1)}(y_t+1,\underline{\boldsymbol{\Psi}}),$$

$$0 = \sum_{t=1}^{n} C_{(t,n)}^{[3]} \begin{bmatrix} 1 - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} \\ -v_{(t,n)}^{[2]} \end{bmatrix} D_{(\beta_2)}(y_t+1,\underline{\boldsymbol{\Psi}}), \text{ and } 0 = \sum_{t=1}^{n} C_{(t,n)}^{[3]} \begin{bmatrix} 1 - \pi^{\nabla(1+y_t;\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}})} \\ -v_{(t,n)}^{[2]} \end{bmatrix} D_{(\underline{\boldsymbol{\mathcal{W}}}_j)}(y_t+1,\underline{\boldsymbol{\Psi}}),$$

where $D_{(\pi)}(y_t+1,\underline{\boldsymbol{\Psi}})$, $D_{(\beta_1)}(y_t+1,\underline{\boldsymbol{\Psi}})$, $D_{(\beta_2)}(y_t+1,\underline{\boldsymbol{\Psi}})$ and $D_{(\underline{\boldsymbol{\mathcal{W}}}_j)}(y_t+1,\underline{\boldsymbol{\Psi}})$ defined above.

Method of L-moments

The L-moments for the population can be obtained from

$$\mathcal{b}_r = \frac{1}{r}\sum_{m=0}^{r-1}\binom{r-1}{m}(-1)^m E(y_{r-m\ :\ m})|_{\ (r\geq 1)}.$$

The first four L-moments are given by

$$\mathcal{b}_1(\pi,\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}) = E(y_{1:1}) = \mu_1' = L_1,\ \mathcal{b}_2(\pi,\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}) = \frac{1}{2}E(y_{2:2}-y_{1:2}) = \frac{1}{2}(\mu_{2:2}'-\mu_{1:2}') = L_2,$$

$$\mathcal{b}_3(\pi,\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}) = \frac{1}{3}E(y_{3:3}-2y_{2:3}+y_{1:3}) = \frac{1}{3}(\mu_{3:3}'-2\mu_{2:3}'+\mu_{1:3}') = L_3,$$

and

$$\mathcal{b}_4(\pi,\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}) = \frac{1}{4}E(y_{4:4}-3y_{3:4}+3y_{2:4}-y_{1:4}) = \frac{1}{4}(\mu_{4:4}'-3\mu_{3:4}'+3\mu_{2:4}'-\mu_{1:4}') = L_4,$$

where $L_t|t = 1.2.3.4$ is the L-moments for the sample.

KE method

The Kolmogorov estimates (KEs) of π, β_1, β_2 and $\underline{\boldsymbol{\mathcal{W}}}_j$ can be obtained directly by minimizing:

$$K = K(\pi,\beta_1,\beta_2,\underline{\boldsymbol{\mathcal{W}}}_j) = \underset{1\leq t\leq n}{may}\left\{t\frac{1}{n} - \mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_{t\ :\ n}), \mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_{t\ :\ n}) - (t-1)\frac{1}{n}\right\}.$$

Bootstrapping method

The estimation of unknown parameters of a probabilistic distribution can be accomplished through the use of a resampling technique known as bootstrapping. It is a non-parametric method that makes use of the existing dataset in order to generate new datasets (which are referred to as bootstrap samples). This is accomplished by randomly selecting from the primary dataset while simultaneously replacing some of the samples. Here's a general outline of the bootstrapping method for estimating unknown parameters:

i. Start with a sample dataset containing observations or measurements from the population of interest.

ii. Generate multiple bootstrap samples by randomly selecting observations from the original dataset with replacement. Each bootstrap sample should have the same size as the original dataset.

iii. For each bootstrap sample, estimate the unknown parameters of the probabilistic distribution of interest. This can be done by applying the desired estimation method (e.g., maximum likelihood estimation, method of moments, etc.) to each bootstrap sample.

iv. Calculate the parameter estimates obtained from all the bootstrap samples. This will give you a distribution of parameter estimates.

v. From the distribution of parameter estimates, calculate the desired confidence intervals. The most common approach is to use the percentile method, where the lower and upper percentiles of the distribution (e.g., 2.5% and 97.5%) correspond to the lower and upper bounds of the confidence interval, respectively.

vi. Interpret the estimated parameters and their confidence intervals in the context of the problem at hand. These estimates provide information about the uncertainty associated with the parameter estimates.

We will see how this works in the following Section (see Hesterberg (2011)).

Left-Tail of the Second Order Anderson Darling method

The left-tail of the second order Anderson Darling estimates (AD2LEs) cam obtained by minimizing

$$\text{AD2LE}(\underline{\boldsymbol{\mathcal{W}}}) = 2\sum_{t=1}^{n}\log[\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_{t\ :\ n})] + \frac{1}{n}\sum_{t=1}^{n}\frac{2t-1}{\mathcal{F}_{\underline{\boldsymbol{\Psi}}}(y_{t\ :\ n})}.$$

Then, the parameter estimates can be obtained by solving the nonlinear equations

$$\partial\left[\text{AD2LE}(\underline{\boldsymbol{w}})\right]/\partial\pi = 0, \partial\left[\text{AD2LE}(\underline{\boldsymbol{w}})\right]/\partial\beta_1 = 0, \partial\left[\text{AD2LE}(\underline{\boldsymbol{w}})\right]/\partial\beta_2 = 0$$

and

$$\partial\left[\text{AD2LE}(\underline{\boldsymbol{w}})\right]/\partial\underline{\boldsymbol{w}}_j = 0.$$

5.2 Bayesian estimation

Bayesian methods offer several advantages, including the ability to incorporate prior knowledge, handle small sample sizes, and provide a flexible framework for modeling complex problems. However, they also require making assumptions about the prior distribution and can be computationally intensive, especially for high-dimensional problems. In practice, Bayesian estimation often involves numerical methods such as Markov chain Monte Carlo (MCMC) techniques or variational inference to approximate the posterior distribution when analytical solutions are not available.

Assume the gamma priors of the parameters $\pi,\ \beta_1,\ \beta_2$ where

$p_{1;(\zeta_1,\xi_1)}(\pi) \sim \text{beta}(\zeta_1,\xi_1), p_{2;(\zeta_2,\xi_2)}(\beta_1) \sim \text{Gamma}(\zeta_2,\xi_2), p_{3;(\zeta_3,\xi_3)}(\beta_2) \sim \text{Gamma}(\zeta_3,\xi_3)$ and

and assume the uniform prior for $\underline{\boldsymbol{w}}_j$ where

$$p_{4;(\zeta_4,\xi_4)}(\underline{\boldsymbol{w}}_j) \sim \text{Uniform}(\zeta_4,\xi_4),$$

The joint prior distribution can be written as

$$p_{(\zeta_t,\xi_t)}(\pi,\beta_1,\beta_2,\underline{\boldsymbol{w}}_j) = \frac{\pi^{\zeta_1}(1-\pi)^{\xi_1}}{(\xi_4-\zeta_4)B(\zeta_1,\xi_1)}\frac{\xi_2^{\zeta_2}}{\Gamma(\zeta_2)}\,\beta_1^{\zeta_2-1}\,exp[-(\beta_1\xi_2)]\frac{\xi_3^{\zeta_3}}{\Gamma(\zeta_3)}\,\beta_2^{\zeta_3-1}\,exp[-(\beta_2\xi_3)]$$

The posterior distribution can be formulated as

$$p\left(\pi,\beta_1,\beta_2,\underline{\boldsymbol{w}}_j|\underline{\boldsymbol{y}}\right) \propto \text{likelihood}(\underline{\boldsymbol{\Psi}}_j|\underline{\boldsymbol{y}}) \times \pi_{(\zeta_t,\xi_t)}(\pi,\beta_1,\beta_2,\underline{\boldsymbol{w}}_j).$$

The means of the marginal posteriors of the variables in question are what the Bayesian estimators correspond to when the squared error loss function is taken into consideration. Applying the previously outlined formulas to the computation procedure will not result in the production of the Bayesian estimates as this is mathematically impossible. The calculation based on the numbers has to be carried out in an approximative manner as a result. We strongly advise that you make use of Markov chain Monte Carlo (MCMC) methods, in particular the Gibbs sampler and the Metropolis Hastings (M-H) approach. Because the conditional posteriors of the parameters cannot be acquired in any standard form, it is recommended that a hybrid MCMC be used for pulling samples from the joint posterior of the parameters. This is because the conditional posteriors of the parameters cannot be acquired in any standard form. This will make it possible to extract the whole conditional posteriors in a much shorter amount of time. One possible formulation of the simulation's algorithm is as follows:

1) Provide the initial values, say π,β_1,β_2 and $\underline{\boldsymbol{w}}_j$ then at $t^{(\text{th})}$ stage,
2) Generate $\pi_{(t)} \sim \pi_1\left(\pi_{(t)}|_{\beta_{1(t)},\beta_{2(t)},\underline{\boldsymbol{w}}_{j(t)}},\underline{\boldsymbol{y}}\right)$;
3) Generate $\beta_{1(t)} \sim \pi_2\left(\beta_{1(t)}|_{\pi_{(t)},\beta_{2(t)},\underline{\boldsymbol{w}}_{j(t)}},\underline{\boldsymbol{y}}\right)$;
4) Generate $\beta_{2(t)} \sim \pi_2\left(\beta_{2(t)}|_{\pi_{(t)},\beta_{1(t)},\underline{\boldsymbol{w}}_{j(t)}},\underline{\boldsymbol{y}}\right)$;
5) Generate $\underline{\boldsymbol{w}}_{j(t)} \sim \pi_3\left(\underline{\boldsymbol{w}}_{j(t)}|_{\pi_{(t)},\beta_{1(t)},\beta_{2(t)}},\underline{\boldsymbol{y}}\right)$;
6) Repeat steps $2-5$, $M = 50000$ times to obtain the sample of size M from the corresponding posteriors of interest. Obtain the Bayesian estimates of π,β_1,β_2 and $\underline{\boldsymbol{w}}_j$ using the following formulae.

$$\hat{\pi} = \frac{1}{M-M_0}\sum_{\hbar=1+M_0}^{M}\pi^{[\hbar]}, \hat{\beta}_1 = \frac{1}{M-M_0}\sum_{\hbar=1+M_0}^{M}\beta_1{}^{[\hbar]}, \hat{\beta}_2 = \frac{1}{M-M_0}\sum_{\hbar=1+M_0}^{M}\beta_2{}^{[\hbar]}, \underline{\boldsymbol{w}}_j = \frac{1}{M-M_0}\sum_{\hbar=1+M_0}^{M}\underline{\boldsymbol{w}}_j{}^{[\hbar]}$$

respectively, where $M_0(\approx 50000)$.

6. Simulations

The MCMC simulation is of great importance due to its ability to handle complex probabilistic distributions, facilitate Bayesian inference, provide flexibility in model formulation, quantify uncertainty, handle intractable integrals, explore posterior distributions, and offer computational efficiency. These features make MCMC an invaluable tool in a wide range of fields, including statistics, machine learning, computational biology, physics, and many more. This assessment is performed under the criteria of mean squared errors (MSEs). First, we generated 1000 samples of the DWGE distribution, where n_1=50, n_2=100, n_3=200, n_4=300. The MSEs are provided in Tables 3, 4 and 5. Based on Tables 3, 4 and 5 we note that, generally, the method of the MLKE is recommended as the best estimations for all

sample sizes with the minimum values for the MSEs, however, one can conclude that the general performance of all estimation methods improves when the sample size increases.

Table 3: MSEs where π= 0:6; β_1 = 0:4; β_2 = 1:5; θ= 0:9.

n		MLKE	ORLS	WLS	CRVM	Bayes	Moment	KE	bootstarp	AD2LE
50	π	0.00212	0.00282	0.00312	0.00261	0.00311	0.00346	0.00225	0.00331	0.00290
	θ	0.00267	0.02421	0.01674	0.02121	0.01600	0.00410	0.00801	0.01170	0.02280
	$\beta1$	0.00219	0.00919	0.01140	0.00754	0.00383	0.00343	0.00496	0.00332	0.00973
	$\beta2$	0.00301	0.07821	0.08989	0.06734	0.04432	0.00533	0.04695	0.02094	0.08321
100	π	0.00107	0.00137	0.00174	0.00151	0.00173	0.00162	0.00109	0.00109	0.00169
	θ	0.00125	0.01043	0.00640	0.01004	0.00323	0.00180	0.00421	0.00367	0.01114
	$\beta1$	0.00109	0.00354	0.00678	0.00393	0.00102	0.00160	0.00228	0.00258	0.00548
	$\beta2$	0.00146	0.03405	0.05289	0.03727	0.01432	0.00211	0.02212	0.00481	0.04810
200	π	0.00050	0.00066	0.00082	0.00068	0.00054	0.00084	0.00057	0.00097	0.00075
	θ	0.00053	0.00461	0.00238	0.00464	0.00154	0.00089	0.00202	0.00107	0.00530
	$\beta1$	0.00050	0.00151	0.00357	0.00154	0.00041	0.00084	0.00114	0.00079	0.00233
	$\beta2$	0.00058	0.01564	0.02701	0.01599	0.00533	0.00098	0.01112	0.00382	0.02113
300	π	0.00031	0.00044	0.00056	0.00043	0.00033	0.00049	0.00036	0.00074	0.00047
	θ	0.00031	0.00313	0.00119	0.00290	0.00129	0.00050	0.00137	0.00088	0.00335
	$\beta1$	0.00031	0.00098	0.00256	0.00093	0.00034	0.00049	0.00075	0.00066	0.00143
	$\beta2$	0.00032	0.01035	0.01919	0.00992	0.00450	0.00051	0.00727	0.00352	0.01329

Table 4: MSEs where π= 0:7; β_1 = 0:7; β_2 = 1:8; θ= 0:8.

n		MLKE	ORLS	WLS	CRVM	Bayes	Moment	KE	bootstarp	AD2LE
50	π	0.00171	0.00203	0.00235	0.00214	0.00328	0.00256	0.00177	0.00160	0.00225
	θ	0.01691	0.21381	0.18406	0.01621	0.05010	0.00564	0.00525	0.01218	0.01628
	β_1	0.00188	0.00927	0.01122	0.00950	0.00860	0.00275	0.00715	0.00169	0.01084
	β_2	0.00236	0.03930	0.04706	0.04113	0.03163	0.00389	0.03109	0.00478	0.04589
100	π	0.00076	0.00104	0.00103	0.00096	0.00076	0.00120	0.00089	0.00149	0.00100
	θ	0.00160	0.03049	0.02567	0.00926	0.00665	0.00180	0.00241	0.00165	0.00929
	β_1	0.00080	0.00483	0.00540	0.00439	0.00194	0.00123	0.00359	0.00130	0.00507
	β_2	0.00091	0.02043	0.02194	0.01889	0.01196	0.00140	0.01572	0.00233	0.02103
200	π	0.00037	0.00050	0.00054	0.00050	0.00039	0.00055	0.00041	0.00106	0.00052
	θ	0.00045	0.00475	0.00395	0.00427	0.00628	0.00072	0.00104	0.00065	0.00428
	β_1	0.00038	0.00237	0.00302	0.00237	0.00152	0.00056	0.00162	0.00096	0.00269
	β_2	0.00039	0.00991	0.01181	0.00996	0.00800	0.00064	0.00716	0.00115	0.01091
300	π	0.00026	0.00034	0.00037	0.00034	0.00027	0.00040	0.00028	0.00024	0.00035
	θ	0.00033	0.00309	0.00267	0.00309	0.00487	0.00050	0.00070	0.00031	0.00309
	β_1	0.00027	0.00164	0.00214	0.00163	0.00084	0.00041	0.00111	0.00024	0.00187
	β_2	0.00030	0.00685	0.00821	0.00688	0.00369	0.00046	0.00491	0.00023	0.00754

Table 5: MSEs where π= 0:1; β_1 = 0:6; β_2 = 0:4; θ= 0:4.

n		MLKE	ORLS	WLS	CRVM	Bayes	Moment	KE	bootstarp	AD2LE
50	π	0.00125	0.00151	0.00147	0.00168	0.00540	0.01144	0.00131	0.00347	0.00173
	θ	0.00234	0.05010	0.04077	0.02830	0.00319	0.01680	0.01064	0.00379	0.02968
	β_1	0.00133	0.06563	0.05237	0.06550	0.01562	0.18678	0.04753	0.00863	0.06565
	β_2	0.00166	0.01053	0.01049	0.00883	0.00680	0.01863	0.00852	0.00506	0.00903
100	π	0.00059	0.00077	0.00077	0.00074	0.00147	0.00952	0.00064	0.00124	0.00076

	θ	0.00080	0.01666	0.01035	0.01563	0.00239	0.01112	0.00529	0.00100	0.01661
	β_1	0.00071	0.03178	0.02601	0.02947	0.00677	0.03043	0.02554	0.00217	0.02957
	β_2	0.00065	0.00464	0.00461	0.00410	0.00329	0.01140	0.00331	0.00186	0.00417
200	π	0.00029	0.00038	0.00039	0.00038	0.00129	0.00837	0.00031	0.00035	0.00039
	θ	0.00033	0.00634	0.00320	0.00570	0.00184	0.00877	0.00229	0.00046	0.00627
	β_1	0.00031	0.01545	0.01336	0.01537	0.00546	0.01023	0.01137	0.00149	0.01544
	β_2	0.00030	0.00217	0.00223	0.00218	0.00300	0.00845	0.00151	0.00036	0.00222
300	π	0.00019	0.00025	0.00026	0.00026	0.00037	0.00409	0.00021	0.00026	0.00026
	θ	0.00023	0.00414	0.00180	0.00332	0.00098	0.00422	0.00152	0.00024	0.00367
	β_1	0.00023	0.01014	0.00896	0.01020	0.00525	0.00408	0.00774	0.00113	0.01024
	β_2	0.00021	0.00145	0.00149	0.00138	0.00270	0.00408	0.00103	0.00033	0.00141

7. Comparing methods under reliability, medicine, agriculture real-count datasets

This Section presents four real dataset applications under the discrete family. Two applications are allocated for modeling the real-count dataset and two are allocated for modeling the real-zero-inflated dataset. Count datasets arise in a variety of fields, including ecology, finance, epidemiology, and social sciences. Modeling count dataset is essential for understanding the relationships between the response variable and its predictors and for making accurate predictions. Modeling zero-inflated dataset is important for understanding the unique characteristics of dataset with an excessive number of zeros. Standard discretization methods may not be appropriate for zero-inflated dataset, so specialized models such as the zero-inflated model and the hurdle model should be used instead. These models allow for the estimation of the probability of excess zeros and the parameters of the count distribution and provide a more accurate representation of the dataset and better predictions.

7.1 Time to failures dataset of 50 devices

Count time to failure dataset refers to dataset where the variable of interest is the number of failures or events that occur within a specific time period. This type of dataset can be modeled using discrete probabilistic distributions, which are appropriate for counting dataset because they can only take on integer values. In this study, the dataset consists of the time to failures of 50 devices, measured in weeks, during a life test. The dataset has been analyzed by Bodhisuwan and Sangpoom (2016). Table 6 presents the estimates obtained using Bayesian and non-Bayesian estimation methods, as well as the KG-SM and $\boldsymbol{p_v}$ statistics for the time to failure dataset. From the results in Table 6, it appears that the ORLSE method is the most suitable, as it yields a KG-SM value of 0.11806 and a $\boldsymbol{p_v}$ value of 0.48873.

7.2 Time to failures of 15 electronic components

This lifetime dataset describes the time to failures for 15 electronic components during an acceleration lifespan test (for detailed information, see Lawless et al. (2003)). The table that follows (Table 7) contains estimates for the dataset pertaining to electronic components. These estimates are presented in the form of KG-SM and $\boldsymbol{p_v}$ statistics, as well as Bayesian and non-Bayesian estimating methods. The KE technique, which has KG-SM equal to 0.09438 and $\boldsymbol{p_v}$ equal to 0.99933, is a candidate for the title of "best method" based on the dataset presented in Table 7.

7.3 Counts of cysts of kidneys

The dataset referred to as cysts of kidneys are dataset in which the response variable is the number of cysts that were observed in the kidneys of persons. Because the dataset on cysts of the kidneys is count dataset, discrete probabilistic distributions can be utilized in order to describe them. The Poisson distribution and the negative binomial distribution are two examples of discrete probabilistic distributions that are frequently utilized in situations like this one. This dataset depicts the counts of cysts of corticosteroid-induced kidneys that are dysmorphogenetic. These cysts are connected with the dysregulated expression of recognized cytogenic molecules as well as Indian hedgehog (for more information, see Chan et al. (2009)). In Table 8, the estimates for the renal dataset are presented using both Bayesian and non-Bayesian estimation methods, as well as KG-SM and $\boldsymbol{p_v}$ statistics. According to Table 8, the MLKE technique with KG-SM equal to 0.14998 and $\boldsymbol{p_v}$ equal to 0.69855 is a contender for the title of "best method"

7.4 Agricultural real-life data

These numbers indicate the total number of European corn-borer larvae parasites that were found in the field (for further information, see Bebington et al. (2012)). Dataset in which the response variable is the count of parasites that attack European corn-borer larvae are referred to as the number of parasites that attack European corn-borer larvae in

this article. Given that these are count dataset, modeling them with discrete probabilistic distributions is a viable option. In Table 9, the estimates for corn-borer larvae parasites dataset are shown under both Bayesian and non-Bayesian estimation methods, as well as KG-SM and p-v statistics. Table 9 shows that the CRVM approach, which has KG-SM equal to 0.96028 and p_v equal to 0.32712, is one of the options for the best method.

Table 6: Estimates, KS and p_v for time to failures of devices.

	π	β_1	β_2	θ	KS	p_v
MLKE	0.6157767958	0.1573461107	0.0000033379	3.3596325474	0.13897	0.28903
ORLS	0.679365447	0.1855967179	0.0019523474	1.8995243873	0.11806	0.48873
WLS	0.6771246495	0.2658284792	0.0002002012	2.3648433070	0.17369	0.09787
CRVM	0.7607090861	0.2176200131	0.0224293761	1.3501642340	0.12500	0.41534
Bayesian	0.6793513549	0.2033032537	0.0019338520	1.9160176137	0.17835	0.08309
Moment	0.6974733370	0.1149606951	0.0011362417	2.1459628209	0.15566	0.17720
KE	0.0231135581	1.3951239157	0.0127164259	0.8093524699	0.15531	0.17913
Bootst.	0.6070316325	0.1612146631	0.000001858	3.5075336967	0.17834	0.08313
AD2LE	0.5278679339	0.1852968998	0.0000078247	3.0520756022	0.11875	0.48115

Table 7: Estimates, KS and p_v for time to failures of electronic components.

	π	β_1	β_2	θ	KS	p_v
MLKE	0.2439934193	0.2498029795	0.0000000379	4.3903617382	0.12316	0.97679
ORLS	0.0649000184	0.7318675329	0.0013791993	1.5001317839	0.11844	0.98447
WLS	0.1346212203	0.6583027489	0.0011088261	1.5916908457	0.10994	0.99347
CRVM	0.0468066652	0.7608445569	0.0011511311	1.4914083278	0.10217	0.99760
Bayesian	0.2364811243	0.2654651850	0.0000000381	4.3382582127	0.14656	0.90404
Moment	0.2574440612	0.4399775046	0.0001062774	2.3533843673	0.10728	0.99525
KE	0.0549459793	0.6288539228	0.0003673616	1.7683089416	0.09438	0.99933
Bootst	0.2315123465	0.2867285737	0.0000000419	4.3917475644	0.139095	0.93369
AD2LE	0.0317952805	0.9419816263	0.0025303378	1.3227447796	0.10883	0.99427

Table 8: Estimates, KS and p_v for kidneys dataset.

	π	β_1	β_2	θ	KS	p_v
MLKE	0.538286604	0.720153555	0.9688003235	0.435380284	0.14998	0.69855
ORLS	0.3044313503	1.0823722268	0.5688744335	0.3691428485	0.38825	0.53322
WLS	0.2062488312	0.7551061653	0.3880181575	0.5833695645	0.40767	0.52315
CRVM	0.3230639855	0.8877792048	0.5624569251	0.4406811894	0.25131	0.61615
Bayesian	0.4940295209	0.7036908397	0.9299949144	0.4389761409	0.6233	0.42982
Moment	0.0419180687	0.0975614653	0.0000018545	5.9974151455	0.54497	0.46038
KE	0.3520133743	0.9075204667	0.4982888878	0.4406811894	2.91018	0.08802
Bootst	0.1867694728	0.6471461723	0.3760235405	0.6636048898	2.13569	0.14391
AD2LE	0.3529607368	1.3428936378	0.6324668229	0.2900647784	0.38277	0.53612

Table 9: Estimates, KS and p_v for corn-borer larvae parasites dataset.

	π	β_1	β_2	θ	KS	p_v
MLKE	0.0468873711	5.5470002735	0.538697811	0.1462873188	1.08781	0.29696
ORLS	0.0265674121	4.6718916679	0.4957373562	0.1888898179	1.05235	0.30497
WLS	0.0982814922	4.8131560409	0.5361854384	0.1768047825	0.96028	0.32712
CRVM	0.0211054595	4.9425178858	0.4994558176	0.1783063213	0.97728	0.32287
Bayesian	0.0580677884	5.3511844765	0.5370877054	0.146794422	1.77406	0.18288
Moment	0.0100595854	0.1729227865	0.0000023539	5.9885425122	1.35159	0.24500
KE	0.0175309812	0.4721164303	0.0655919866	0.0099672330	1463.00	<0.0001
Bootst.	0.0768506047	5.2742164412	0.5404704993	0.1606818352	119.542	<0.0001

AD2LE	0.026674495	4.9348095934	0.502307644	0.1849002612	1.34243	0.24661

8. Comparing discrete models under reliability, medicine, agriculture real-count datasets

We illustrate the adaptability of the DWGE and DWGW distributions by using four distinct applications from the real world, and we also discuss the significance of these distributions. The log-likelihood function, AIC, CAIC, χ^2_{ϑ} with degree of freedom (d.f), KG-SM, and its $\wp_{\vartheta}$ are the approaches that are used in the process of analyzing and comparing the fitted distributions (see Table 10). Table 8 provides a rundown of the various offerings made by the rival companies.

Table 10: The competitive models.

Discrete distribution	Abbreviation
Standard exponential, Standard Weibull and Standard Rayleigh	DE, DW and DR
Standard negative binomial (Dougherty (1992))	NB
Standard inverted Rayleigh	DINR
Standard inverted Weibull	DINW
Standard Lindley and Standard Poisson	DLN and Poisson
Standard Lindley of the second type	DLN-II
Standard loglogistic, Standard Lomax and Standard Pareto	DLL, DLX and DPAR
Standard Burr of the type XII	DBTXII
Standard exponentiated Weibull	EDW
Standard exponentiated Lindley	EDLN
discrete generated exponential of the second type	DGEXII

8.1 Time to failures (times to failure dataset for 50 devices)

Dataset on count time to failures are dataset sets whose response variable is the number of failures or occurrences of an event during a predetermined period of time. These dataset sets are referred to as count time to failures dataset. Modeling dataset on count time to failures can be accomplished with the help of discrete probabilistic distributions. Discrete distributions are suitable for counting dataset since they can only take on integer values. This makes them more precise. As a result of the findings of Bebbington et al. (2012), we evaluate the fits of the DWGW distribution in comparison to those of a number of other models, including EDW, DW, DINW, DLN-II, EDLN, DLL, and DPAR. The box graph, the quantile-quantile graph, and the total time in test graph for the time to failures dataset are presented in Figure 3. Table 11 contains a listing of the MLKEs as well as the standard errors (SEs). Tables 12 contain a listing of the statistics on the test's appropriateness. According to Table 12, the DWGW offers the best fits in comparison to all other competitor models, with a value of -ℓ equal to 219.452, an AIC equal to 446.905, a CAIC equal to 447.794, KG-SM equal to 0.13897, and $\wp_{\vartheta}$ equal to 0.28903. The estimated HRF (EHRF), the estimated SF (ESF) or the Kaplan-Meier SF, and the Probability-Probability (PR-PR) graphs for time to failures dataset are presented in Figure 4.

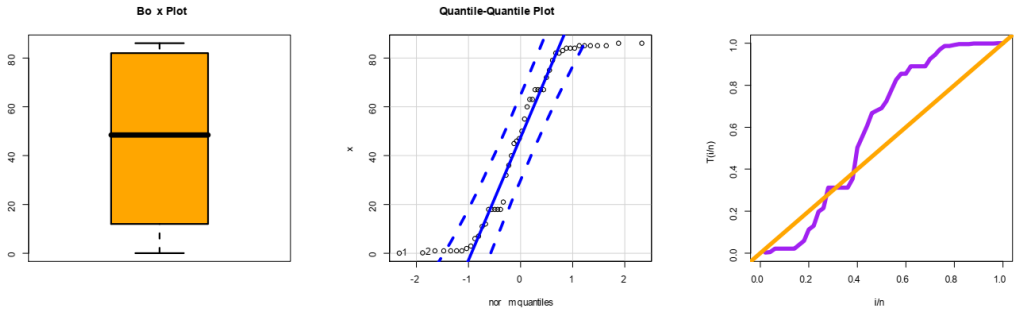


Figure 3: Box graph, Q-Q graph and TTT graph for the time to failures dataset.

Table 11: MLKEs (SEs) for time to failures dataset.

Model	$\hat{\pi}$ MLKE (SE)	$\widehat{\beta_1}$ MLKE (SE)	$\widehat{\beta_2}$ MLKE (SE)	$\hat{\theta}$ MLKE (SE)
DWGW	0.61582 (0.0761)	0.15735 (0.0377)	0.00003 (<0.0001)	3.3596 (0.49523)
EDW	0.98914 (0.1644)	1.13778 (3.2271)	0.78474 (3.0533)	
DW	0.9809	1.02331		

	(0.0112)	(0.1313)
DINW	0.01808	0.58211
	(0.01311)	(0.0612)
DLN-II	0.9692	0.05812
	(0.0052)	(0.0271)
EDLN	0.97182	0.4802
	(0.0051)	(0.0868)
DLLc	1.00013	0.43925
	(0.3246)	(0.0624)
DPAR	0.73924	
	(0.0343)	

Table 12: The statistics for the time to failures dataset about the goodness of fit.

	DWGW	EDW	DW	DINW	DLN-II	EDLN	DLLc	DPAR
-ℓ	219.452	240.212	241.616	261.877	240.631	240.307	294.93	275.882
AIC	446.905	486.719	487.225	527.824	485.198	484.553	593.88	553.741
CAIC	447.794	487.204	487.555	528.101	485.455	484.888	594.05	553.885
KG-SM	0.13897	0.19488	0.18656	0.25777	0.18632	0.19453	0.5354	0.33535
$\mathcal{P}_v$	0.28903	0.04512	0.06111	0.00322	0.06422	0.04525	< 0.0010	< 0.0010

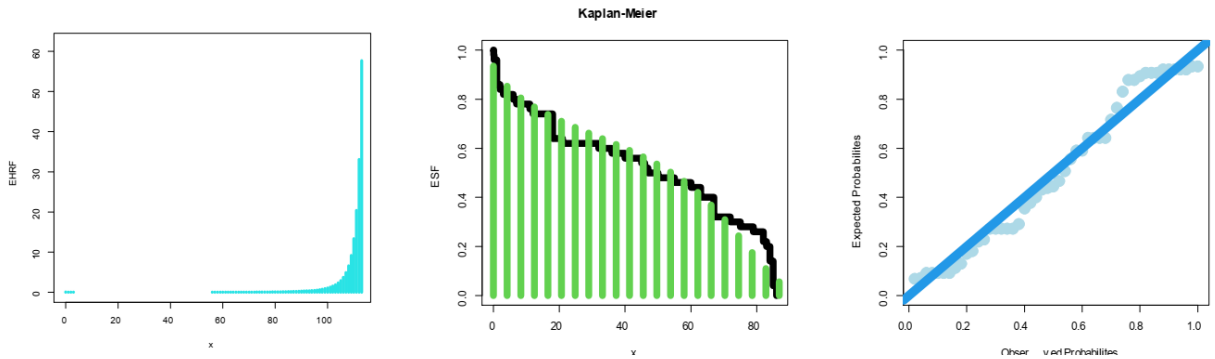


Figure 4: The EHRF, ESF and PR-PR graphs for time to failures dataset.

8.2 Time to failures (times to failure dataset for 15 devices)

Modeling count time to failure dataset using discrete probabilistic distributions is possible; the Poisson distribution and the negative binomial distribution are the two that are utilized most frequently in this endeavor. The properties of the dataset and the research issue that are being addressed both play a role in the decision on which distribution to choose. For the purpose of this application, we evaluate the degree to which the discrete DE, DGEXII, DR, DIR, DINW, DLX, DBTXII, and DPAR models match the dataset in comparison to the fits provided by the DWGE and DWGW distributions. The MLKEs and SEs are detailed in Tables 13, which may be found here. Tables 14 contain a listing of the statistics pertaining to the appropriateness of the exam. According to the results of Table 14, the DWGW model offers the best fits in comparison to all other competitive models, with a value of -ℓ equal to 62.958, AIC equal to 133.915, CAIC equal to 137.915, KG-SM equal to 0.12315, and $\mathcal{P}_v$ equal to 0.97679. The box graph, the Q-Q graph, and the TTT graph for the second time to failures dataset are presented in Figure 5. The EHRF, ESF, and PR-PR graphs for the second time to failures dataset are presented in Figure 6.

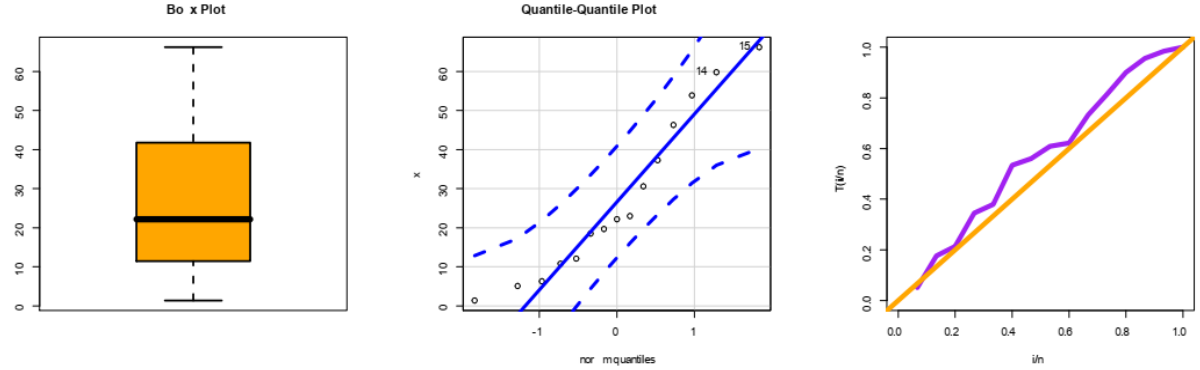


Figure 5: Box graph, Q-Q graph and TTT graph for the second time to failures dataset.

Table 13: MLKEs (SEs) for second time to failures dataset.

	$\hat{\pi}$	$\widehat{\beta_1}$	$\widehat{\beta_2}$	$\hat{\theta}$
Model	MLKE (SE)	MLKE (SE)	MLKE (SE)	MLKE (SE)
DWGW	0.24399	0.24980	0.000004	4.39036
	(0.1943)	(0.25474)	(<0.0001)	(4.6153)
DGEXII	0.95623	1.49121		
	(0.01311)	(0.53516)		
DINW	2.21×10^{-4}	0.87551		
	(7.83×10^{-4})	(0.1644)		
DLX	0.01233	104.514		
	(0.03910)	(84.431)		
DBTXII	0.97513	13.3676		
	(0.05132)	(27.785)		
DR	0.99952			
	(2.58×10^{-4})			
DINR	1.82×10^{-7}			
	(0.05553)			
DPAR	0.72034			
	(0.0615)			
DE	0.96532			
	(0.00915)			

Table 14: The statistics reflect the degree to which the data fits the second attempt to the failures.

	DWGW	DE	DGEXII	DR	DINR	DINW	DLX	DB-XII	DPAR
-ℓ	62.958	65.011	64.420	66.392	89.095	68.703	65.863	75.724	77.401
AIC	133.915	134.24	134.83	134.814	180.178	141.441	135.728	155.433	156.887
CAIC	137.915	136.33	135.82	136.110	180.503	142.436	136.728	156.466	157.136
KG-SM	0.12315	0.1778	0.1288	0.21615	0.6982	0.20925	0.20506	0.38829	0.4052
$\wp_v$	0.97679	0.6731	0.9376	0.43317	< 0.0010	0.48215	0.4905	0.01516	0.00961

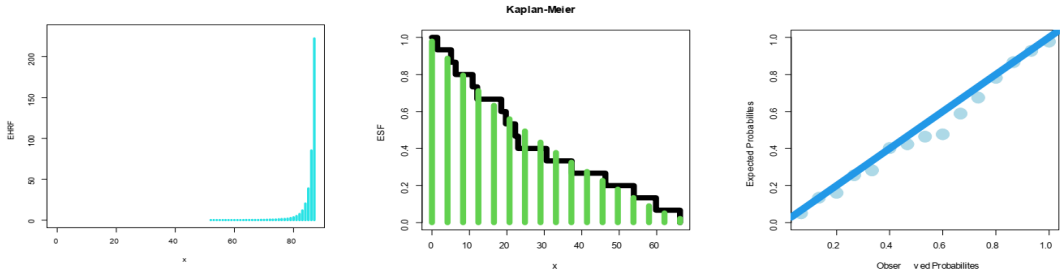


Figure 6: The EHRF, ESF and PP graphs for second time to failures dataset.

8.3 Counts of cysts of kidneys

Discrete probabilistic distributions can be used to represent counts of kidney cysts; the Poisson distribution and the negative binomial distribution are the two that are utilized the most frequently. The features of the dataset, such as the mean and the variance, as well as the research issue that is being addressed, all play a part in determining which distribution to choose. When analyzing this particular dataset set, we evaluate the adequacy of the DWGE and DWGW distributions in relation to other prevalent statistical models, including the DLX, DW, DLN-II, DINW, DR, DLN, DE, and the standard Poisson distributions. The MLKEs and SEs are detailed in Table 15, which may be found here. The information regarding the goodness of fit can be seen in table 16. According to Table 16, the DWGW model has the best fit when compared to all other competitive models with the following values: $-\ell$=166.846, AIC =341.693, CAIC =342.074, χ^2_V =0.150, and $\wp_v$=0.699. The box graph, the Q-Q graph, and the TTT graph for the kidney's dataset are presented in Figure 7. The EPMF, ESF, EHRF, ECDF, and PP graphs for the kidney's dataset are presented in Figure 8.

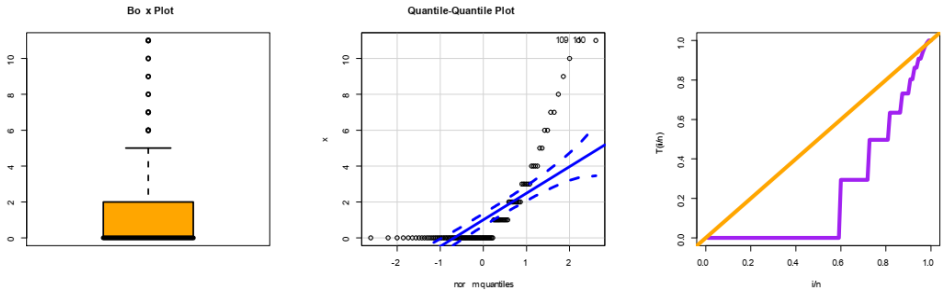


Figure 7: The kidneys dataset includes a box graph, a Q-Q graph, and a TTT graph.

Table 15: MLKEs (SEs) for kidneys dataset.

	$\hat{\pi}$	$\widehat{\beta_1}$	$\widehat{\beta_2}$	$\hat{\theta}$
Model	MLKE	MLKE	MLKE	MLKE
DWGW	0.5383	0.72015	0.9688	0.43543
	(0.3928)	(0.54461)	(0.8922)	(0.3513)
DW	0.75032	0.43151		
	(0.0844)	(0.3404)		
DINW	0.58135	1.04944		
	(0.0482)	(0.146).5		
DLN-II	0.5813	0.00165		
	(0.0457)	(0.05842)		
DLX	0.14032	1.83049		
	(0.09839)	(0.9516)		
Poi.	1.39024			
	(0.1126)			
DE	0.58124R			
	(0.03036)			
DR	0.901433			
	(0.00883)			
DLN	0.436223			
	(0.02617)			

Table 16: The statistics for the kidney's dataset about the appropriateness of the fit.

Z	OF	DWGW	DW	DINW	DR	DEx	DLN	DLN-II	DLX	Poi.
0	65	64.484	59.01	63.91	11.00	46.09	40.25	46.031	61.888	27.421
1	14	14.583	19.84	20.70	26.83	26.78	29.83	26.771	21.015	38.081
2	10	8.9816	10.78	8.05	29.55	15.56	18.36	15.560	9.650	26.470
3	6	6.139	6.260	4.23	22.23	9.04	10.35	9.0452	5.230	12.26
4	4	4.368	4.190	2.60	12.49	5.25	5.527	5.270	3.173	4.26
5	2	3.165	2.0115	1.751	5.424	3.05	2.861	3.063	2.062	1.176
6	2	2.309	1.991	1.259	1.852	1.777	1.400	1.777	1.421	0.272
7	2	1.688	1.327	0.951	0.500	1.031	0.71	1.0416	1.021	0.050
8	1	1.231	0.988	0.741	0.111	0.601	0.353	0.6016	0.761	0.010
9	1	0.893	0.866	0.599	0.0173	0.359	0.1724	0.355	0.577	0.000
10	1	0.645	0.764	0.468	0.000	0.203	0.087	0.200	0.462	0.000
11	2	0.462	1.991	4.741	0.000	0.281	0.071	0.282	2.741	0.000
-ℓ		166.846	170.14	172.93	277.78	178.77	189.15	178.85	170.48	246.21
AIC		341.693	344.28	349.87	557.56	359.53	380.23	361.54	344.96	494.42
CAIC		342.074	344.39	349.98	557.59	359.57	380.35	361.60	345.07	494.46
$\chi^2_{\hat{v}}$		0.1503	3.1255	6.4636	321.073	22.881	43.483	22.894	3.3165	294.105
d.f		1	3	3	4	4	4	3	3	4
p_v		0.6993	0.3731	0.0913	<0.0010	0.0010	<0.0010	<0.0010	0.34510	<0.0010

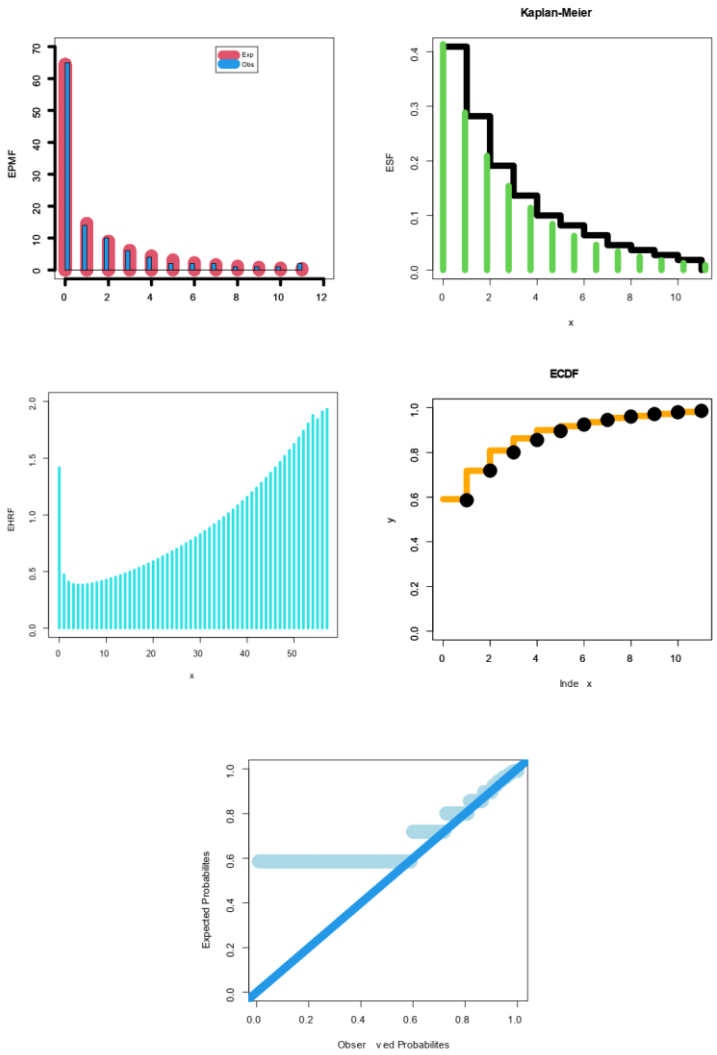


Figure 8: The EPMF, ESF, EHRF, ECDF, PP graphs for kidneys dataset.

8.4 The number of parasitic European corn borer larvae

It is possible to simulate the number of parasitic European corn borer larvae using discrete probabilistic distributions. The Poisson distribution and the negative binomial distribution are the two that are utilized the most frequently in this modeling process. The features of the dataset, such as the mean and the variance, as well as the research issue that is being addressed, all play a part in determining which distribution to choose. In the following part, we are going to evaluate the DWGW distribution in relation to other popular distribution models such as the DGINW, DINW, DBTXII, DIR, DR, NB, and DPAR distributions, as well as the Poisson distribution. The MLKEs and SEs are detailed in Tables 17, which may be found here. The information regarding the goodness of fit can be seen in table 18. According to Table 18, the DWGW offers the best fits in comparison to all other competitor models, with a value of -ℓ equal to 200.438, AIC equal to 408.875, CAIC equal to 409.223, χ^2_{ϑ} equal to 1.088, and p_{ϑ} equal to 0.297. The box graph, the Q-Q graph, and the TTT graph for the corn-borer larvae dataset are presented in Figure 7. Graphs of the EPMF, ESF, EHRF, ECDF, and PP variables for corn-borer larvae dataset are shown in Figure 8. Figure 9 shows the box graph, Q-Q graph and TTT graph for corn-borer larvae dataset. However, Figure 10 gives the EPMF, ESF, EHRF, ECDF, PP graphs for corn-borer larvae dataset.

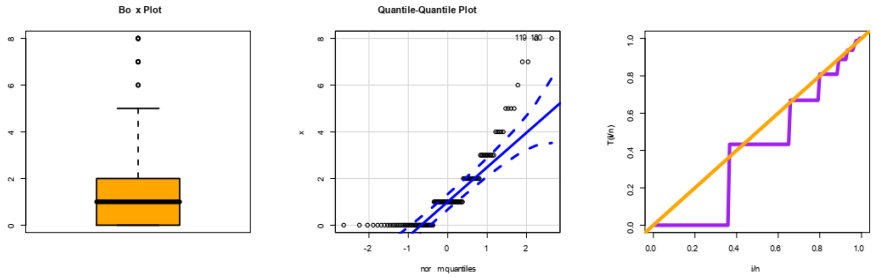


Figure 9: Box graph, Q-Q graph and TTT graph for corn-borer larvae dataset.

Table 17: MLKEs (SEs) for corn-borer larvae dataset.

	$\hat{\pi}$	$\widehat{\beta_1}$	$\widehat{\beta_2}$	$\hat{\theta}$
Model	MLKE	MLKE	MLKE	MLKE
DWGW	0.0469	5.5470	0.5387	0.1463
	(0.2321)	(3.539)	(0.1376)	(0.095)
DGW	0.04502	2.5395	2.15936	0.4791
	(0.429)	(4.703)	(2.6974)	(0.465)
DINW	0.34525	1.5411		
	(0.0433)	(0.1561)		
DBTXII	0.51939	2.35815		
	(0.05124)	(0.3668)		
DINR	0.3188			
	(0.0423)			
NB	0.87024	9.95611		
	(0.0365)	(0.09516)		
DR	0.86733			
	(0.0124e)			
DPAR	0.32915			
	(0.03437)			
Poi	1.483515			
	(0.02515)			

Table 18: The goodness of fit statistics for corn-borer larvae dataset.

Z	OF	DWGW	DINW	DBTXII	DINR	DR	NB	DPAR	Poi
0	43	45.107	41.37	43.84	38.28	15.933	30.122	64.45	27.22
1	35	30.082	41.85	39.61	51.90	36.13	38.866	20.16	40.39
2	17	18.827	15.42	15.62	15.51	34.575	27.611	9.693	29.96
3	11	11.297	7.170	7.201	6.04	21.023	14.262	5.655	14.83
4	5	6.567	3.936	3.910	2.913	8.888	5.988	3.678	5.490
5	4	3.718	2.400	2.371	1.611	2.702	2.171	2.577	1.630
6	1	2.058	1.614	1.564	0.977	0.600	0.700	1.900	0.400
7	2	1.116	1.130	1.095	0.638	0.090	0.210	1.460	0.090
8	2	0.594	5.090	4.807	2.141	0.020	0.060	10.44	0.020
-ℓ		200.438	204.810	204.293	208.440	235.23	211.52	220.63	219.19
AIC		408.875	413.621	412.587	418.881	472.45	427.05	443.24	440.38
CAIC		409.223	413.723	412.689	418.915	472.49	427.14	443.27	440.41
χ^2_{ν}		1.088	5.511	4.664	14.274	70.688	20.367	32.462	38.478
d.f		1	3	3	4	4	3	4	4
p_v		0.297	0.138	0.198	< 0.0010	< 0.0010	0.0010	< 0.0010	< 0.0010

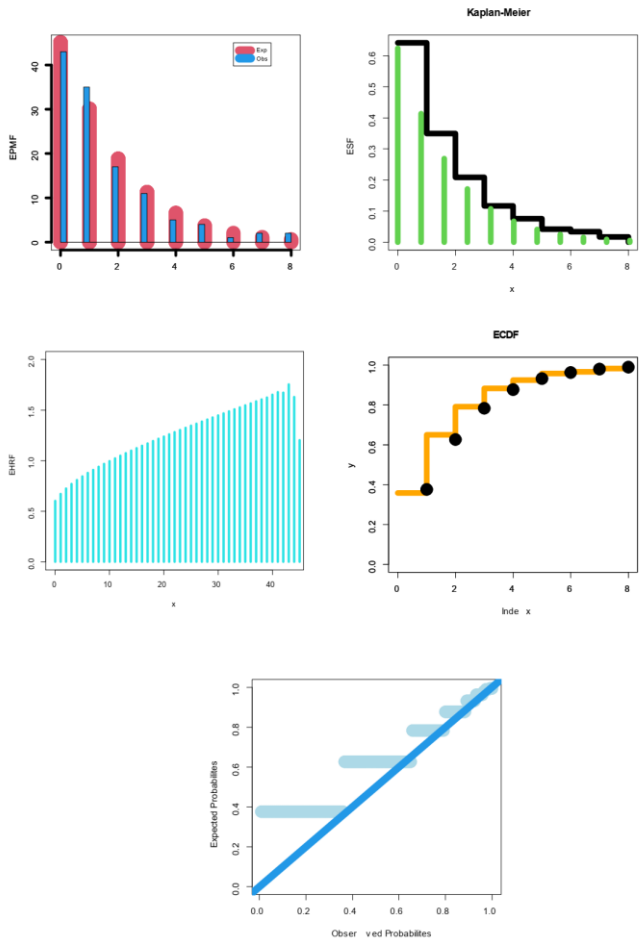


Figure 10: The EPMF, ESF, EHRF, ECDF, PP graphs for corn-borer larvae dataset.

9. Application in biological science

Genetics, a branch of biological science, investigates the mechanisms responsible for both similarities and differences within closely related species. The term "genetic" originates from the Greek word "genesis," signifying growth into or becoming. Thus, genetics is the examination of heredity and the variations inherent in it. This scientific discipline focuses on the transmission of physical traits—both similarities and differences—from parents to offspring, along with the laws governing this process. Any disparities among individual organisms or groups within a species, stemming from either genetic distinctions or the influence of environmental factors, are termed variations. These variations can manifest in diverse aspects such as physical appearance, metabolism, behavior, learning capabilities, mental aptitude, and other observable traits.

Recently, Hassan et al. (2020) presented a new discrete distribution with applications to two count biological data sets. Figure 11 (the first three plots) gives the Box graph, Q-Q graph and TTT graph for the first count biological dataset. Figure 11 (the second three plots) gives the Box graph, Q-Q graph and TTT graph for the second count biological dataset. Based on the box plot and the Q-Q plots, it is noted that the two data sets have some extreme values. Based on the TTT plot, it is noted that the two data sets have a decreasing HRF.

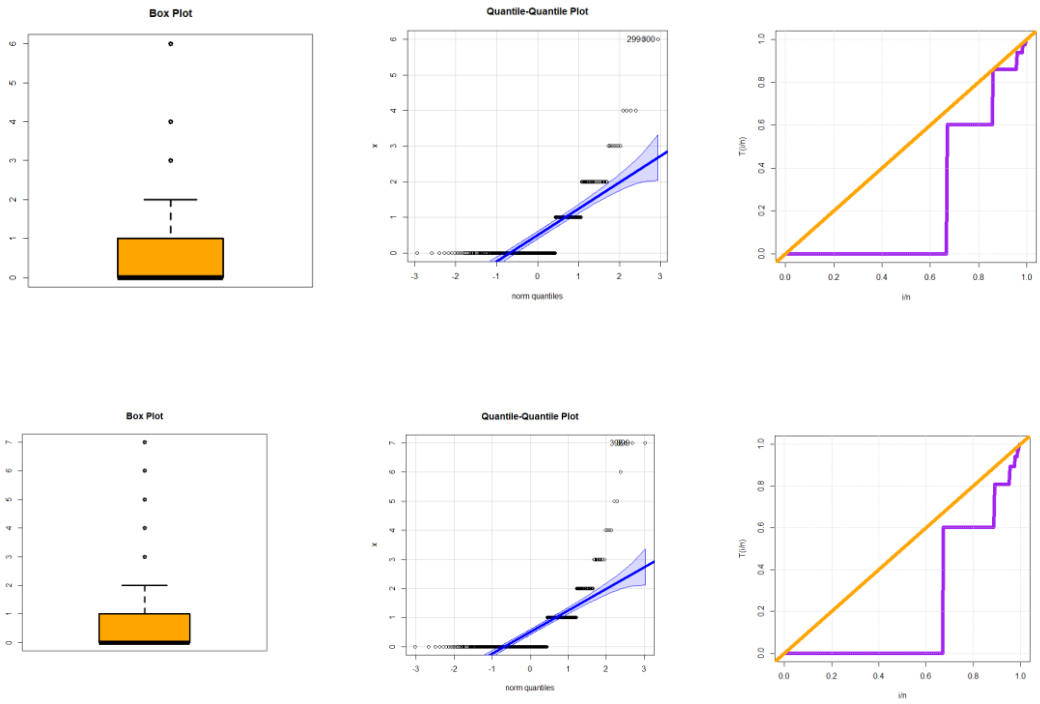


Figure 11: The Box graph, Q-Q graph and TTT graph for the two count biological datasets.

Table 19: The MLKEs and the goodness of fit statistics for first count biological data.

Number of Aberrations	OF	Fitted Distribution				
		Poi	PAD	DWGW	DRW	DWW
0	268	231.3	246.1	270.876	272.308	272.701
1	87	126.1	106.1	78.193	74.276	75.34
2	26	34.7	34.3	29.602	30.688	29.974
3	9	6.3	9.8	12.068	13.114	12.586
4	4	0.8	2.6	5.136	5.603	5.379
5	2	0.1	0.6	2.251	2.367	2.308
6	1	0.1	0.1	1.009	0.983	0.988
7+	3	0.1	0.04	0.46	0.401	0.42
Total	400	400	400	400	400	400
$\chi^2_{\hat{\nu}}$		40.8	10.05	2.4	0.68077	0.68175
MLKEs		$\hat{\theta}$=0.55	$\hat{\alpha}$=1.82	$\hat{p}=0.00051$ $\widehat{\beta_1}=32.892$ $\widehat{\beta_2}=0.66463$ $\hat{\theta}=0.0192$	$\hat{p}=0.6793$ $\hat{\theta}=0.2490$	$\hat{p}=0.00062$ $\hat{\lambda}=0.04862$ $\hat{\beta}=1.7643$ $\hat{\theta}=0.40002$
$\mathcal{P}_v$		<0.0001	0.018	0.12221	<0.0001	<0.0001

In this application, the new distribution was harnessed within the framework of count biological data modeling, and the new distribution was compared with other new competing and important distributions such as the Poisson distribution, Poisson Ailamujia distribution, discrete Rayleigh Weibull (DRW) distribution and discrete Weibull-Weibull (DWW) distribution. Table 19 gives the MLKEs and the goodness of fit statistics for first count biological data. Table 20 provides the MLKEs and the goodness of fit statistics for second count biological data. Based on Table

19, the DWGW provides the best results with $\mathscr{p}_v$=0.12221. However, Based on Table 20, the DWGW provides the best results with $\mathscr{p}_v$=0.13129.

Table 20: The MLKEs and the goodness of fit statistics for second count biological data.

Number of Aberrations	OF	Fitted Distribution				
		Poi	PAD	DWGW	DRW	DWW
0	200	172.5	184.6	200.583	199.558	199.357
1	57	95.4	79.5	59.104	60.215	60.78
2	30	26.4	25.7	24.341	24.382	24.046
3	7	7.3	4.9	9.957	9.804	9.686
4	4	1.9	0.7	3.889	3.823	3.836
5	0	.5	0.1	1.427	1.436	1.471
6	2	0.1	0.0	0.487	0.518	0.542
Total	300	300	300	300	300	300
χ^2_v		29.6	9.4	2.27721	0.66519	0.6645
MLKEs		$\hat{\theta}$=0.55	$\hat{\alpha}$=1.82	$\hat{p}=0.0731$ $\widehat{\beta_1}=1.1552$ $\widehat{\beta_2}=0.3881$ $\hat{\theta}=0.5924$	$\hat{p}=0.6903$ $\hat{\theta}=0.26603$	$\hat{p}=0.0006$ $\hat{\lambda}=0.0846$ $\hat{\beta}=1.0359$ $\hat{\theta}=0.7772$
$\mathscr{p}_v$		<0.0001	0.09	0.13129	<0.0001	<0.0001

For illustrating the high fitting of the new model, we present Figure 12 which contains six plots. Figure 12 (the first three plots) gives the EPMF, ESF, and PP graphs for the first count biological datasets, Figure 12 (the second three plots) gives the EPMF, ESF, and PP graphs for the second count biological datasets.

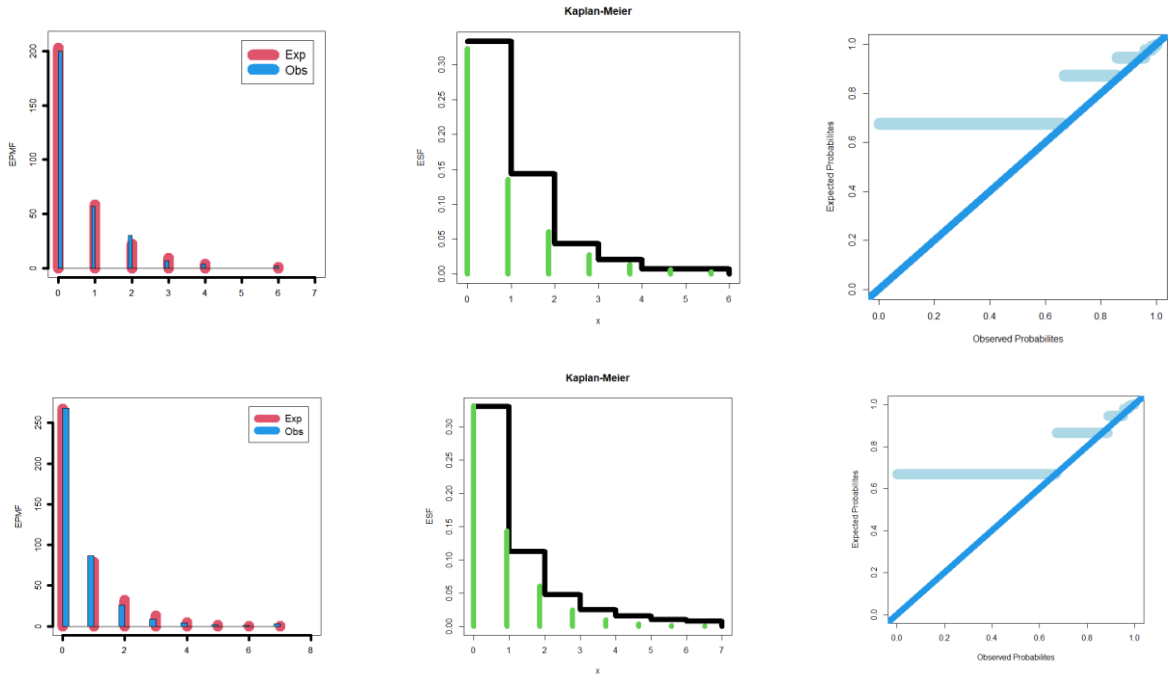


Figure 12: The EPMF, ESF, and PP graphs for two count biological datasets.

10. Concluding remarks

We proposed and investigated a novel discrete family of distributions that we dubbed the discrete Weibull generated G (DWG-G) family of distributions in light of the continuous G family that was presented by Yousof et al. (2018). Numerous helpful statistical features are developed and evaluated, some examples of which are the probability generating function, ordinary moments, central moment, moment generating function, cumulant generating function, and dispersion index (DisIx). The DWGW model, which is a specific discrete member of the DWG-G family based on the standard Weibull model, is studied graphically, theoretically, and numerically. "Increasing-constant," "decreasing-constant-increasing (U)," "constant," "U-constant," "decreasing," and "J-HRF" are the possible values for the hazard rate function of the DWGW model. The probability mass function of the DWGW model can take on a variety of useful shapes, including "asymmetric left skewed," "right skewed with wide peak," "right skewed," "bimodal," "symmetric," and "right skewed." Other shapes include "symmetric" and "bimodal".

Traditional approaches to estimation were taken into consideration and utilized. The Bayesian estimating approach with the squared error loss function was another one of the methods that were taken into consideration. Both the Gibbs sampler and the Metropolis Hastings algorithm were used and put into practice. In order to evaluate Bayesian estimate in comparison to other methods, more conventional methods, simulations built with the Markov Chain Monte Carlo method are run. Despite this, the accuracy of any estimation method increases in parallel with the total number of observations as they are added. On the other hand, increasing the size of the sample is a straightforward way to enhance the accuracy of any and all estimation techniques. On the other hand, improving the performance of any and all ways of estimating results in improved performance when the sample size is increased. Nevertheless, increasing the size of the sample population results in an improvement in the performance of all estimating methods. The maximum likelihood estimation strategy is the best, with the lowest value of mean squared errors for any sample size; however, increasing the sample size also enhances the performance of other estimation strategies. The maximum likelihood estimation strategy is the best, with the least amount of mean squared errors for any sample size.

In addition, six distinct collections of actual datasets were utilized in order to examine and contrast the Bayesian and traditional estimation strategies. For the purpose of illustrating how flexible the DWGW model is, six distinct collections of previously collected data were used. The DWGW distribution provided a more accurate fit than any of the other sixteen distributions that were in the running.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Dataset Availability Statement: The dataset can be provided upon requested.

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

References

1. Aboraya, M., M Yousof, H., Hamedani, G. G. and Ibrahim, M. A New Family of Discrete Distributions with Mathematical Properties, Characterizations, Bayesian and Non-Bayesian Estimation Methods. Mathematics, 2020, 8(10), 1648.
2. Alizadeh, M., Cordeiro, G. M., Pinho, L. G. B., and Ghosh, I. The Gompertz-G family of distributions. Journal of Statistical Theory and Practice, 2017, 11(1), 179-207.
3. Alizadeh, S. M., Bagheri, S. F., Bahrami Samani, E., Ghobadi, S. and Nadarajah, S. (2018). Exponentiated power Lindley power series class of distributions: Theory and applications. Communications in Statistics-Simulation and Computation. 47(9), 2499-2531
4. Bahrami Samani, E. and Ganjali, M. (2014). Mixed correlated bivariate ordinal and negative binomial longitudinal responses with nonignorable missing values. Communications in Statistics-Theory and Methods. 43 (13), 2659-2673
5. Bahrami Samani, E., Ganjali, M. and Amirian, Y. (2012). Zero-inflated power series joint model to analyze count dataset with missing responses. Journal of Statistical Theory and Practice 6, 334-343
6. Bebbington, M.; Lai, C.D. Wellington, M.; Zitikis, R. The discrete additive Weibull distribution: A bathtub-shaped hazard for discontinuous failure dataset. Reliab. Eng. Syst. Saf. 2012, 106, 37-44.
7. 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.
8. Bourguignon, M.; Silva, R.B.; Cordeiro, G.M. The Weibull-G family of probabilistic distributions. J. Dataset Sci. 2014, 12, 53-68.
9. 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.

10. 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.
11. Gomez-Déniz, E. Another generalization of the geometric distribution. Test 2010, 19, 399-415.
12. Gomez-Déniz, E.; Caldern-Ojeda, E. The discrete Lindley distribution: Properties and applications. J. Stat. Comput. Simul. 2011, 81, 1405-1416.
13. Hassan, A., Shalbaf, G. A., Bilal, S., & Rashid, A. (2020). A new flexible discrete distribution with applications to count data. Journal of Statistical Theory and Applications, 19(1), 102-108.
14. Hesterberg, T. Bootstrap, Wiley Interdisciplinary Reviews: Computational Statistics, 2011, 3(6), 497-526.
15. Hussaine, T.; Aslam, M.; Ahmad, M. A Two parameter discrete Lindley distribution. Rev. Colomb. Estadistica 2016, 39, 45-61.
16. Jazi, A.M.; Lai, D.C.; Alamatsaz, H.M. Inverted Weibull distribution and estimation of its parameters. Stat. Methodol. 2010, 7, 121-132.
17. Kemp, AW: Classes of discrete lifetime distributions. Commun. Stat. Theor. Methods. 33(12), 3069-3093 (2004).
18. Lawless, J.F. Statistical Models and Methods for Lifetime Dataset; Wiley: New York, NY, USA, 2003.
19. Nadarajah,S., Bagheri, S. F., Alizadeh, M. and Bahrami Samani, E. (2018). Estimation of the stress strength parameter for the generalized exponential-Poisson distribution. Journal of Testing and Evaluation. 46(5), 1-28.
20. Nakagawa, T.; Osaki, S. The discrete Weibull distribution. IEEE Trans. Reliab. 1975, 24, 300—301
21. Para, B.A.; Jan, T.R. Discrete version of loglogistic distribution and its applications in genetics. Int. J. Mod. Math. Sci. 2016, 14, 407-422.
22. Para, B.A.; Jan, T.R. On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count dataset from medical science. Biom. Biostat. Int. J. 2016, 4, 1-15.
23. 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.
24. Razie,F., Bahrami Samani, E. and Ganjali, M. (2016). A latent variable model for analyzing mixed longitudinal (k,l)-inflated count and ordinal responses. Journal of Applied Statistics 43 (12), 2203-2224
25. Roy, D. The discrete normal distribution. Commun. Stat. Theor. Methods. 2003, 32(10), 1871-1883.
26. Roy, D. Discrete Rayleigh distribution. IEEE Trans. Reliab. 2004, 53, 255-260.
27. Nakagawa, T. and Osaki, S. The discrete Weibull distribution, IEEE Transactions on Reliability, 1975, 24(5), 300-301.
28. Yousof, H. M., Majumder, M., Jahanshahi, S. M. A., Masoom Ali, M. and Hamedani, G. G. (2018). A new Weibull class of distributions: theory, characterizations and applications. Journal of Statistical Research of Iran JSRI, 15(1), 45-82.
29. Tabrizi, E., Bahrami Samani, E. and Ganjali, M. (2020). A note on the identifiability of latent variable models for mixed longitudinal dataset. Statistics & Probability Letters 167, 108882.