Filomat 39:26 (2025), 9249–9274
https://doi.org/10.2298/FIL2526249S

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

A discrete odd exponentiated half-logistic-G class: Mathematical and statistical theory with Goodness-of-fit dispersion data analysis across varied failure profiles

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

[a]*Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia.*
[b]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*
[c]*Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia.*

Abstract. The paper introduces a novel discrete probability class specifically designed to extend the odd exponentiated half-logistic-G family, providing a flexible framework for generalizing various discrete baseline models. The study begins with the formulation of the new discrete class, followed by an in-depth analysis of a particular discrete model within this framework. A comprehensive investigation of its mathematical and statistical properties is conducted, covering key characteristics such as the probability mass function, hazard rate function, crude moments, index of dispersion, entropy measures, order statistics, and L-moments. The findings demonstrate the model's ability to effectively capture both symmetric and asymmetric data distributions while accommodating a broad range of kurtosis structures. Additionally, the proposed class proves adept at addressing overdispersion and underdispersion in datasets with outliers, as well as modeling diverse hazard rate patterns, including monotonically increasing and decreasing trends, bathtub-shaped, unimodal-bathtub, J-shaped, inverse J-shaped, and other complex forms. Parameter estimation for the proposed class is carried out using the maximum likelihood method, with the accuracy and efficiency of the estimation process assessed through Markov chain Monte Carlo simulations. To validate its practical applicability, the new probability generator is applied to three real-world datasets, demonstrating its robustness and effectiveness in capturing real data patterns.

1. Introduction

In count data modeling, discrete probability distributions play a crucial role in representing the discrete outcomes of underlying continuous-time processes. Observed count data are often treated as discrete

2020 *Mathematics Subject Classification*. Primary 62E99; Secondary 62E15.
Keywords. Discrete random variables, Failure analysis; Esitmation theory; Extreme count datasets; Simulation; Data analysis and decision making.
Received: 19 August 2024; Revised: 09 March 2025; Accepted: 28 May 2025
Communicated by Biljana Popović
This project was supported by the deanship of scientific research at Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz Universityproject number (PSAU/2025/R/1446).
* Corresponding author: Mahmoud El-Morshedy
Email addresses: hendsalah@mans.edu.eg (Hend S. Shahen), mseliwa@mans.edu.eg (Mohamed S. Eliwa), mah_elmorshedy@mans.edu.eg (Mahmoud El-Morshedy)
ORCID iDs: https://orcid.org/0000-0001-5619-210X (Mohamed S. Eliwa), https://orcid.org/0000-0002-7585-5519 (Mahmoud El-Morshedy)

representations of continuous phenomena. For instance, the number of customer arrivals at a service center can be modeled as discrete instances of a continuous Poisson process, while the spread of infectious diseases can be viewed as discrete outcomes stemming from an evolving epidemiological process. Traditional discrete distributions such as the Poisson, negative binomial, and geometric models are commonly used in these cases. However, these models frequently fail to capture the complexities encountered in real-world scenarios.

A primary limitation of the Poisson distribution, for example, is its assumption of equidispersion, where the mean and variance are equal. In practice, many empirical datasets exhibit overdispersion (where the variance exceeds the mean) or underdispersion (where the variance is less than the mean). These discrepancies highlight the need for more flexible models that can accommodate varying dispersion levels. To address such limitations, statisticians have developed advanced discretization techniques. Methods such as the infinite series method, survival-based discretization, reversed hazard function discretization, and the compound two-phase method have been introduced to enhance the flexibility of count data modeling (see Chakraborty, 2015). Among these, survival-based discretization is often favored due to its simplicity and effectiveness in generating discrete distributions that better align with empirical data. To elaborate further, if we consider a continuous random variable X with a survival function (SF) denoted as $S(x,.) = \Pr(X > x)$, the probability mass function (PMF) of its corresponding discrete random variable X, which takes non-negative integer values, can be expressed as

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

where x represents the specific integer values within the support set of the discrete random variable (see Roy, 2004). This approach results in a discrete distribution that mirrors the functional characteristics of the SF associated with the continuous distribution. Examples of models derived through this technique include the discrete Weibull (Nakagawa and Osaki, 1975), discrete inverse Weibull (Jazi et al., 2010), discrete Lindley (Gómez-Déniz and Calderín-Ojeda, 2011), discrete Gompertz-G family (Eliwa et al., 2020), discrete generalized odd Lindley-Weibull (Aryuyuen et al., 2020), discrete pseudo Lindley (Irshad et al., 2021), discrete Marshall-Olkin inverted Topp-Leone (Almetwally, 2022), discrete generalized Pareto (Haj Ahmad and Almetwally, 2022), discrete DAL generator (Eliwa et al., 2023), discrete exponentiated generalized class (Abd EL-Hady, et al., 2023), discrete half-logistic (Teamah, 2024), a novel version of geometric model (Alosey and Gemeay, 2025), and others. As research advances, the significance of discrete probability models remains unquestionable. These models are essential for identifying complex patterns in data, refining statistical inference, and advancing predictive modeling across diverse fields. Additionally, understanding the link between discrete and continuous models has led to the development of hybrid approaches that leverage the strengths of both frameworks. These hybrid methods not only enhance model interpretability but also provide computational efficiency, ensuring that data complexity across various domains is effectively managed.

This paper presents a discrete version of the odd exponentiated half-logistic-G family, developed using the survival discretization method. The motivation for this approach lies in the flexibility of the exponentiated half-logistic (EHL) distribution, which excels at modeling and analyzing various patterns in real-world datasets across different fields. Many researchers have further expanded the EHL model, exploring its applications in diverse contexts. Notable extensions include the exponentiated half-logistic family (Cordeiro et al., 2014), odd exponentiated half-logistic-G family (Afify et al., 2017), Type II exponentiated half-logistic generated family (Al-Mofleh et al., 2020), bivariate exponentiated half-logistic (Alotaibi et al., 2021), Type II exponentiated half-logistic Topp-Leone Marshall-Olkin-G family (Moakofi et al., 2021), Type II exponentiated half-logistic-Topp-Leone-G power series class (Moakofi et al., 2021), unit-exponentiated half-logistic (Hassan et al., 2022), exponentiated half-logistic skew-t (Adubisi et al., 2022), Type II exponentiated half-logistic Gompertz-G family (Moakofi and Oluyede, 2023), Topp-Leone Type II exponentiated half-logistic-G family (Gabanakgosi and Oluyede, 2023), exponentiated half-logistic (El-Awady et al., 2024), among others. The development of the discrete odd exponentiated half-logistic-G (DOEHL-G) class is motivated by several key factors that enhance its adaptability and usefulness in statistical modeling:

- Versatility in Count Data Modeling: The DOEHL-G class is designed to accommodate both symmetric and asymmetric count data, making it particularly useful for datasets with zero-inflated count observations.

- Handling Outliers: It effectively models not only moderate outliers but also extreme outlier count observations, ensuring robustness in practical applications.
- Overdispersion and Underdispersion Analysis: This class serves as a valuable probabilistic tool for investigating overdispersion and underdispersion in scattered data, which is essential for many real-world scenarios.
- Kurtosis Variability: The DOEHL-G class enables the analysis of different kurtosis shapes, including leptokurtic (peaked) and platykurtic (flat) distributions, allowing for greater adaptability in data characterization.
- Generation of New Flexible Models: By leveraging the underlying distribution, this framework facilitates the development of novel discrete probability models with enhanced flexibility.
- Comprehensive Failure Rate Modeling: It is capable of capturing a wide range of failure rate behaviors, such as monotonic (increasing or decreasing), bathtub-shaped, unimodal-bathtub, J-shaped, and inverse J-shaped failure rates, making it highly applicable in reliability analysis.
- Heavy-Tailed Data Suitability: The DOEHL-G class is particularly well-suited for modeling datasets with heavy tails, which commonly arise in various fields, including finance, insurance, and environmental studies.

The article is structured as follows for an engaging and detailed exploration: Section 2 unveils the mathematical framework of the DOEHL-G class, delving into its structural intricacies. Section 3 lists a comprehensive analysis of various distributional statistics associated with the DOEHL-G family. In Section 4, a specific sub-model within the DOEHL-G is meticulously examined, highlighting its unique characteristics. The estimation of DOEHL-G family parameters utilizing the maximum likelihood approach is elaborated upon in Section 5, showcasing the precision of the technique. Section 6 features a thorough simulation study, demonstrating the practical applicability and performance of the DOEHL-G class. To showcase the versatility and effectiveness of the DOEHL-G class, Section 7 conducts in-depth analyses on three real-world datasets, illustrating its remarkable capabilities. Finally, Section 8 consolidates the findings and insights gained throughout the article, offering conclusive remarks and potential avenues for future research.

2. The Mathematical Framework of the DOEHL-G Family

Afify et al. (2017) presented the odd exponentiated half-logistic-G (OEHL-G) family, emphasizing its mathematical and statistical properties. The CDF of the OEHL-G class can be expressed as

$$W_Y(y;\varepsilon,\alpha,\mathbf{\Phi}) = \left(\frac{1-e^{-\varepsilon\frac{G(y;\mathbf{\Phi})}{1-G(y;\mathbf{\Phi})}}}{1+e^{-\varepsilon\frac{G(y;\mathbf{\Phi})}{1-G(y;\mathbf{\Phi})}}}\right)^{\alpha};\ y\in\mathbb{R}, \tag{2}$$

where $\mathbf{\Phi}$ represents a parameter vector, with $\varepsilon > 0$ and $\alpha > 0$. Based on Equations (1) and (2), the CDF of the DOEHL-G family is given by

$$F_X(x;\alpha,p,\mathbf{\Phi}) = \left(\frac{1-p^{Q(x+1;\mathbf{\Phi})}}{1+p^{Q(x+1;\mathbf{\Phi})}}\right)^{\alpha};\ x\in\mathbb{N}_0, \tag{3}$$

where $Q\left(x;\mathbf{\Phi}\right) = \frac{G(x;\mathbf{\Phi})}{1-G(y;\mathbf{\Phi})}$, $0 < p < 1$ where $p = e^{-\varepsilon}$ and $\mathbb{N}_0 = 0, 1, 2, 3, ...$ The associated probability mass function (PMF) of Equation (3) can be expressed as

$$P_X\left(x;\alpha,p,\mathbf{\Phi}\right) = \left(\frac{1-p^{Q(x+1;\mathbf{\Phi})}}{1+p^{Q(x+1;\mathbf{\Phi})}}\right)^{\alpha} - \left(\frac{1-p^{Q(x;\mathbf{\Phi})}}{1+p^{Q(x;\mathbf{\Phi})}}\right)^{\alpha};\ x\in\mathbb{N}_{0,} \tag{4}$$

where p is a scale parameter and α is a shape parameter. The hazard rate function (HRF) is given by

$$h(x;\alpha,p,\mathbf{\Phi})=\left[\left(\frac{1-p^{Q(x+1;\mathbf{\Phi})}}{1+p^{Q(x+1;\mathbf{\Phi})}}\right)^{\alpha}-\left(\frac{1-p^{Q(x;\mathbf{\Phi})}}{1+p^{Q(x;\mathbf{\Phi})}}\right)^{\alpha}\right]\left[1-\left(\frac{1-p^{Q(x;\mathbf{\Phi})}}{1+p^{Q(x;\mathbf{\Phi})}}\right)^{\alpha}\right]^{-1};\ x\in\mathbb{N}_0. \tag{5}$$

The HRF for discrete random variables plays a significant role in various applications. In epidemiology, the HRF is employed to analyze disease progression in discrete time intervals, aiding in the estimation of disease transmission rates and the effectiveness of interventions. In reliability engineering, it is used to model the failure rates of discrete systems over time. For instance, in the context of electronic components, the HRF helps predict the likelihood of component failure at different stages of operation, guiding decisions on maintenance schedules and replacement strategies. Further, in actuarial science, the HRF assists in calculating insurance premiums for discrete events such as accidents or policyholder mortality, by quantifying the risk associated with these events at different time points. Another statistical concept closely related to reliability theory is known as the reversed hazard rate function (RHRF). Specifically for the new generator, the expression of the RHRF is as follows

$$r(x;\alpha,p,\mathbf{\Phi})=1-\left(\frac{\left(1-p^{Q(x;\mathbf{\Phi})}\right)\left(1+p^{Q(x+1;\mathbf{\Phi})}\right)}{\left(1+p^{Q(x;\mathbf{\Phi})}\right)\left(1-p^{Q(x+1;\mathbf{\Phi})}\right)}\right)^{\alpha};\ x\in\mathbb{N}_0. \tag{6}$$

The RHRF holds significant value in the analysis of censored data and is widely applicable across various fields. Its relevance is especially notable in forensic studies, where it plays a crucial role in determining the time to an event or failure, essential for investigative purposes. Moreover, in actuarial sciences, the RHRF contributes to estimating remaining life expectancies accurately and crafting insurance products that align with individual risk profiles, ensuring effective risk management strategies.

3. Some Statistical Properties

3.1. The moment generating function and r^{th} moments

The moment generating function (MGF) and r^{th} moments provide a comprehensive understanding of the characteristics and behavior of random variables, aiding in statistical analysis, modeling, and inference. The MGF, say $M_X(t)$, of a random variable X is defined as $M_X(t)=E(e^{Xt})$, where t is a parameter. The MGF uniquely determines the distribution of X if it exists in a neighborhood of $t=0$. It is a powerful tool for deriving moments of X since $M_X^{(r)}(0)=\mu_r'$, where $M_X^{(r)}(t)$ denotes the r^{th} derivative of $M_X(t)$ with respect to t. If we suppose that a random variable X follows a DOEHL-G distribution, then the MGF can be expressed as

$$M_X(t)=\sum_{x=0}^{\infty}\sum_{m=0}^{\infty}\frac{(xt)^m}{m!}\left[\left(\frac{1-p^{Q(x+1;\mathbf{\Phi})}}{1+p^{Q(x+1;\mathbf{\Phi})}}\right)^{\alpha}-\left(\frac{1-p^{Q(x;\mathbf{\Phi})}}{1+p^{Q(x;\mathbf{\Phi})}}\right)^{\alpha}\right];x\in\mathbb{N}_0. \tag{7}$$

Alternatively, based on Equation (1), the r^{th} moment of X can be expressed as

$$\mu_r'=E(X^r)=\sum_{x=0}^{\infty}x^r\Pr(X=x)=\sum_{x=0}^{\infty}x^r[S(x)-S(x+1)],$$

using summation by parts (discrete integration by parts) along with the properties of the SF, particularly that $S(0)=1$ and $S(\infty)=0$, the r^{th} moment can be reported as

$$\begin{aligned}\mu_r'&=\sum_{x=1}^{\infty}[x^r-(x-1)^r]\,S(x)\\&=\sum_{x=1}^{\infty}[x^r-(x-1)^r]\left(\frac{1-p^{Q(x;\mathbf{\Phi})}}{1+p^{Q(x;\mathbf{\Phi})}}\right)^{\alpha};\ r=1,2,3,....\end{aligned} \tag{8}$$

For more details on summation by parts, refer to the publications of Davis and Polonsky (1965) and Riley (2006). The transformed sum expresses the moment as a sum over the SF weighted by the difference $x^r - (x-1)^r$. This formulation is useful in cases where the survival function is easier to work with than the PMF, such as in reliability analysis or actuarial science. Using Equation (8), the mean "$E(X)$", variance "$Var(X)$", skewness "$Sk(X)$" and kurtosis "$Ku(X)$" can be determined as follows

$$E(X) = \sum_{x=1}^{\infty} - \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha}, \; Var(X) = \sum_{x=1}^{\infty} (1-2x) \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha} - \mu_1^{'2},$$

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

The dispersion index (DsI), as determined from the mean and variance measures, serves as a metric for assessing potential over-dispersion or under-dispersion within the dataset under consideration. The expression for the DsI of the DOEHL-G family can be formulated as

$$DsI\,(X) = \frac{\sum_{x=1}^{\infty} (2x-1) \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha}}{\sum_{x=1}^{\infty} \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha}} + \sum_{x=1}^{\infty} \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha}. \tag{9}$$

A DsI value greater than (less than) 1 suggests over-dispersion (under-dispersion), whereas a DsI of 1 indicates equi-dispersion.

3.2. Rényi and Shannon entropies

Rényi and Shannon entropies refer to measures of uncertainty or information content in a system, with Rényi entropy being a generalized form of Shannon entropy. The Rényi and Shannon entropies for the DOEHL-G family can be listed as

$$I_{\omega}\,(X) = \frac{1}{1-\omega} \log \sum_{x=0}^{\infty} \left[\left(\frac{1 - p^{Q(x+1;\mathbf{\Phi})}}{1 + p^{Q(x+1;\mathbf{\Phi})}} \right)^{\alpha} - \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha} \right]^{\omega}; \; x \in \mathbb{N}_0,$$

and

$$I\,(X) = - \sum_{x=0}^{\infty} \left[\left(\frac{1 - p^{Q(x+1;\mathbf{\Phi})}}{1 + p^{Q(x+1;\mathbf{\Phi})}} \right)^{\alpha} - \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha} \right] \log \left[\left(\frac{1 - p^{Q(x+1;\mathbf{\Phi})}}{1 + p^{Q(x+1;\mathbf{\Phi})}} \right)^{\alpha} - \left(\frac{1 - p^{Q(x;\mathbf{\Phi})}}{1 + p^{Q(x;\mathbf{\Phi})}} \right)^{\alpha} \right],$$

respectively, where $\omega \in (0,1)$ and $\omega \neq 1$. The Shannon entropy can be derived as a specific instance of the Rényi entropy when the parameter ω approaches 1.

3.3. Order Statistics

Consider a set of n random variables, denoted as $X_1, X_2, ..., X_n$ arranged in ascending order and labeled as $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$. In the context of order statistics, there's no requirement for the X_i's to be independent or identically distributed (IID). However, numerous established findings regarding order statistics are derived under the classical assumption of independence and IID among the X_i's. The CDF of the i^{th} order statistic is given by

$$\begin{aligned} F_{i:n}\,(x;\alpha,p,\mathbf{\Phi}) &= \sum_{k=i}^{n} \binom{n}{k} [F_i\,(x;\alpha,p,\mathbf{\Phi})]^k \, [1 - F_i\,(x;\alpha,p,\mathbf{\Phi})]^{n-k} \\ &= \sum_{k=i}^{n} \sum_{j=0}^{n-k} \Psi_{(n,k)}^{(j)} F_i\,(x;\alpha\,(k+j)\,,p,\mathbf{\Phi})\,, \end{aligned} \tag{10}$$

where

$$\Psi_{(n,k)}^{(j)} = (-1)^j \binom{n}{k}\binom{n-k}{j}.$$

Further, the corresponding PMF of the i^{th} order statistic can be expressed as

$$\begin{aligned} f_{i:n}(x;\alpha,p,\mathbf{\Phi}) &= F_{i:n}(x;\alpha,p,\mathbf{\Phi}) - F_{i:n}(x-1;\alpha,p,\mathbf{\Phi}) \\ &= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\Psi_{(n,k)}^{(j)}\left(F_i(x;\alpha(k+j),p,\mathbf{\Phi}) - F_i(x-1;\alpha(k+j),p,\mathbf{\Phi})\right). \end{aligned} \tag{11}$$

Thus, the r^{th} moments of $X_{i:n}$ can be reported as

$$E\left(X_{i:n}^r\right) = \sum_{x=0}^{\infty}\sum_{k=i}^{n}\sum_{j=0}^{n-k}\Psi_{(n,k)}^{(j)}x^r\left(F_i(x;\alpha(k+j),p,\mathbf{\Phi}) - F_i(x-1;\alpha(k+j),p,\mathbf{\Phi})\right). \tag{12}$$

L-moments serve as summary statistics for probability distributions, akin to ordinary moments but derived from linear functions of the ordered data values. The L-moment of X can be expressed as

$$\Theta_\delta = \frac{1}{\delta}\sum_{i-0}^{\delta-1}(-1)^i\binom{\delta-1}{i}E\left(X_{\delta-i:\delta}\right);\ \delta = 1,2,3,... \tag{13}$$

Using Equation (13), several statistical measures can be computed based on L-moment statistics, including L-mean = Θ_1, L-skewness = $\frac{\Theta_3}{\Theta_2}$, L-kurtosis = $\frac{\Theta_4}{\Theta_2}$, and L-coefficient of variation = $\frac{\Theta_2}{\Theta_1}$.

4. Special model: The DOEHL-Weibull distribution

The Weibull distribution is a versatile probability distribution commonly used in reliability engineering, survival analysis, and various other fields. The distribution is characterized by its flexibility in modeling a wide range of behaviors, making it a valuable tool in statistical analysis. The CDF of Weibull model can be expressed as

$$G(x;a,b) = 1 - e^{-ax^b};\ x > 0, \tag{14}$$

where $a > 0$ represents a scale parameter, while $b > 0$ denotes a shape parameter. Then, the PMF of the DOEHL-Weibull (DOEHL-W) distribution can be reported as

$$P_X(x;\alpha,p,\theta,b) = \left(\frac{1-p^{\left(\theta^{-(x+1)^b}-1\right)}}{1+p^{\left(\theta^{-(x+1)^b}-1\right)}}\right)^{\alpha} - \left(\frac{1-p^{\left(\theta^{-x^b}-1\right)}}{1+p^{\left(\theta^{-x^b}-1\right)}}\right)^{\alpha};\ x \in \mathbb{N}_0, \tag{15}$$

where $\theta = e^{-a}$ and $0 < \theta < 1$. Figures 1 and 2 illustrate the PMF and HRF of the DOEHL-W model, respectively. Notably, the PMF of the DOEHL-W model serves as a valuable probability tool for analyzing and discussing data with asymmetric or symmetric characteristics, whether unimodal or bimodal. Additionally, the HRF can exhibit a range of shapes including monotone increasing, monotone decreasing, bathtub, unimodal-bathtub, J-shaped, and inverse J-shaped. The diverse shapes of the PMF and HRF in the DOEHL-W model make it a preferred choice for researchers in statistical fields, enabling the modeling and analysis of various datasets across different domains.

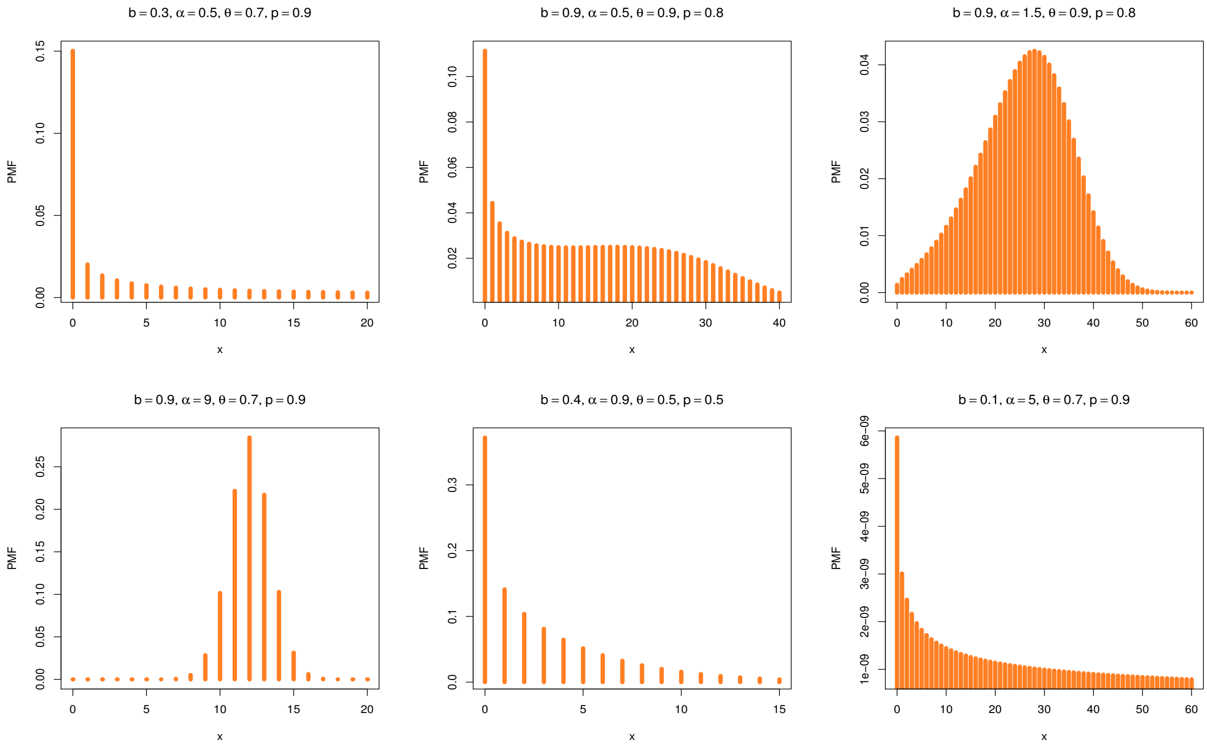


Figure 1. The PMF plots of the DOEHL-W distribution.

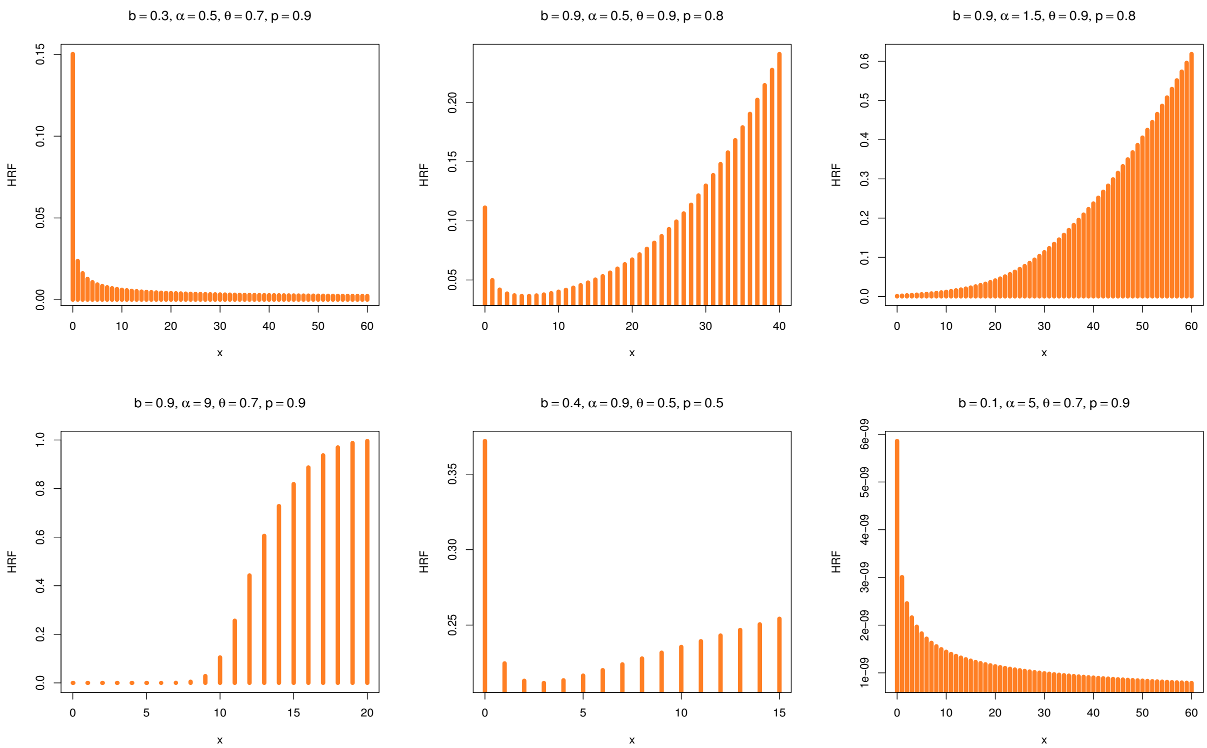


Figure 2. The HRF plots of the DOEHL-W distribution.

Tables 1 and 2 present computations of central tendency measures for the DOEHL-W distribution across various model parameter values, while Tables 3 and 4 list numerical values for the entropy.

Table 1. Numerical descriptive statistics of the DOEHL-W distribution at b=0.3, α=0.2.

Measure	$\theta \downarrow p \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	0.3	0.0020	0.0115	0.0338	0.0778	0.1592	0.3111	0.6132	1.3079	3.5682
$E(X)$	0.5	0.0819	0.2284	0.4650	0.8442	1.4666	2.5327	4.4266	7.6164	11.200
	0.9	5.5123	5.5764	5.4563	5.2693	5.0421	4.7778	4.4663	4.0768	3.5131
	0.3	0.0024	0.0170	0.0633	0.1870	0.5034	1.3342	3.7377	12.307	62.576
$Var(X)$	0.5	0.3001	1.3372	3.9801	10.229	24.928	60.218	143.94	316.81	542.96
	0.9	282.49	291.24	288.11	280.53	270.32	257.85	242.68	223.23	194.41
	0.3	28.394	15.006	10.645	8.3665	6.8812	5.7739	4.8610	4.0299	3.1535
$Sk(X)$	0.5	11.871	9.1880	7.7473	6.7290	5.8760	5.0074	4.0245	3.0481	2.3967
	0.9	3.8494	3.8190	3.8663	3.9450	4.0469	4.1742	4.3382	4.5681	4.9638
	0.3	1013.3	309.78	161.18	100.77	68.465	48.323	34.381	23.848	15.006
$Ku(X)$	0.5	220.16	130.29	91.548	68.245	51.035	35.778	22.242	12.658	8.0343
	0.9	18.302	17.957	18.326	18.991	19.885	21.043	22.591	24.867	29.064
	0.3	1.1991	1.4834	1.8708	2.4049	3.1625	4.2889	6.0959	9.4100	17.537
$DsI(X)$	0.5	3.6639	5.8538	8.5590	12.117	16.997	23.776	32.517	41.596	48.478
	0.9	51.247	52.228	52.803	53.238	53.613	53.968	54.335	54.757	55.337

Table 2. Numerical descriptive statistics of the DOEHL-W distribution at b=0.6, α=1.2.

Measure	$\theta \downarrow p \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	0.4	0.0755	0.2133	0.3975	0.6343	0.9436	1.3643	1.9755	2.9707	5.0368
$E(X)$	0.6	0.6259	1.1328	1.6984	2.3763	3.2321	4.3741	6.0147	8.6675	14.153
	0.8	3.6802	5.7999	8.1058	10.839	14.270	18.833	25.373	35.912	54.419
	0.4	0.0746	0.2091	0.3930	0.6444	1.0003	1.5274	2.3617	3.8372	7.1363
$Var(X)$	0.6	0.8512	1.7829	3.0244	4.7376	7.1792	10.813	16.584	26.828	49.816
	0.8	14.620	29.316	48.870	75.848	114.31	171.57	262.59	423.25	759.51
	0.4	3.5725	2.0881	1.5211	1.1985	0.9634	0.7612	0.5649	0.3507	0.0729
$Sk(X)$	0.6	1.7469	1.3979	1.1803	1.0048	0.8450	0.6871	0.5194	0.3246	0.0604
	0.8	1.5175	1.2882	1.1164	0.9648	0.8194	0.6709	0.5094	0.3111	0.1675
	0.4	15.412	7.0755	5.0426	4.1224	3.5473	3.1312	2.8139	2.5811	2.4776
$Ku(X)$	0.6	6.5836	5.1496	4.3780	3.8395	3.4218	3.0817	2.8020	2.5862	2.4868
	0.8	5.8968	4.8974	4.2625	3.7843	3.3968	3.0722	2.8000	2.5590	2.2338
	0.4	0.9883	0.9802	0.9887	1.0159	1.0600	1.1196	1.1955	1.2917	1.4168
$DsI(X)$	0.6	1.3601	1.5740	1.7808	1.9937	2.2212	2.4719	2.7573	3.0952	3.5199
	0.8	3.9727	5.0545	6.0289	6.9975	8.0104	9.1105	10.349	11.786	13.957

Table 3. Numerical entropy of DOEHL-W distribution at b=0.2, α=0.3, ω=3.

$\theta \downarrow p \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.3	0.0042	0.0211	0.0543	0.1066	0.1809	0.2820	0.4193	0.6153	0.9452
0.5	0.0903	0.1824	0.2785	0.3812	0.4943	0.6237	0.7804	0.9886	1.3248
0.9	0.9277	1.0877	1.2177	1.3403	1.4658	1.6030	1.7646	1.9756	2.3133

Table 4. Numerical entropy of DOEHL-W distribution at b = 0.1, α=3, ω=2.

$\theta \downarrow p \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.4	0.3738	1.0428	1.9002	2.9258	4.1387	5.6061	7.4789	10.094	14.484
0.6	2.5342	4.1013	5.5456	7.0030	8.5608	10.316	12.423	15.208	19.696
0.8	7.4368	9.4791	11.174	12.787	14.446	16.268	18.418	21.229	25.730

Based on the findings in Tables 1 and 2, the analysis showcases the DOEHL-W model's efficacy in accurately representing symmetric and asymmetric data with various kurtosis shapes like leptokurtic and platykurtic. Additionally, it effectively handles both overdispersion and underdispersion situations. For a constant parameter p, increasing the parameter θ in Tables 3 and 4 leads to a notable rise in entropy values for the DOEHL-W model, suggesting increased uncertainty as θ increases and emphasizing the model's sensitivity to parameter changes, which remains consistent as p increases while θ remains constant.

5. Maximum Likelihood Estimation

In this segment, we delve into the parameter estimation of the DOEHL-G family through the maximum likelihood method, utilizing a complete sample. Consider a random sample $X_1, X_2, \ldots, X_n$ from the DOEHL-G class. Subsequently, the log-likelihood function (L) can be derived as follows

$$L(\alpha, p, \mathbf{\Phi}|\underline{\mathbf{x}}) = \sum_{i=1}^{n} \ln\left[\left(\frac{1-p^{Q(x_i+1;\mathbf{\Phi})}}{1+p^{Q(x_i+1;\mathbf{\Phi})}}\right)^{\alpha} - \left(\frac{1-p^{Q(x_i;\mathbf{\Phi})}}{1+p^{Q(x_i;\mathbf{\Phi})}}\right)^{\alpha}\right]. \tag{16}$$

Differentiating the log-likelihood with respect to α, p and $\mathbf{\Phi}$ respectively, and setting the result equal to zero, we have

$$\frac{\partial L(\alpha, p, \mathbf{\Phi}|\underline{\mathbf{x}})}{\partial \alpha} = \sum_{i=1}^{n} \frac{(\Lambda_1(x_i+1;\mathbf{\Phi}))^{\alpha} \ln(\Lambda_1(x_i+1;\mathbf{\Phi})) - (\Lambda_1(x_i;\mathbf{\Phi}))^{\alpha} \ln(\Lambda_1(x_i;\mathbf{\Phi}))}{(\Lambda_1(x_i+1;\mathbf{\Phi}))^{\alpha} - (\Lambda_1(x_i;\mathbf{\Phi}))^{\alpha}}, \tag{17}$$

$$\frac{\partial L(\alpha, p, \mathbf{\Phi}|\underline{\mathbf{x}})}{\partial p} = -2\alpha \sum_{i=1}^{n} \frac{(\Lambda_1(x_i+1;\mathbf{\Phi}))^{\alpha-1} Q(x_i+1;\mathbf{\Phi})\, \Lambda_2(x_i+1;\mathbf{\Phi}) - (\Lambda_1(x_i;\mathbf{\Phi}))^{\alpha-1} Q(x_i;\mathbf{\Phi})\, \Lambda_2(x_i;\mathbf{\Phi})}{(\Lambda_1(x_i+1;\mathbf{\Phi}))^{\alpha} - (\Lambda_1(x_i;\mathbf{\Phi}))^{\alpha}}, \tag{18}$$

$$\frac{\partial L(\alpha, p, \mathbf{\Phi}|\underline{\mathbf{x}})}{\partial \mathbf{\Phi}_j} = -2\alpha \ln p \sum_{i=1}^{n} \frac{(\Lambda_1(x_i+1;\mathbf{\Phi}))^{\alpha-1} \omega(x_i+1;\mathbf{\Phi})\, \Lambda_2(x_i+1;\mathbf{\Phi}) - (\Lambda_1(x_i;\mathbf{\Phi}))^{\alpha-1} \omega(x_i;\mathbf{\Phi})\, \Lambda_2(x_i;\mathbf{\Phi})}{(\Lambda_1(x_i+1;\mathbf{\Phi}))^{\alpha} - (\Lambda_1(x_i;\mathbf{\Phi}))^{\alpha}}, \tag{19}$$

where $\Lambda_1(x_i;\mathbf{\Phi}) = \frac{1-p^{Q(x_i;\mathbf{\Phi})}}{1+p^{Q(x_i;\mathbf{\Phi})}}$, $\Lambda_2(x_i;\mathbf{\Phi}) = p^{Q(x_i;\mathbf{\Phi})-1}\left(1+p^{Q(x_i;\mathbf{\Phi})}\right)^{-2}$ and $[\omega(x_i;\mathbf{\Phi})]_{\mathbf{\Phi}_j} = \frac{\partial Q(x_i;\mathbf{\Phi})}{\partial \mathbf{\Phi}_j}$; $j = 1, 2, ..., m$. To address a maximum likelihood estimation (MLE) issue involving a nonlinear likelihood function in R, one can utilize optimization techniques offered by packages such as **optim**, **nlminb**, or **nloptr**.

6. Testing the Estimation Technique Using Different Simulation Schemes

In this section, the performance of the MLE is evaluated through a simulation approach for the DOEHL-W model. The *distr* package offers a more structured, object-oriented way to define discrete distributions using the *DiscreteDistribution* class and generate samples with the *r*() function. While the *sample*() function from base R is adequate for most discrete distributions, the *distr* package provides enhanced flexibility and advanced features for more complex scenarios. The choice of method depends on the complexity of the distribution: simple sampling can be handled with *sample*(), whereas the *distr* package is better suited for more intricate discrete distributions.

A total of $N = 10000$ samples are generated, comprising five different sizes ($n = 20, 50, 100, 200, 500$), drawn from the DOEHL-W distribution across various values of the model parameters outlined in Tables 5 and 6. The simulation evaluates bias, mean squared error (MSE), and mean relative error (MRE) as criteria. Both numerical results (presented in Tables 5 and 6) and graphical representations (Figures 3-6) demonstrate that bias, MSE, and MRE decline with increasing sample size n. This highlights the consistency property of the MLE across all parameter combinations, affirming its efficacy in accurately estimating the parameters of the DOEHL-W distribution.

Table 5. Simulation results for the DOEHL-W parameters.

		$\alpha = 0.5, p = 0.9, \theta = 0.7, b = 0.3$			$\alpha = 0.5, p = 0.8, \theta = 0.9, b = 0.9$		
n	**Parameter**	**Bias**	**MSE**	**MRE**	**Bias**	**MSE**	**MRE**
α	20	0.63861984	0.41735783	0.49468312	0.30465171	0.24374581	0.28463975
	50	0.42894761	0.28478711	0.32717840	0.21946955	0.13658963	0.18409761
	100	0.29947971	0.12947092	0.22846913	0.13058981	0.09847913	0.10574961
	200	0.10289476	0.00648461	0.08496139	0.08465921	0.00080536	0.04659761
	500	0.00428648	0.00004718	0.00095621	0.00074975	0.00000981	0.00009478
p	20	0.32163480	0.22846898	0.30463881	0.28463966	0.24649461	0.25083745
	50	0.23624708	0.13338460	0.21194781	0.23745915	0.18464962	0.21364803
	100	0.13269640	0.08804671	0.10994648	0.14746819	0.10389476	0.13856919
	200	0.07496915	0.00024648	0.01398569	0.08485913	0.00946751	0.04443856
	500	0.00055481	0.00000891	0.00029477	0.00064869	0.00006482	0.00020471
θ	20	0.19476496	0.16484692	0.17494685	0.25484619	0.22474816	0.24438468
	50	0.15484691	0.13204639	0.14438461	0.19888462	0.14383691	0.16384692
	100	0.10648461	0.07084639	0.09564397	0.12749616	0.02484620	0.08409047
	200	0.02295476	0.00648615	0.00993748	0.00443873	0.00026392	0.00104875
	500	0.00084836	0.00006493	0.00038987	0.00005410	0.00000076	0.00000854
b	20	0.53845817	0.39457461	0.45538451	0.49377617	0.39565983	0.43327451
	50	0.38256371	0.24378393	0.30574977	0.34387468	0.25548610	0.31057826
	100	0.21104873	0.13874638	0.18849486	0.22947615	0.13498469	0.19469675
	200	0.10946849	0.00463941	0.03328452	0.11048734	0.00885648	0.07453814
	500	0.00648812	0.00014846	0.00259946	0.00776581	0.00006487	0.00053862

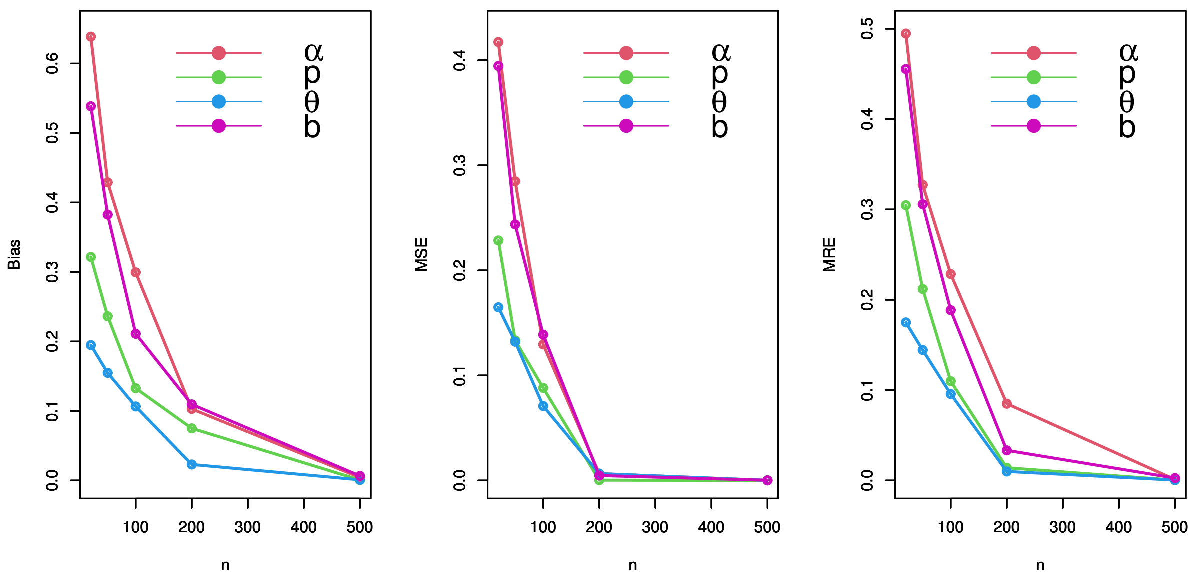


Figure 3. Simulation plots for the DOEHL-W parameters "$\alpha = 0.2, p = 0.3, \theta = 0.8, b = 0.2$".

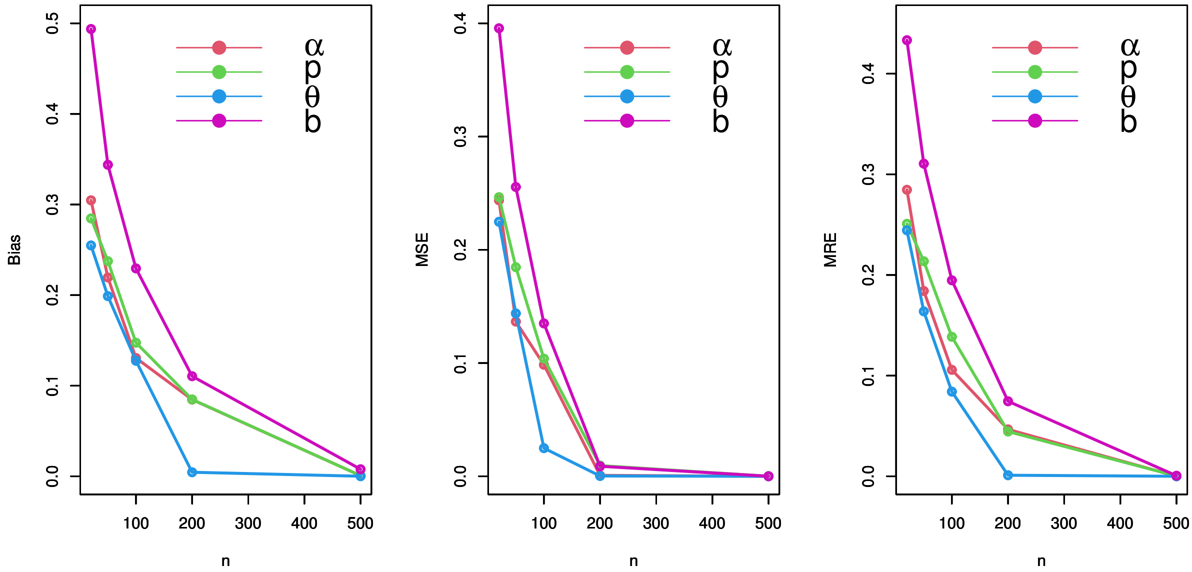


Figure 4. Simulation plots for the DOEHL-W parameters "$\alpha = 0.6, p = 0.2, \theta = 0.7, b = 0.5$".

Table 6. Simulation results for the DOEHL-W parameters.

		$\alpha = 0.1, p = 0.5, \theta = 0.2, b = 0.8$			$\alpha = 0.9, p = 0.2, \theta = 0.5, b = 0.5$		
n	**Parameter**	**Bias**	**MSE**	**MRE**	**Bias**	**MSE**	**MRE**
α	20	0.331100	0.280256	0.297488	0.369527	0.325994	0.354477
	50	0.263240	0.224543	0.245540	0.288626	0.209496	0.238802
	100	0.181024	0.122564	0.162637	0.184977	0.096039	0.122025
	200	0.099023	0.011286	0.016918	0.006469	0.00385	0.001523
	500	0.001442	0.000110	0.000663	0.000078	0.000001	0.000012
p	20	0.807687	0.592769	0.683213	0.643756	0.514544	0.564442
	50	0.577430	0.366797	0.459787	0.473680	0.333333	0.403836
	100	0.320656	0.209646	0.290980	0.298431	0.185128	0.253199
	200	0.164303	0.007019	0.049939	0.145784	0.011546	0.096935
	500	0.009860	0.000225	0.003949	0.010181	0.000085	0.000705
θ	20	0.383172	0.254588	0.301163	0.578838	0.461825	0.541626
	50	0.257369	0.173720	0.199186	0.422040	0.260613	0.351819
	100	0.179688	0.078977	0.139092	0.248170	0.187113	0.201320
	200	0.061737	0.003956	0.051724	0.160893	0.001533	0.088558
	500	0.002572	0.000029	0.000582	0.001431	0.000019	0.000181
b	20	0.257308	0.182775	0.243922	0.341568	0.296078	0.301048
	50	0.188998	0.106814	0.169794	0.285853	0.223745	0.256481
	100	0.106157	0.070431	0.088007	0.177942	0.134517	0.166350
	200	0.059975	0.000197	0.011183	0.102234	0.011406	0.053533
	500	0.000444	0.000004	0.000237	0.000781	0.000078	0.000247

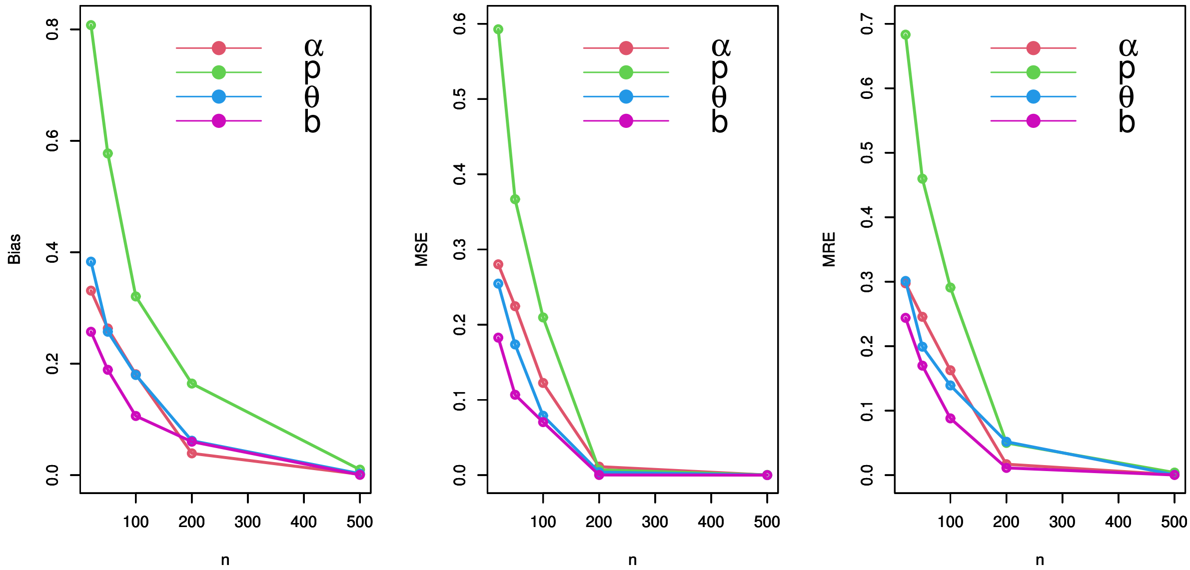


Figure 5. Simulation plots for the DOEHL-W parameters "$\alpha = 0.1, p = 0.5, \theta = 0.2, b = 0.8$".

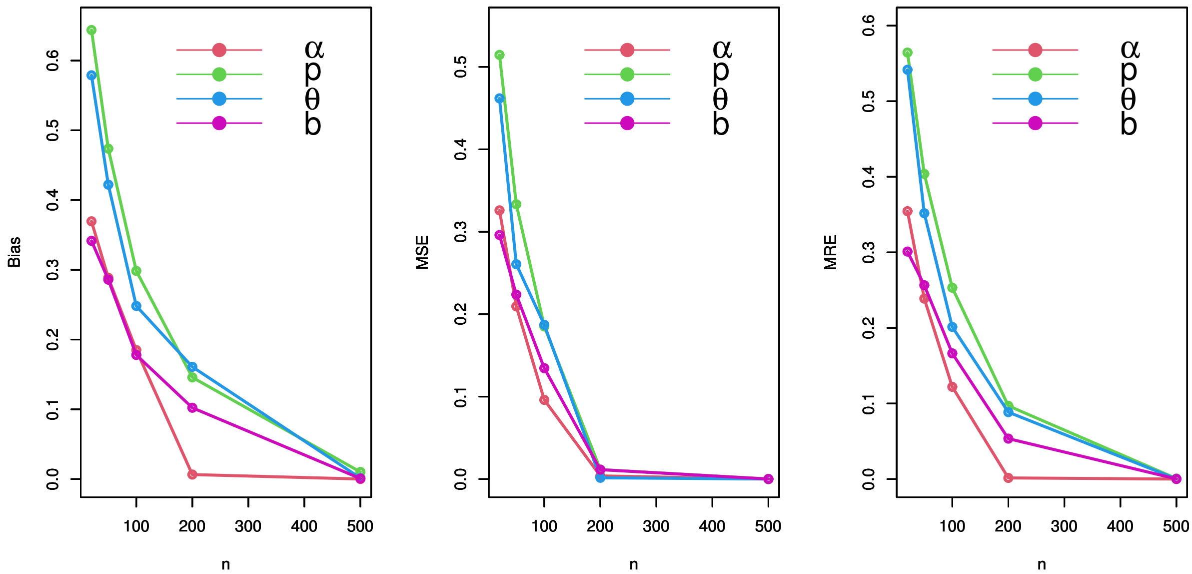


Figure 6. Simulation plots for the DOEHL-W parameters "$\alpha = 0.9, p = 0.2, \theta = 0.5, b = 0.5$".

7. Fitting and Analyzing Data: Non-Parametric and Parametric Infographics

This section demonstrates the practical uses of the DOEHL-W distribution with three actual datasets. Nonparametric plots were initially employed to discuss the datasets. Nonparametric plots, such as box plot, violin plot, normal quantile-quantile (QQ) plot, total time in test (TTT) plot, strip plot, and scatter|scores plot, are graphical tools that do not assume an underlying probability distribution, making them advantageous for exploratory data analysis and understanding data distribution and relationships without relying on specific distribution assumptions. Following data visualization, we assess the fitted models using various criteria: Negative log-likelihood ($-L$), Akaike information criterion (AIC), corrected Akaike information criterion (CAIC), Hannan-Quinn information criterion (HQIC), Chi-square (χ^2) with degrees of freedom (DF), and its corresponding p-value. When performing a χ^2 test, it is crucial to ensure that the expected frequencies in each category are sufficiently large to maintain the validity of the test. To achieve this, categories with expected frequencies less than 3 are often merged. By combining these categories, we increase the expected frequencies, which helps satisfy the test's assumptions and ensures reliable results. This adjustment also impacts the DF for the test, as merging categories reduces the number of independent categories, thereby lowering the DF. Furthermore, when comparing p-values, merging categories with expected frequencies less than 5 can lead to different p-values than when only categories with expected frequencies less than 3 are combined. This change may affect the strength and interpretation of the test's results. The DOEHL-W distribution will be compared against other competitive models outlined in Table 7.

Table 7. The competitive models.

Distribution	Abbreviation	Author(s)
Discrete Rayleigh	DR	Roy (2004)
Discrete inverse Rayleigh	DIR	Hussain and Ahmad (2014)
Discrete inverse Weibull	DIW	Jazi et al. (2010)
Poisson	Pois	Poisson (1837)
Discrete Burr	DB	Krishna and Pundir (2009)
Discrete Pareto	DPa	Krishna and Pundir (2009)
Discrete Burr-Hatke	DBH	El-Morshedy et al.(2020)
Discrete log-logistic	DLogL	Para and Jan (2016)
Negative Binomial	NeBi	-
One parameter discrete Lindley	DLi-I	Gómez-Déniz and Calderín-Ojeda (2011)
One parameter discrete flexible model	DFx-I	Eliwa and El-Morshedy (2021)
Discrete Lomax distribution	DLo	Para and Jan (2016)
Two parameters discrete Lindley	DLi-II	Hussain et al. (2016)
Three parameters discrete Lindley	DLi-III	Eliwa et al. (2020)
Geometric	Geo	-

7.1. *Dataset I: Cysts of kidneys (COK)*

The dataset consists of counts of COK with steroid usage (refer to Chan et al., 2010). In this section, we calculate the maximum likelihood estimators of the DOEHL-W parameters using the COK data. Before conducting the analysis on the COK data, it is essential to visualize the data to understand its behavior. Figure 7 illustrates various nonparametric infographics. These plots reveal the presence of extreme values and outliers, indicating a positively skewed distribution with a decreasing HRF shape.

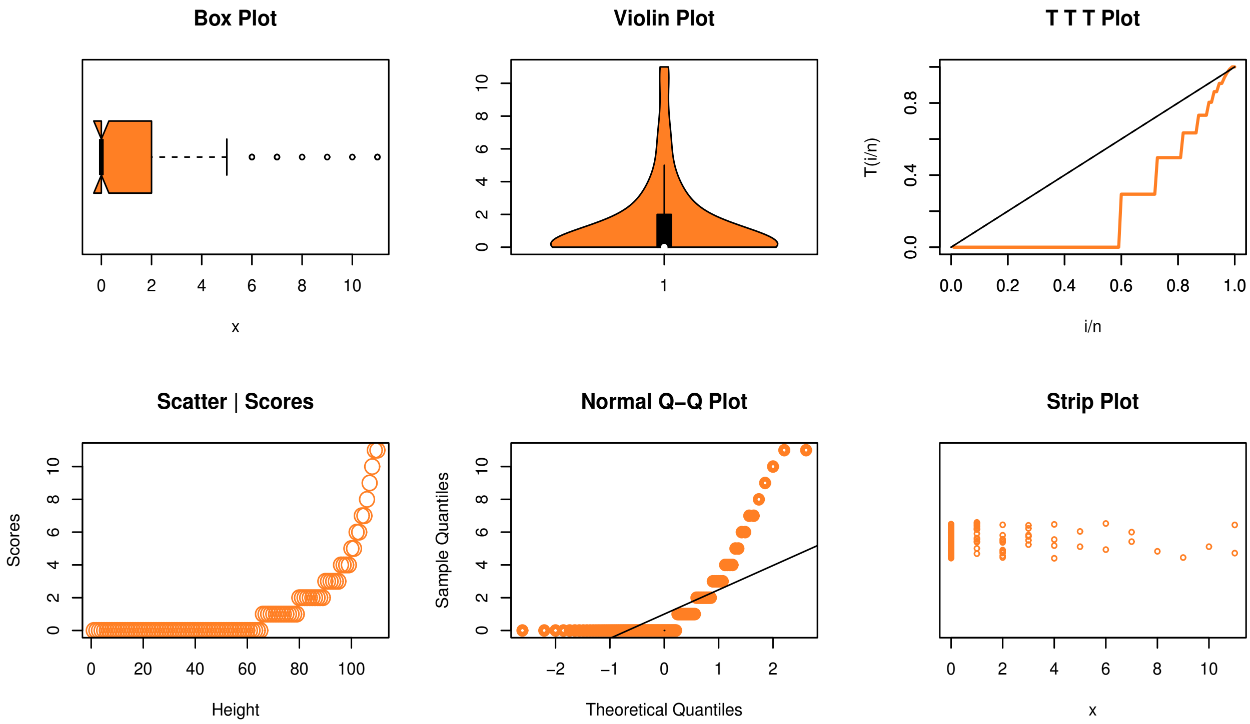


Figure 7. Non-parametric infographics of the COK data.

Tables 8, 9, and 10 display the maximum likelihood estimates (MLEs) for the parameter(s) along with their corresponding standard errors (Std-er), as well as the results of the goodness-of-fit test (GOFT).

Table 8. The MLEs with their corresponding Std-er and GOFT for the COK data.

Model	MLEs (Std − er)	−L	AIC	CAIC	HQIC
DOEHLW	$\widehat{\alpha} = 0.386\ (0.471)$, $\widehat{b} = 0.689\ (0.928)$, $\widehat{\theta} = 0.837\ (0.689)$, $\widehat{p} = 0.067\ (0.648)$	166.864	341.728	342.109	346.109
DR	$\widehat{\theta} = 0.901\ (0.009)$	277.778	557.556	557.593	558.651
DIR	$\widehat{\theta} = 0.554\ (0.049)$	186.547	375.094	375.131	376.189
DIW	$\widehat{\beta} = 1.049\ (0.146)$, $\widehat{\theta} = 0.581\ (0.048)$	172.935	349.869	349.982	352.060
Pois	$\widehat{\alpha} = 1.390\ (0.112)$	246.210	494.420	494.457	495.515
DB	$\widehat{\alpha} = 0.278\ (0.045)$, $\widehat{\beta} = 1.053\ (0.167)$	171.139	346.278	346.391	348.469
DLogL	$\widehat{\alpha} = 0.780\ (0.136)$, $\widehat{\beta} = 1.208\ (0.159)$	171.717	347.430	347.550	349.620
DLi-I	$\widehat{\alpha} = 0.436\ (0.026)$	189.110	380.220	380.257	381.316
DFx-I	$\widehat{\alpha} = 0.623\ (0.031)$	182.288	366.575	366.612	367.671
DLo	$\widehat{\alpha} = 0.152\ (0.098)$, $\widehat{\beta} = 1.830\ (0.952)$	170.481	344.961	345.073	347.152
DLi-II	$\widehat{\alpha} = 0.581\ (0.045)$, $\widehat{\beta} = 0.001\ (0.058)$	178.767	361.534	361.646	363.724
DLi-III	$\widehat{\alpha} = 0.582\ (0.005)$, $\widehat{\theta} = 0.001\ (20.698)$, $\widehat{\beta} = 358.728\ (11863.37)$	178.767	363.533	363.759	366.819
Geo	$\widehat{\theta} = 0.582\ (0.030)$	178.767	359.533	359.570	360.629

Table 9. The GOFT of the COK data "part I".

		ExFr						
X	ObFr	DOEHLW	DR	DIR	DIW	Pois	DB	DLogL
0	65	65.159	10.890	60.888	63.910	27.398	64.743	63.192
1	14	13.659	26.618	33.990	20.699	38.084	19.177	20.101
2	10	8.776	29.448	8.123	8.053	26.468	8.484	8.644
3	6	6.212	22.296	6.999	6.833	12.264	7.495	7.520
4	4	7.850	12.629			5.786		
5	2		8.119		10.505		10.101	10.543
6	2	8.344						
7	2							
8	1							
9	1							
10	1							
11	2							
Total	110	110	110	110	110	110	110	110
χ^2		0.674	306.52	40.456	6.445	89.277	2.587	4.033
DF		1	4	2	2	3	2	2
P-value		0.412	< 0.001	< 0.001	0.092	< 0.001	0.274	0.258

Table 10. The GOFT of the COK data "part II".

		ExFr						
X	ObFr	DOEHLW	DLi-I	DFx-I	DLo	DLi-II	DLi-III	Geo
0	65	65.159	40.286	45.256	61.615	46.026	46.008	45.980
1	14	13.659	29.834	29.094	21.023	26.768	26.765	26.760
2	10	8.776	18.357	16.508	9.687	15.568	15.570	15.575
3	6	6.212	10.336	8.893	5.275	9.054	9.058	9.064
4	4	7.850	5.523	10.249	5.285	5.266	5.269	5.275
5	2		50664			7.318	7.330	7.346
6	2	8.344			7.115			
7	2							
8	1							
9	1							
10	1							
11	2							
Total	110	110	110	110	110	110	110	110
χ^2		0.674	34.635	31.702	3.238	19.091	19.096	19.109
DF		1	4	3	3	3	2	4
P-value		0.412	< 0.001	< 0.001	0.356	< 0.001	< 0.001	< 0.001

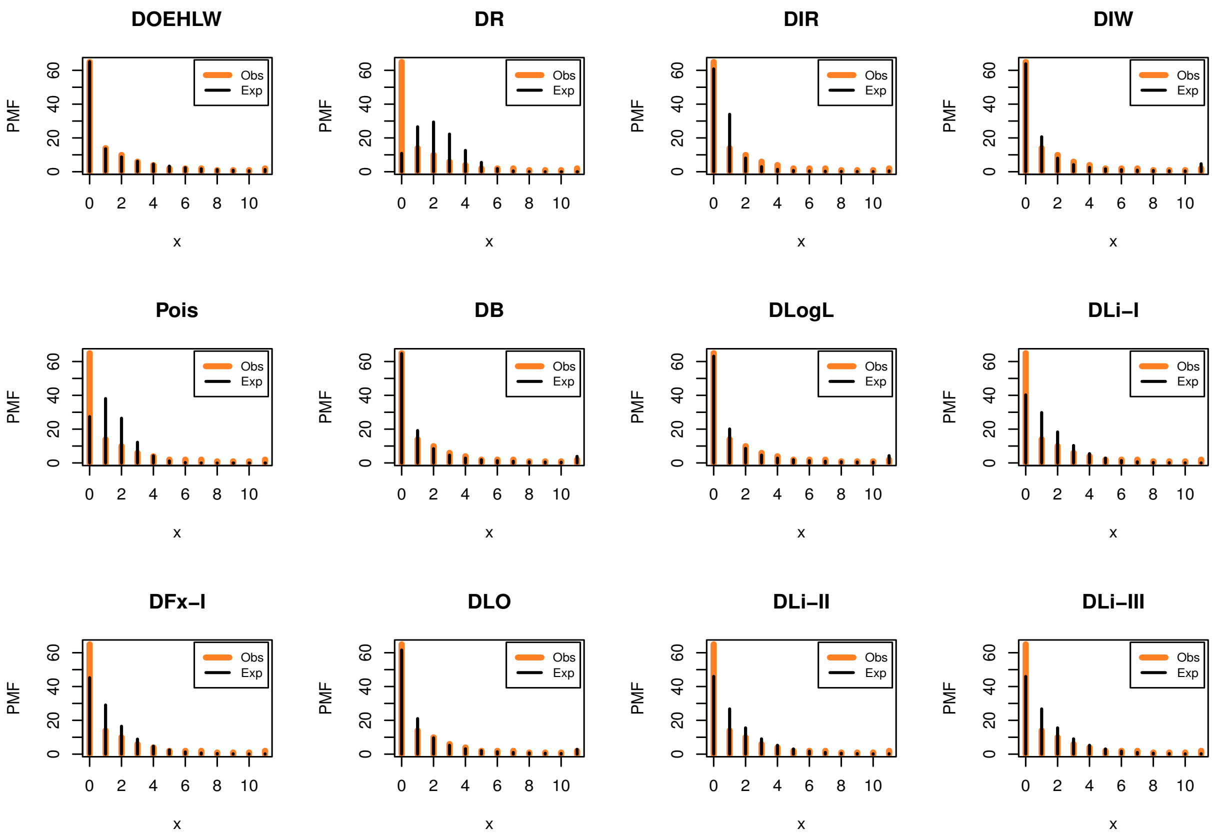
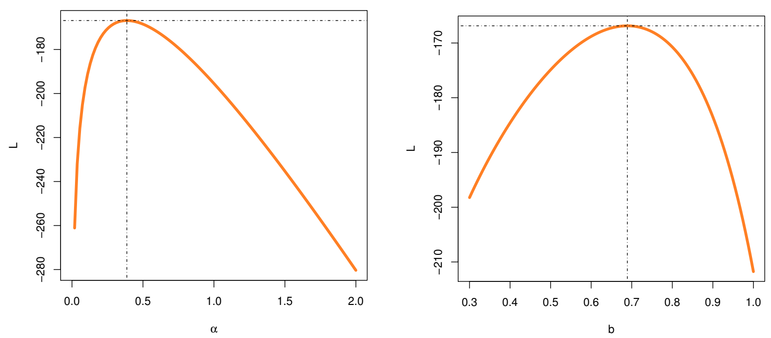


Figure 8. The observed and expected PMFs for dataset I.

The DIW, DB, DLogL, and DLo models are applicable for analyzing the COK dataset, with the DOEHLW model identified as the most suitable choice for this data. Figure 8 displays the fitted PMF for all test models, validating the results mentioned in Tables 8-10. Figure 9 depicts the L profile of the parameters for the DOEHLW model using the COK data, revealing that the estimators exhibit a unique "unimodal-shaped" distribution.



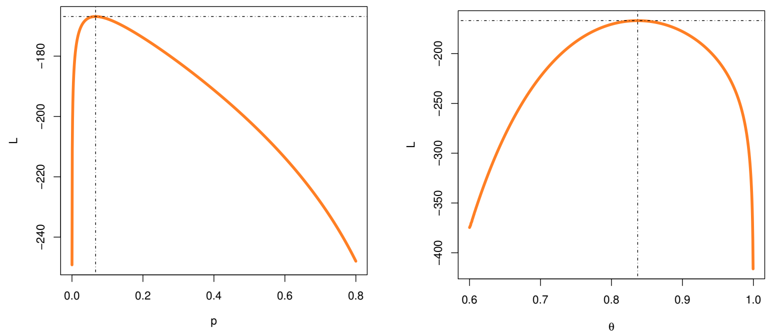


Figure 9. The L profile of parameters of DOEHLW for dataset I.

Figure 10 includes contour diagrams to further demonstrate the unique properties of the model estimators.

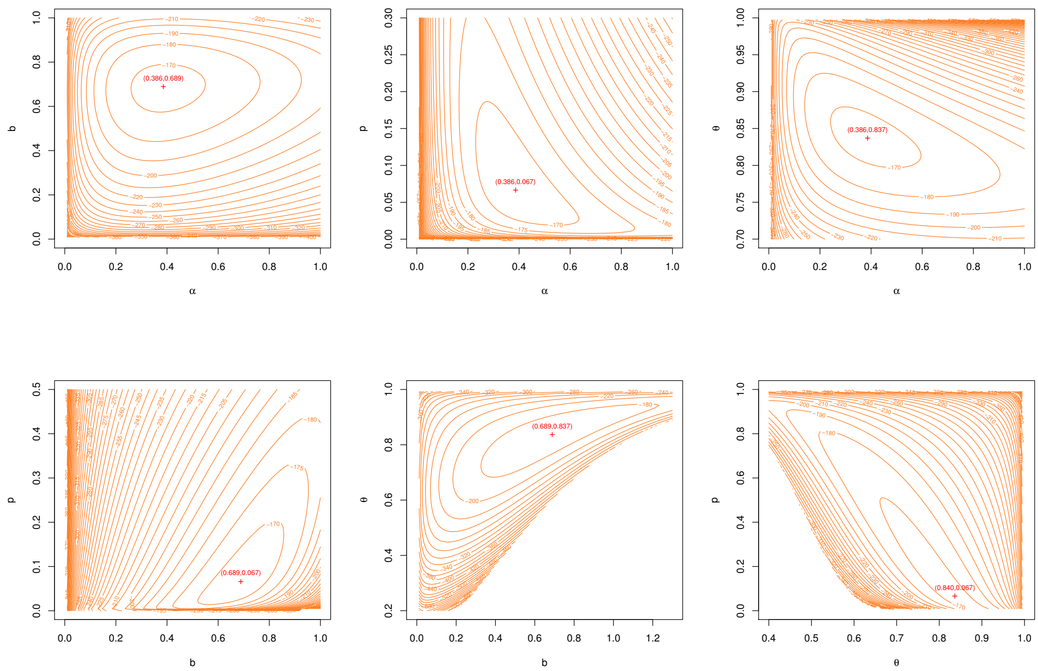


Figure 10. Contour diagrams of the DOEHLW estimators based on dataset I.

Table 11 displays the comparison between experimental and observational descriptive statistics, revealing that the two measures are nearly identical.

Table 11. Descriptive statistics of the COK data.

Model	Mean	Variance	Skewness	Kurtosis	Index of dispersion
Observed	1.3909	6.1118	2.2926	8.1735	4.3941
Empirical	1.3943	6.0547	2.4786	10.4734	4.3426

The data exhibits a positively skewed distribution with a leptokurtic shape and shows signs of overdispersion.

7.2. Dataset II: COVID-19 in South Korea

This dataset presents the daily new deaths of COVID-19SK spanning from 15 February 2020 to 12 December 2020 (source: *https://www.worldometers.info/coronavirus/country/south-korea/*). In this section, we assess the maximum likelihood estimators of the DOEHLW parameters based on the COVID-19SK data. Figure 11 displays nonparametric plots to analyze the data's behavior, revealing the presence of extreme observations with a positively skewed shape with bathtub HRF shape.

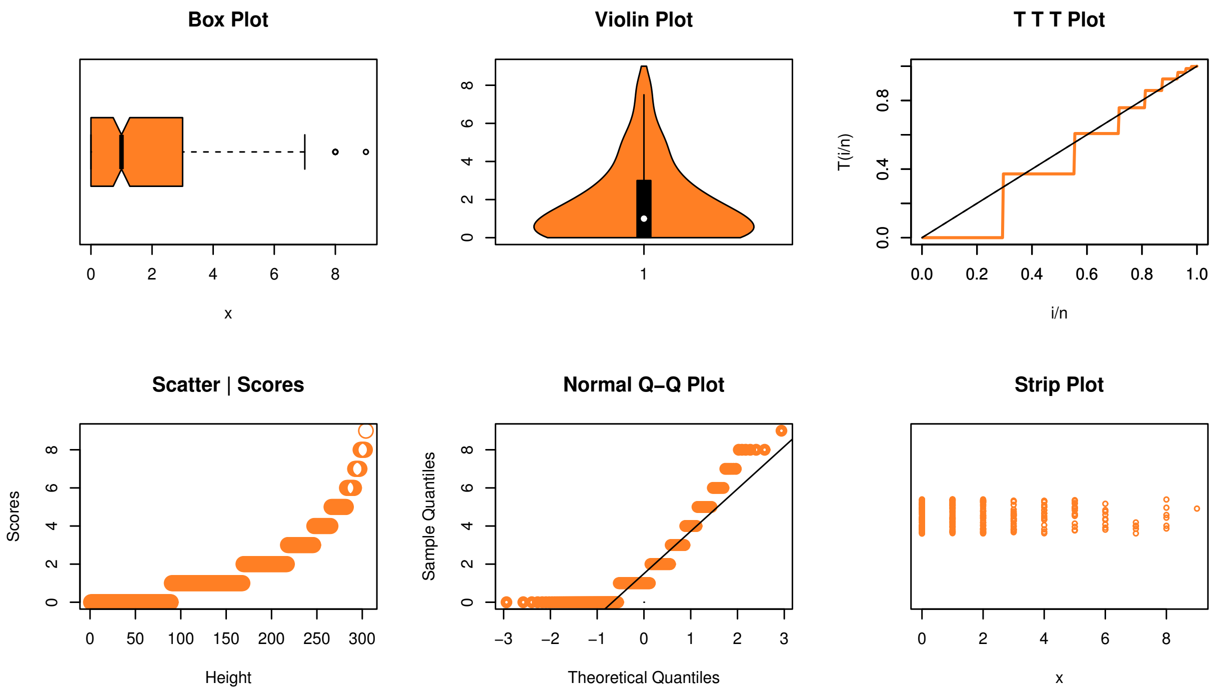


Figure 11. Non-parametric infographics of the COKID-19SK data.

The ML estimators for the parameter(s) as well as GOFT are listed in Tables 12 and 13. The DOEHLW model has been determined to be the most appropriate option for this dataset. Figure 12 showcases the fitted PMF for all test models, confirming the findings presented in Tables 12 and 13.

Table 12. The MLEs (Std-er) and GOFT for the COVID-19SK data.

Model	MLEs (Std − er)	−L	AIC	CAIC	HQIC
DOEHLW	$\widehat{\alpha} = 1.671\ (1.044)\,, \widehat{b} = 0.529\ (0.389)\,,$ $\widehat{\theta} = 0.662\ (0.542)\,, \widehat{p} = 0.126\ (0.521)$	564.707	1137.414	1137.548	1143.362
DR	$\widehat{\theta} = 0.903\ (0.005)$	638.905	1279.810	1279.823	1281.297
DIR	$\widehat{\theta} = 0.229\ (0.023)$	606.870	1215.740	1215.754	1217.227
DIW	$\widehat{\alpha} = 0.271\ (0.025)\,, \widehat{\beta} = 1.411\ (0.083)$	586.855	1177.711	1177.751	1180.684
Pois	$\widehat{\alpha} = 1.901\ (0.079)$	621.098	1244.195	1244.208	1245.682
DB	$\widehat{\alpha} = 0.591\ (0.031)\,, \widehat{\beta} = 2.466\ (0.248)$	587.652	1179.304	1179.344	1182.278
DPa	$\widehat{\alpha} = 0.377\ (0.021)$	633.531	1269.061	1269.075	1270.548
DBH	$\widehat{\alpha} = 0.904\ (0.020)$	620.466	1242.932	1242.945	1244.419
DLogL	$\widehat{\alpha} = 1.716\ (0.095)\,, \widehat{\beta} = 1.878\ (0.107)$	577.011	1158.023	1158.063	1160.997

Table 13. The GOFT of the COVID-19SK data.

		ExFr								
X	ObFr	DOEHL	DR	DIR	DIW	Pois	DB	DPa	DBH	DLogL
0	89	90.459	29.500	69.890	82.351	45.408	92.887	149.396	166.600	80.931
1	79	71.794	72.408	140.613	103.702	86.336	97.788	50.500	54.598	92.775
2	49	51.831	80.783	47.685	44.390	82.074	42.676	25.478	26.667	51.431
3	29	34.757	61.938	19.125	22.317	52.014	21.174	15.379	15.540	27.336
4	19	22.181	35.680	9.333	12.926	24.725	12.172	10.312	10.017	15.559
5	17	13.672	15.984	5.198	8.248	13.443	7.754	7.387	6.885	9.518
6	9	8.211	7.707	5.260	5.636		5.308	5.533	8.635	6.193
7	6	11.095			7.083		6.710	7.874		7.308
8	6			6.896					15.058	
9	1				17.347		17.531	32.181		12.949
Total	304	304	304	304	304	304	304	304	304	304
χ^2		3.524	184.999	92.204	41.868	115.896	44.784	128.631	109.333	25.019
DF		3	5	6	6	4	6	7	6	6
P-value		0.318	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

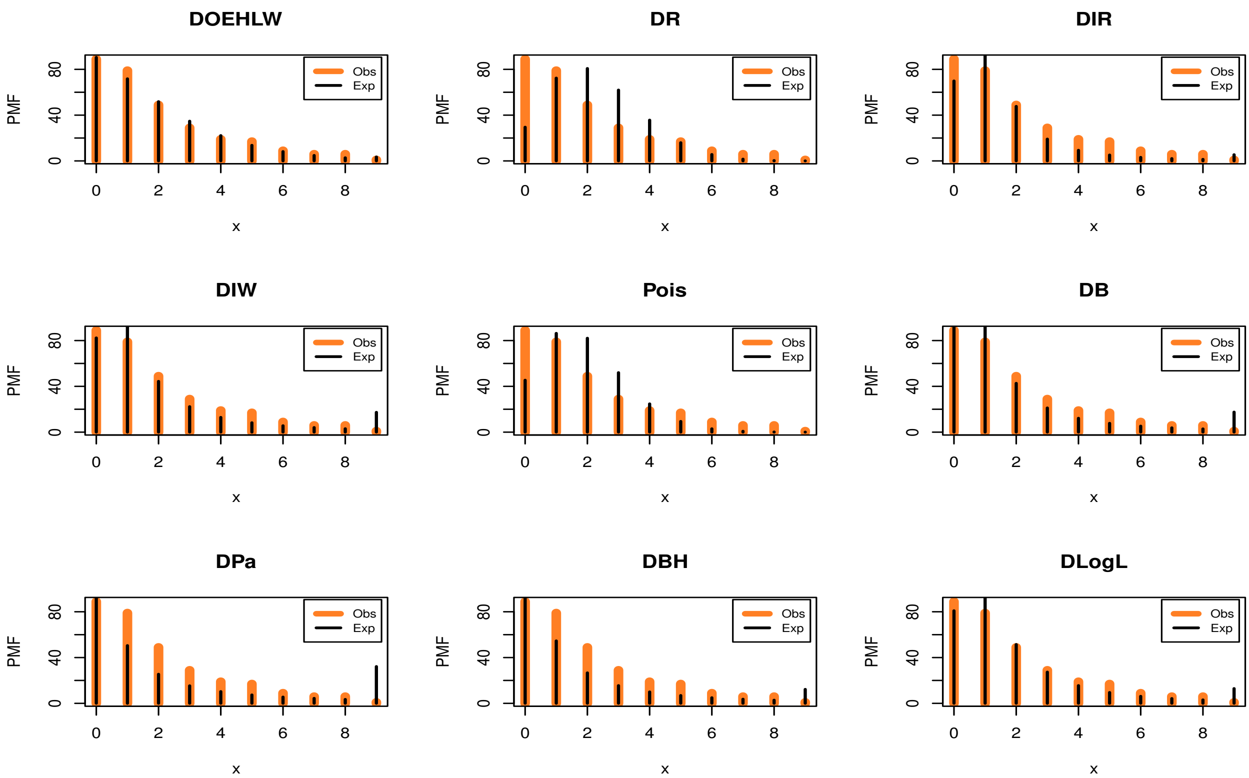


Figure 12. The fitted PMFs for the COVID-19SK data.

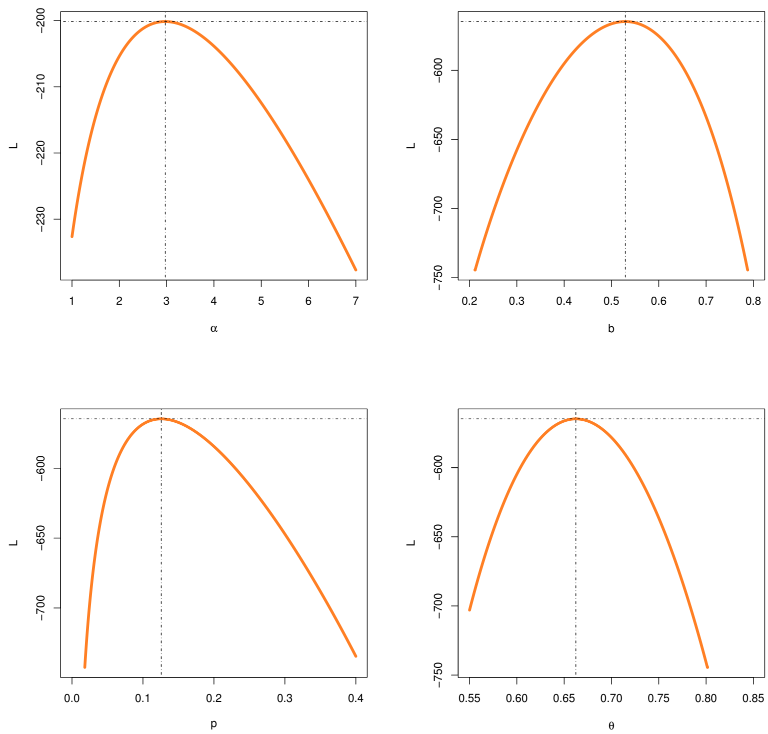


Figure 13. The L profile of parameters of DOEHLW for dataset II.

Figures 13 and 14 present the L profile and the contour diagrams of the DOEHLW parameters based on the COVID-19SK data. It's evident that the estimators exhibit uniqueness.

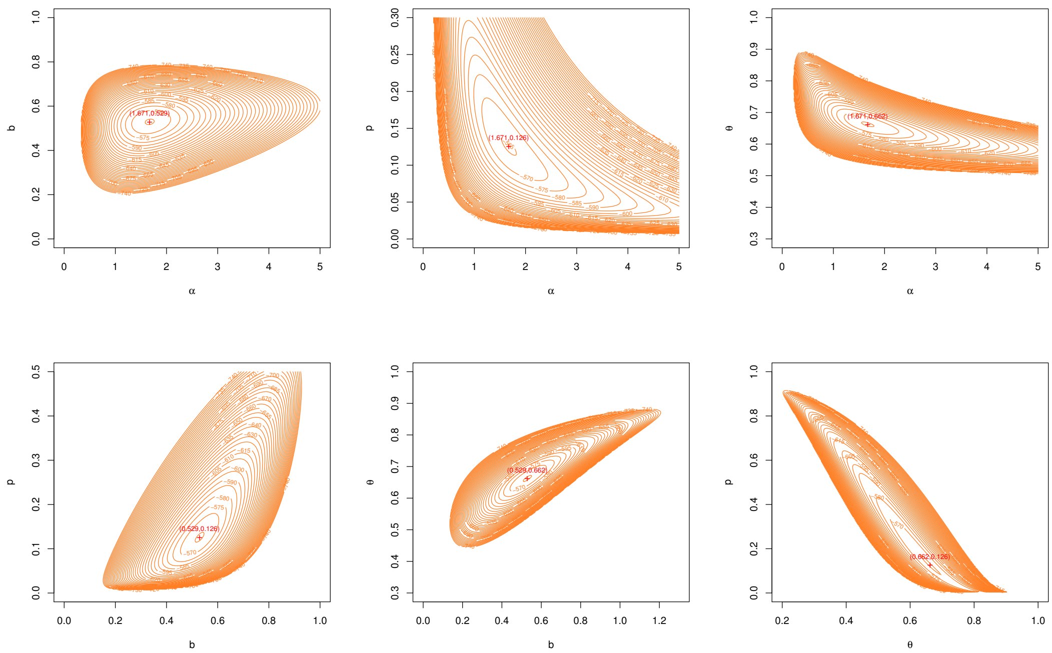


Figure 14. Contour diagrams of the DOEHLW estimators based on dataset II.

Table 14 outlines the variance between experimental and observational descriptive statistics, with both measures being approximately equivalent.

Table 14. Descriptive statistics of the COVID-19SK data.

Model	Mean	Variance	Skewness	Kurtosis	Index of dispersion
Observed	1.8567	3.9978	1.3657	4.2367	2.1531
Empirical	1.8997	4.1734	1.5436	6.1714	2.1969

The COVID-19SK data displays positive skewness with a leptokurtic shape and indicates overdispersion.

7.3. Dataset III: Biological experiment (BE)

This dataset records the count of European corn-borer larvae (Pyrausta) in a field (refer to Bodhisuwan and Sangpoom, 2016). The experiment was conducted randomly on eight hills with 15 replications each, where the researcher recorded the number of borers per hill of corn. In this section, we assess the maximum likelihood estimators of the DOEHLW parameters using the BE data. Figure 17 presents nonparametric plots to analyze the data's behavior, highlighting the presence of extreme observations with a positively skewed shape with a decreasing HRF shape.

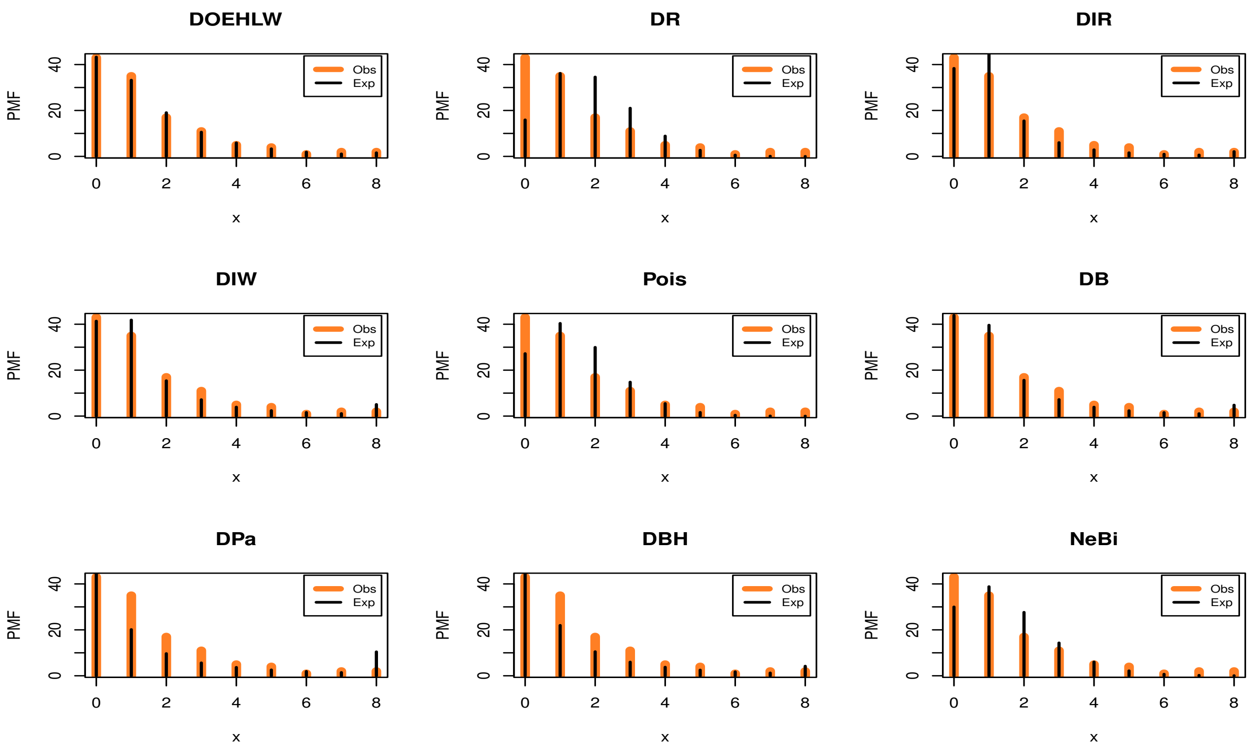


Figure 17. Non-parametric infographics of the BE dataset.

The ML estimators for the parameter(s) as well as GOFT are listed in Tables 15 and 16.

Table 15. The MLEs (Std-er) and GOFT for the BE dataset.

Model	MLEs (Std − er)	−L	AIC	CAIC	HQIC
DOEHLW	$\widehat{\alpha} = 2.971\ (3.386)$, $\widehat{b} = 0.424\ (0.297)$, $\widehat{\theta} = 0.641\ (0.454)$, $\widehat{p} = 0.042\ (0.219)$	200.135	408.271	408.618	412.799
DR	$\widehat{\theta} = 0.867\ (0.012)$	235.227	472.453	472.487	473.585
DIR	$\widehat{\theta} = 0.319\ (0.042)$	208.440	418.881	418.915	420.013
DIW	$\widehat{\alpha} = 0.345\ (0.043)$, $\widehat{\beta} = 1.541\ (0.156)$	204.810	413.621	413.723	415.885
Pois	$\widehat{\alpha} = 1.483\ (0.025)$	219.188	440.376	440.410	441.508
DB	$\widehat{\alpha} = 0.519\ (0.051)$, $\widehat{\beta} = 2.358\ (0.366)$	204.293	412.587	412.689	414.851
DPa	$\widehat{\alpha} = 0.329\ (0.034)$	220.618	443.236	443.270	444.368
DBH	$\widehat{\alpha} = 0.865\ (0.039)$	214.049	430.098	430.132	431.230
NeBi	$\widehat{\alpha} = 0.870\ (0.036)$, $\widehat{\beta} = 9.956\ (0.096)$	211.526	427.052	427.155	429.316

Table 16. The GOFT of the BE dataset.

		ExFr								
X	ObFr	DOEHLW	DR	DIR	DIW	Pois	DB	DPa	DBH	NeBi
0	43	43.275	15.919	38.347	41.374	27.226	43.831	64.452	68.073	29.994
1	35	33.215	36.170	51.876	41.850	40.385	39.607	20.149	21.967	38.821
2	17	19.065	34.577	15.490	15.425	29.952	15.623	9.686	10.513	27.646
3	11	10.590	21.025	6.028	7.169	14.810	7.206	5.647	5.983	14.323
4	5	5.906	12.309	8.259	6.368	7.627	6.825	6.260	6.261	9.216
5	4	7.949								
6	1				7.814		7.448	13.806	7.203	
7	2									
8	2									
Total	120	120	120	120	120	120	120	120	120	120
χ^2		0.615	70.688	14.274	5.511	38.478	4.664	32.462	26.824	20.367
DF		1	3	3	3	3	3	4	4	2
P-value		0.433	< 0.001	< 0.001	0.138	< 0.001	0.198	< 0.001	< 0.001	< 0.001

The BE dataset can be analyzed using the DIW and DB models, but the most appropriate choice is the DOEHLW model. Figure 18 depicts the fitted PMF for all test models, confirming the results presented in Tables 15 and 16.

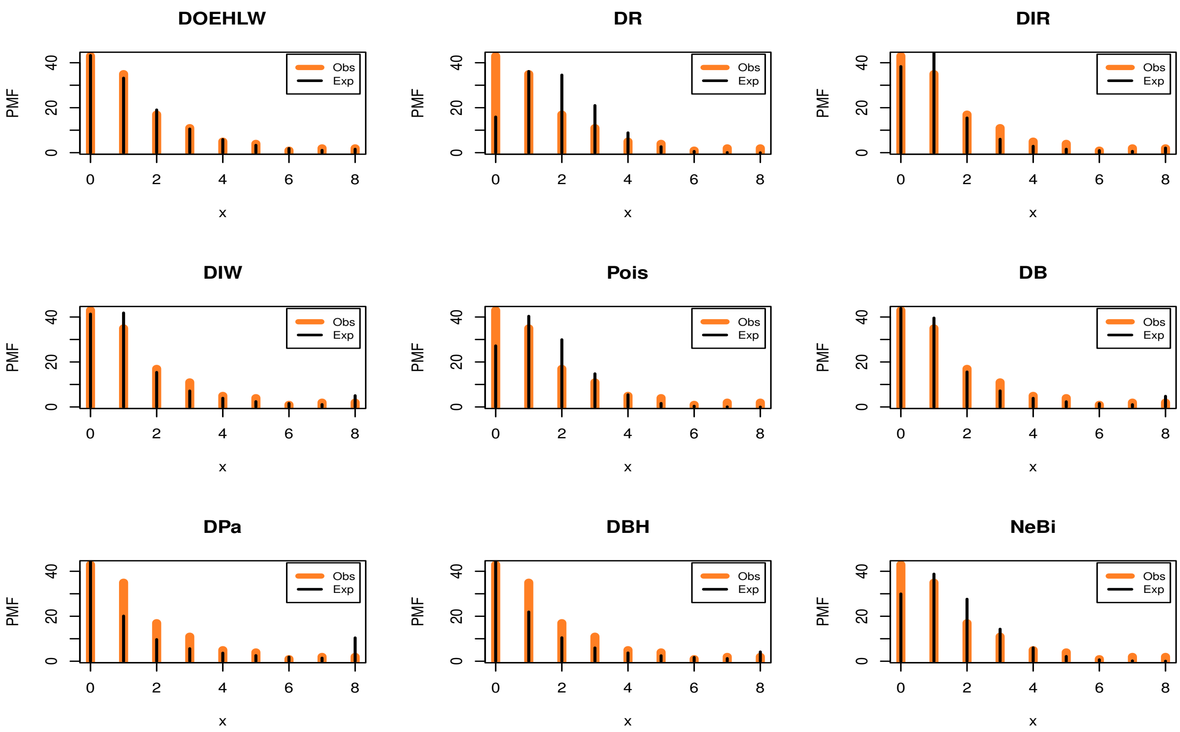
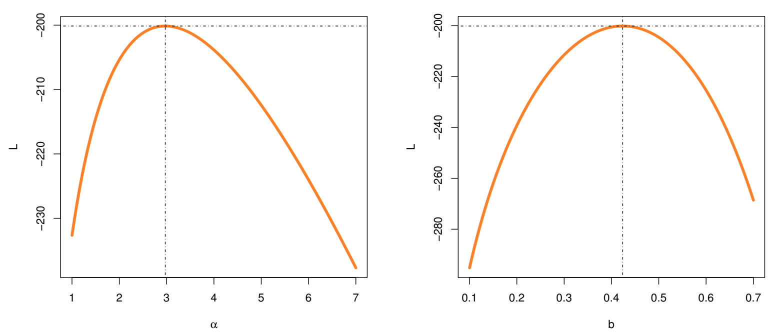


Figure 18. The fitted PMFs for the BE data.

Figures 19 and 20 showcase the L profile and contour diagrams of the DOEHLW parameters derived from the BE dataset, demonstrating the uniqueness of the estimators.



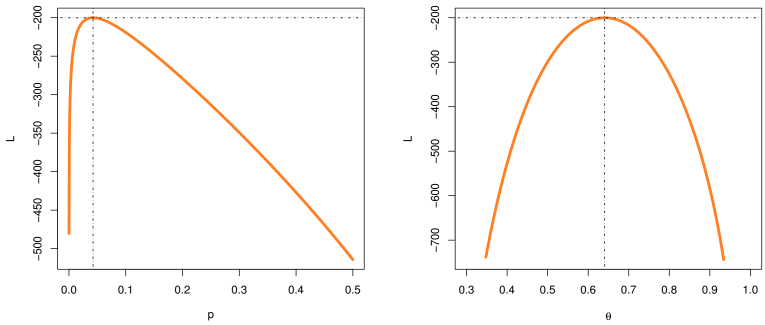


Figure 19. The L profile of parameters of DOEHLW for dataset III.

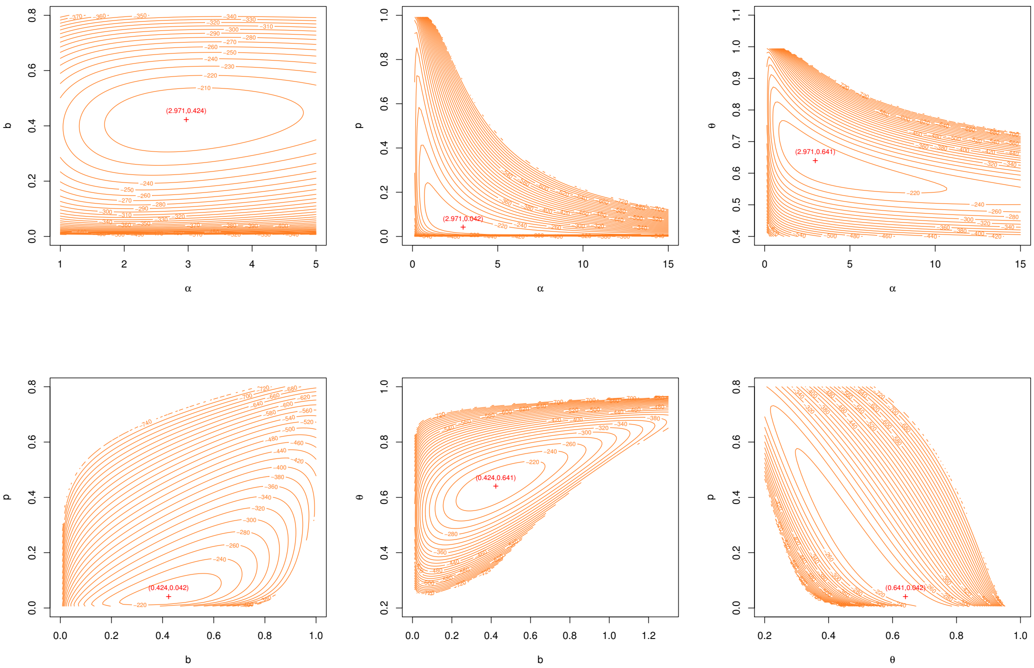


Figure 20. Contour diagrams of the model estimators based on dataset III.

Table 17 outlines the variance between experimental and observational descriptive statistics, with both measures being nearly equal.

Table 17. Descriptive statistics of the BE dataset.

Model	Mean	Variance	Skewness	Kurtosis	Index of dispersion
Observed	1.4833	3.1930	1.6948	6.0311	2.1526
Empirical	1.4813	3.2649	2.0920	9.8708	2.2040

The BE dataset shows positive skewness with a leptokurtic shape, suggesting overdispersion.

8. Summary and Prospects

In this article, a discrete analogue of the odd exponentiated half-logistic generator of distributions was derived and discussed in detail. After introducing the new class, a special discrete model was investigated. Some distributional properties, including the probability mass function, hazard rate function, moments, index of dispersion, entropies, order statistics, and L-moment statistics, were derived. It was found that the new probabilistic discrete class could be applied to discuss symmetric and asymmetric data under various kinds of kurtoses, including platykurtic and leptokurtic shapes. It could be used to analyze over-(under-) dispersed count data with zero-inflated, extreme, or outlier observations. Moreover, it could be used to create many discrete models, including different hazard rate forms. The class parameters were estimated via the maximum likelihood approach. The evaluation of the estimation technique using MCMC simulation revealed its effectiveness in accurately estimating model parameters across a range of sample sizes. Finally, to illustrate the flexibility of the new probability generator, three real datasets were analyzed and discussed in detail. The results proved that the new class provided high efficiency in data modeling among all competing models. As future work, the authors planned to propose the bivariate extension of this family with its applications.

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

- Acknowledgement. This project was supported by the deanship of scientific research at Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz Universityproject number (PSAU/2025/R/1446).

References

[1] Abd EL-Hady, A. E., Hegazy, M. A., & EL-Helbawy, A. A. (2023). Discrete exponentiated generalized family of distributions. Computational Journal of Mathematical and Statistical Sciences, 2(2), 303-327.

[2] Adubisi, O. D., Abdulkadir, A., Farouk, U. A., & Chiroma, H. (2022). The exponentiated half logistic skew-t distribution with GARCH-type volatility models. Scientific African, 16, e01253.

[3] Alosey, A. R. E., & Gemeay, A. M. (2025). A novel version of geometric distribution: Method and application. Computational Journal of Mathematical and Statistical Sciences, 4(1), 1-16.

[4] Afify, A. Z., Altun, E., Alizadeh, M., Ozel, G., & Hamedani, G. G. (2017). The odd exponentiated half-logistic-G family: properties, characterizations and applications. Chilean Journal of Statistics, 8(2), 65-91.

[5] Almetwally, E. M., Abdo, D. A., Hafez, E. H., Jawa, T. M., Sayed-Ahmed, N., & Almongy, H. M. (2022). The new discrete distribution with application to COVID-19 Data. Results in Physics, 32, 104987.

[6] Al-Mofleh, H., Elgarhy, M., Afify, A., & Zannon, M. (2020). Type II exponentiated half logistic generated family of distributions with applications. Electronic Journal of Applied Statistical Analysis, 13(2), 536-561.

[7] Alotaibi, R. M., Rezk, H. R., Ghosh, I., & Dey, S. (2021). Bivariate exponentiated half logistic distribution: Properties and application. Communications in Statistics-Theory and Methods, 50(24), 6099-6121.

[8] Aryuyuen, S., Bodhisuwan, W., & Volodin, A. (2020). Discrete generalized odd Lindley–Weibull distribution with applications. Lobachevskii Journal of Mathematics, 41, 945-955.

[9] Bodhisuwan, W., & Sangpoom, S. (2016). The discrete weighted Lindley distribution. In 2016 12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA) (pp. 99-103). IEEE.

[10] Chakraborty, S. (2015). Generating discrete analogues of continuous probability distributions-A survey of methods and constructions. Journal of Statistical Distributions and Applications, 2, 1-30.

[11] Chan, S. K., Riley, P. R., Price, K. L., McElduff, F., Winyard, P. J., Welham, S. J., Woolf, A. S., & Long, D. A. (2010). Corticosteroid-induced kidney dysmorphogenesis is associated with deregulated expression of known cystogenic molecules, as well as Indian hedgehog. American Journal of Physiology-Renal Physiology, 298(2), F346-F356.

[12] Cordeiro, G. M., Alizadeh, M., & Ortega, E. M. (2014). The exponentiated half-logistic family of distributions: Properties and applications. Journal of Probability and Statistics, 2014.

[13] Davis, P. J., & Polonsky, I. (1965). Handbook of Mathematical Functions, ed. M. Abramowitz and IA Stegun.

[14] El-Awady, M. M., El-Monsef, M. M. E. A., & Elbaz, I. M. (2024). Exponentiated half-logistic Weibull distribution with reliability inference. Quality and Reliability Engineering International.

[15] El-Awady, M. M., El-Monsef, M. M. E. A., & Elbaz, I. M. (2024). Exponentiated half-logistic Weibull distribution with reliability inference. Quality and Reliability Engineering International.

[16] Eliwa, M. S., & El-Morshedy, M. (2021). A one-parameter discrete distribution for over-dispersed data: Statistical and reliability properties with applications. Journal of Applied Statistics. Advance online publication. doi:10.1080/02664763.2021.1905787

[17] Eliwa, M. S., Alhussain, Z. A., & El-Morshedy, M. (2020). Discrete Gompertz-G family of distributions for over-and under-dispersed data with properties, estimation, and applications. Mathematics, 8(3), 358.

[18] Eliwa, M. S., Altun, E., El-Dawoody, M., & El-Morshedy, M. (2020). A new three-parameter discrete distribution with associated INAR(1) process and applications. IEEE Access, 8, 91150-91162. doi:10.1109/ACCESS.2020.2993593.

[19] Eliwa, M. S., El-Morshedy, M., & Shahen, H. S. (2023). Modelling dispersed count data under various shapes of failure rates: A discrete probability analogue of odd Lomax generator. Filomat, 37(18), 6177-6203.

[20] Eliwa, M. S., Tahir, M. H., Hussain, M. A., Almohaimeed, B., Al-Bossly, A., & El-Morshedy, M. (2023). Univariate probability-G classes for scattered samples under different forms of hazard: continuous and discrete version with their inferences tests. Mathematics, 11(13), 2929.

[21] El-Morshedy, M., Eliwa, M. S., & Altun, E. (2020). Discrete Burr-Hatke distribution with properties, estimation methods and regression model. IEEE Access, 8, 74359-74370.

[22] Gabanakgosi, M., & Oluyede, B. (2023). The Topp-Leone type II exponentiated half logistic-G family of distributions with applications. International Journal of Mathematics in Operational Research, 25(1), 85-117.

[23] Gómez-Déniz, E., & Calderín-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. Journal of statistical computation and simulation, 81(11), 1405-1416.

[24] Gómez-Déniz, E., & Calderín-Ojeda, E. (2011). The discrete Lindley distribution: Properties and applications. Journal of Statistical Computation and Simulation, 81, 1405-1416.

[25] Haj Ahmad, H., & Almetwally, E. M. (2022). Generating optimal discrete analogue of the generalized Pareto distribution under Bayesian inference with applications. Symmetry, 14(7), 1457.

[26] Hassan, A. S., Fayomi, A., Algarni, A., & Almetwally, E. M. (2022). Bayesian and non-Bayesian inference for unit-exponentiated half-logistic distribution with data analysis. Applied Sciences, 12(21), 11253.

[27] Hussain, T., & Ahmad, M. (2014). Discrete inverse Rayleigh distribution. Pakistan Journal of Statistics, 30, 203-222.

[28] Hussain, T., Aslam, M., & Ahmad, M. (2016). A two-parameter discrete Lindley distribution. Revista Colombiana De Estadistica, 39, 45-61.

[29] Irshad, M. R., Chesneau, C., D'cruz, V., & Maya, R. (2021). Discrete pseudo Lindley distribution: properties, estimation and application on INAR (1) process. Mathematical and Computational Applications, 26(4), 76.

[30] Jazi, M. A., Lai, C. D., & Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. Statistical Methodology, 7(2), 121-132.

[31] Jazi, M. A., Lai, C. D., & Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. Statistical Methodology, 7, 121-132.

[32] Krishna, H., & Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. Statistical Methodology, 6, 177-188.

[33] Moakofi, T., & Oluyede, B. (2023). Type II exponentiated half-logistic Gompertz-G family of distributions: properties and applications. Mathematica Slovaca, 73(3), 785-810.

[34] Moakofi, T., Oluyede, B., & Chipepa, F. (2021). Type II exponentiated half-logistic Topp-Leone Marshall-Olkin-G family of distributions with applications. Heliyon, 7(12).

[35] Moakofi, T., Oluyede, B., & Chipepa, F. (2021). Type II exponentiated half-logistic-Topp-Leone-G power series class of distributions with applications. Pakistan Journal of Statistics and Operation Research, 885-909.

[36] Para, B. A., & Jan, T. R. (2016). Discrete version of log-logistic distribution and its applications in genetics. International Journal of Modern Mathematical Sciences, 14, 407-422.

[37] Para, B. A., & Jan, T. R. (2016). On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science. Biometrics & Biostatistics International Journal, 4, 1-15.

[38] Poisson, S. D. (1837). Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités. Paris: Bachelier.

[39] Riley, K. F., Hobson, M. P., & Bence, S. J. (2006). Mathematical methods for physics and engineering: a comprehensive guide. Cambridge university press.

[40] Roy, D. (2004). Discrete Rayleigh distribution. IEEE Transactions on Reliability, 53(2), 255-260.