



---

*Research article*

**A weighted hybrid discrete probability model: Mathematical framework, statistical analysis, estimation techniques, simulation-based ranking, and goodness-of-fit evaluation for over-dispersed data**

**Mahmoud El-Morshedy<sup>1</sup>, Mohamed S. Eliwa<sup>2,\*</sup>, Mohamed El-Dawoody<sup>1</sup> and Hend S. Shahren<sup>3</sup>**

<sup>1</sup> Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

<sup>2</sup> Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia

<sup>3</sup> Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

\* **Correspondence:** Email: m.eliwa@qu.edu.sa.

**Abstract:** Data modeling played a crucial role in a variety of research domains due to its widespread practical applications, especially when handling complex datasets. This study explored a specific discrete distribution, characterized by a single parameter, developed using the weighted combining discretization method. The statistical properties of this distribution were rigorously derived and expressed mathematically, covering essential aspects such as moments, skewness, kurtosis, covariance, index of dispersion, order statistics, entropies, mean residual life, residual coefficient of variation function, stress-strength models, and premium principles. These properties highlighted the model's suitability for analyzing right-skewed data with heavy tails, making it a powerful tool for probabilistic modeling in situations where data exhibited overdispersion and increasing failure rates. The research introduced a range of estimation techniques, including maximum product of spacings, method of moments, Anderson-Darling, right-tail Anderson-Darling, maximum likelihood, least squares, weighted least squares, Cramer-Von-Mises, and percentile, each explained in detail. A ranking simulation study was performed to assess the performance of these estimators, with ranking techniques used to determine the most effective estimator across various sample sizes. The study further applied the proposed model to real-world datasets, demonstrating its ability to address complex data scenarios and showcasing its superior performance in comparison to traditional models such as the geometric, Poisson, and negative binomial distributions. Overall, the results emphasized the proposed model's potential as a versatile and effective tool for modeling over-dispersed and skewed data, with promising implications for future research in diverse fields.

**Keywords:** statistical model; weighted combining discretization approach; failure analysis; entropy;

residual coefficient of variation; exponential premium principle; estimation methods; simulation; data analysis

---

**Abbreviations:** HDWC: Hybrid discrete weighted combining; PMF: Probability mass function; CDF: Cumulative distribution function; HRF: Hazard rate function; RHRF: Reversed hazard rate function; IFR: Increasing failure rate; MGF: Moment generating function; PGF: Probability generating function; COV: Covariance; IOD: Index of dispersion; Sk and Ku: Skewness and kurtosis; MRL: Mean residual life; SF: Survival function; RCOV: Residual coefficient of variation; SSM: Stress-strength model; EVP: Expected value principle; EPP: Exponential premium principle; MLE: Maximum likelihood estimation; ME: Moment estimation; MPSE: Maximum product of spacings estimator; PCE: Percentile estimator; ADE: Anderson-Darling estimator; RADE: Right-tail Anderson-Darling estimator; LSE: Least-squares estimator; WLSE: Weighted least-squares estimator; CVME: Cramer-Von-Mises estimator; MCMC: Markov chain Monte Carlo; MSE: Mean squared errors; GOF: Goodness-of-fit; AIC: Akaike information criterion; CAIC: Corrected Akaike information criterion; BIC: Bayesian information criterion; HQIC: Hannan-Quinn information criterion; D.F: Degree of freedom; EDS: Empirical descriptive statistics; KS: Kolmogorov-Smirnov statistic; P-P: Probability-Probability

## 1. Introduction

In recent years, there has been remarkable progress in developing discrete probability distributions, which are essential for modeling count data across a wide range of disciplines, including statistical quality control, engineering, epidemiology, biology, agriculture, medicine, and social sciences. These distributions are used to represent diverse scenarios, such as the number of species in an ecological community, customer purchases, regional accidents, or votes received by candidates in an election. Count data, which generally represent the occurrence of random events within a specific time frame, are often viewed as the outcomes of underlying continuous-time count processes. Discretizing continuous data is frequently preferred because it simplifies analysis and interpretation in many practical applications. Such data, consisting of nonnegative integers (including zero), typically lack an upper limit, making them ideal candidates for modeling with discrete probability distributions.

While classical models like the Poisson, geometric, and negative binomial distributions have long been used for this purpose, they come with notable limitations. To address these challenges, statisticians have developed various methods for generating discrete distributions as counterparts to continuous ones, enriching the tools available for analyzing count data. These include the survival discretization approach (see Nakagawa and Osaki [1]), infinite series discretization technique (see Kulasekera and Tonkyn [2] and Sato et al. [3]), round-off method (see Smith [4]), binning method (see Roederer et al. [5]), minimizing Cramer-von Mises-type distances method (see Barbiero and Hitaj [6]), reversed hazard function discretization method (see Ghosh et al. [7]), compound two-phase approach (see Chakraborty [8]), and others. For more details (see Kotsiantis and Kanellopoulos [9]). These methods offer different approaches to discretizing continuous distributions, each with its own advantages and implications. For a comprehensive understanding of these discretization methods and their differences, interested readers are encouraged to refer to Chakraborty [8], which provides a

detailed survey on the topic.

Additionally, there is another approach for proposing a discrete probability model known as the combining approach. This method involves combining two probability mass functions (PMFs) to create a new PMF, a technique commonly used in probability theory and statistics. The combining approach is often applied in situations where the outcomes of interest can be modeled by multiple distributions, aiming to produce a combined distribution that captures the characteristics of both original distributions. Various methods can be employed for combining PMFs, including convolution, mixture models, and weighted averages. Convolution, a mathematical operation, merges two functions to generate a third function. In the context of probability distributions, convolution merges the PMFs of two random variables to derive the PMF of their sum, as described in Casella and Berger [10], which is particularly beneficial for dealing with independent random variables. Mixture models, on the other hand, are probabilistic models representing a variable's distribution as a mixture of several component distributions. Each component distribution is weighted by a mixing coefficient, resulting in an overall distribution that is a weighted sum of these components, as outlined in Bishop [11]. This approach is valuable for modeling data originating from different subpopulations, each with its distinct distribution. Furthermore, weighted averages can also be employed to combine PMFs by taking the weighted averages of their values, as discussed in Gelman et al. [12]. This method is frequently utilized in Bayesian statistics, where prior distributions are combined with likelihood functions to yield a posterior distribution. The weighted combining approach is vital for analyzing count-type data across diverse sectors, especially in situations where accurately measuring quality characteristics on a numerical scale is challenging. This challenge is particularly pronounced in statistical quality control scenarios involving defects in manufactured items or service error frequencies. By integrating dependencies or correlations between observations within integer-valued time series data models, this approach enhances the accuracy and reliability of performance evaluation, addressing a critical aspect in count-type data analysis where the assumption of observation independence in traditional control charts may not hold. As a result, the discrete weighted PMF stands out as a potent tool that boosts the flexibility, accuracy, and interpretability of probabilistic models, making it indispensable in various analytical and modeling tasks across different domains.

Building upon these principles, the proposed paper introduces a new PMF with a single parameter, known as the hybrid discrete weighted combining (HDWC) distribution. The HDWC model is distinct within the family of statistical models due to its capacity to handle complex data behaviors that other classical distributions may not effectively represent. While the geometric, Poisson, and negative binomial distributions each have their strengths, they struggle with data exhibiting heavy tails, overdispersion, or extreme values. In contrast, the HDWC model is specifically designed to address these challenges. The geometric distribution is useful for modeling the number of trials until the first success in Bernoulli trials, where the data follows a memoryless property. However, it cannot account for overdispersion or heavy tails, limitations that the HDWC model overcomes with its ability to model skewed and heavy-tailed data. Similarly, while the Poisson distribution is widely used to model count data with a constant mean and variance, it fails to handle overdispersion or data with greater variability than the mean. The HDWC model excels in this area, as it accommodates both overdispersion and outlier-sensitive, heavy-tailed data, offering a more robust fit for real-world datasets. The negative binomial distribution extends the Poisson model to address overdispersion by

introducing an additional parameter. However, it assumes a specific variance structure, which may not always match observed data. The HDWC model offers greater flexibility, effectively capturing leptokurtic shapes and data with increasing failure rates, making it especially valuable for analyzing extreme data and phenomena with high variability.

The HDWC model presents distinct advantages compared to the geometric, Poisson, and negative binomial distributions, especially in its capacity to handle data with heavy tails, overdispersion, and extreme values. Its closed-form statistical properties and versatility in addressing complex datasets make it a valuable tool for researchers in fields like engineering sustainability, actuarial analysis, and medical data. Thus, the motivation for introducing the HDWC model can be summarized as follows:

- It allows for expressing statistical properties in closed forms with a simple formulation.
- It is suitable for modeling positively skewed data with a leptokurtic shape.
- It can be used to analyze overdispersed data.
- It enables the discussion of extreme and outlier data, fitting well under heavy-tailed shapes.
- It is considered a suitable choice for researchers analyzing the phenomenon of increasing failure rates.

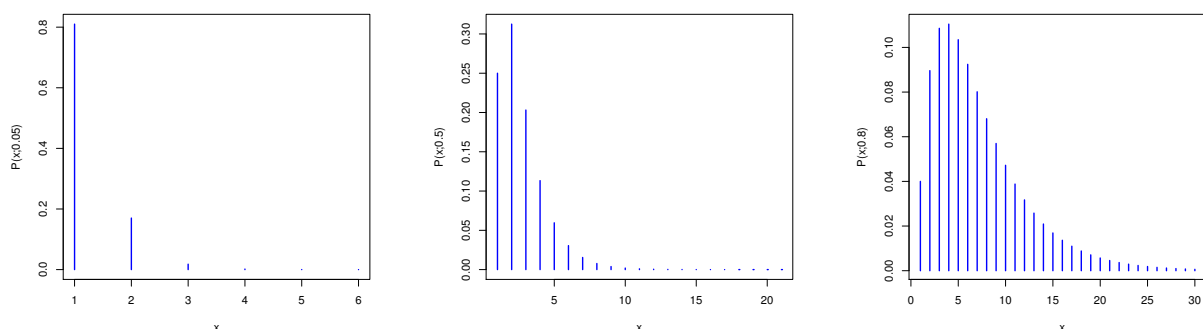
The paper is organized as follows: Section 2 introduces the mathematical framework of the HDWC distribution, while Section 3 explores its key statistical properties. Section 4 derives various mathematical estimation methods for practical applications. In Section 5, these methods are evaluated and compared using a ranked-simulation approach to identify optimal model parameters across different sample sizes. Section 6 demonstrates the HDWC model's flexibility in fitting real-world datasets, highlighting its performance against well-established PMFs in the statistical literature, with applications in system reliability and treatment analysis. Finally, Section 7 concludes with a summary of key findings and suggestions for future research directions.

## 2. Mathematical framework of the hybrid discrete weighted combining distribution

Let  $f_1(x; \theta) = (1 - \theta)\theta^x$  represent a geometric distribution where the probability of success in each trial is  $\theta$ , and the probability of failure is  $1 - \theta$ . On the other hand,  $f_2(x; \theta) = (1 - \theta^2)\theta^{2x}$  can be viewed as a geometric distribution applied specifically to even integers  $x$ , rather than the conventional geometric distribution that applies to all nonnegative integers. Let  $f(x; \theta) = 2f_1(x; \theta) - f_2(x; \theta)$ , where  $f_1(x; \theta)$  and  $f_2(x; \theta)$  represent two PMFs such that  $2f_1(x; \theta) - f_2(x; \theta) \geq 0$  for all  $x = 0, 1, 2, \dots$  and  $0 < \theta < 1$ . Then, the PMF of the HDWC distribution can be formulated as

$$f(x; \theta) = 2(1 - \theta)\theta^x - (1 - \theta^2)\theta^{2x}; \quad x = 0, 1, 2, \dots, \quad 0 < \theta < 1. \quad (2.1)$$

The conditions for a valid PMF, including nonnegativity and normalization, have been verified. Figure 1 illustrates the potential PMF shapes of the HDWC distribution, revealing unimodal right-skewed patterns with extended right tails.



**Figure 1.** The PMF of the HDWC model.

The corresponding cumulative distribution function (CDF), denoted as  $F(x; \theta)$ , can be written as

$$F(x; \theta) = \sum_{t=0}^x \Pr(X = t; \theta) = 1 - 2\theta^{x+1} + \theta^{2x+2}; \quad x = 0, 1, 2, \dots, \quad 0 < \theta < 1. \quad (2.2)$$

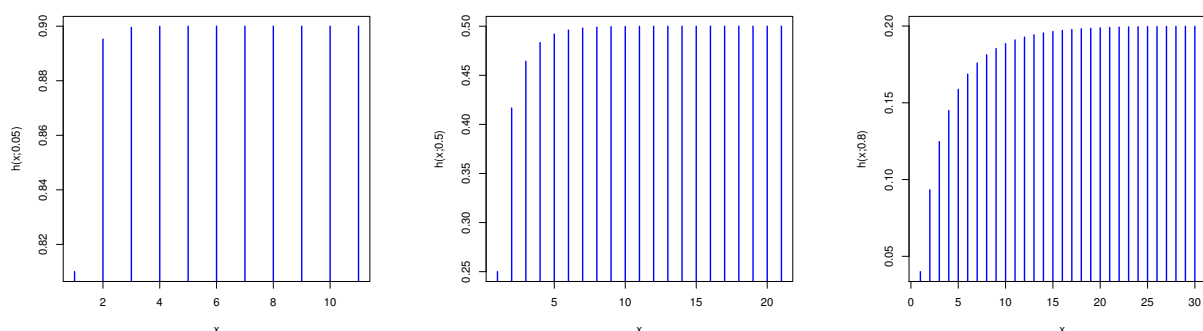
The hazard rate function (HRF) can be listed as

$$h(x; \theta) = \frac{(1 - \theta)[(1 + \theta)\theta^x - 2]}{\theta^x - 2}; \quad x = 0, 1, 2, \dots, \quad 0 < \theta < 1. \quad (2.3)$$

The reversed hazard rate function (RHRF) can be reported as

$$r(x; \theta) = \frac{2(1 - \theta)\theta^x - (1 - \theta^2)\theta^{2x}}{1 - 2\theta^{x+1} + \theta^{2x+2}}; \quad x = 0, 1, 2, \dots, \quad 0 < \theta < 1. \quad (2.4)$$

Figure 2 presents the HRF for the HDWC model across different parameter values. The analysis indicates that the failure rate of the HDWC distribution exhibits a monotonically increasing or increasing-constant shape. A log-concave distribution has discrete increasing failure rate (IFR) (see Barlow [13]). The relationship between the IFR and other discrete aging classes like new better than used (NBU) can be explored in Lai and Xie [14].



**Figure 2.** The HRF of the HDWC model.

The finding that the failure rate of the HDWC distribution shows either a consistently increasing trend or a steady, constant increase has significant implications across various domains of data analysis and analytics. In reliability engineering, this behavior aids in predicting system failures and scheduling maintenance. Healthcare analytics can leverage it to assess medical treatment effectiveness over time. Financial risk management benefits from understanding investment failure rates, while manufacturing and quality control can optimize processes based on product failure trends. Additionally, in survival analysis and event history studies, this insight proves valuable for analyzing time-to-event data in epidemiology, social sciences, and marketing, enhancing decision-making and risk mitigation strategies across diverse sectors.

### 3. Statistical characteristics of the hybrid discrete weighted combining distribution

#### 3.1. Recurrence relation for probabilities and long-tailedness

The PMF described in Eq (2.1) exhibits log-concavity across all  $\theta$  values, with the ratio of successive probabilities  $\frac{f(x+1;\theta)}{f(x;\theta)}$  being a decreasing function in  $x$  for all  $\theta$  values. Given that for a discrete distribution, the value of  $B = \lim_{x \rightarrow \infty} \frac{f(x+1;\theta)}{f(x;\theta)}$  indicates the relative long-tailedness of the distribution, and comparing it to the Poisson distribution (which has  $B = 0$ ), we can conclude that the HDWC distribution exhibits longer tails than the Poisson distribution. Long-tailedness refers to the slow decay of probabilities in the tail of a distribution, where the probability of observing large values remains higher compared to short-tailed distributions. Unlike light-tailed distributions such as the Poisson, long-tailed distributions decay more slowly, with their tail probabilities taking longer to approach zero. This characteristic is particularly important in fields like risk assessment, finance, and insurance, where rare but extreme events can have significant consequences.

#### 3.2. Log-concavity

The HDWC distribution exhibits log-concavity. Specifically, since the ratio of successive probabilities is a strictly decreasing function for  $x \geq 0$ , it follows that

$$f(x; \theta)^2 > f(x-1; \theta).f(x+1; \theta), \quad x = 0, 1, 2, \dots$$

This condition is sufficient for the log-concavity of a PMF. A log-concave distribution exhibits an increasing failure hazard rate, is strongly unimodal, has all moments that exist, retains its log-concavity when truncated, and its convolution with any other discrete distribution remains both unimodal and log-concave.

#### 3.3. The moments and associated metrics

The moment generating function (MGF) and probability generating function (PGF) are fundamental tools in probability theory and statistics with diverse practical applications. The MGF, denoted as  $M_X(t)$ , for a random variable  $X$ , is defined as the expected value of  $e^{tX}$  where  $t$  is a parameter. It is a powerful tool for finding moments of a distribution. Specifically, the  $n$ th moment of  $X$  can be obtained by differentiating the MGF  $n$  times and evaluating at  $t = 0$ , i.e.,  $E(X^n) = M_X^{(n)}(0)$ . This makes the MGF useful for calculating moments without directly integrating over the distribution. The PGF, denoted as  $\psi_X(z)$ , for  $X$  is defined as the expected value of  $z^X$  where  $z$  is a parameter typically restricted to

the interval  $[0, 1]$ . It is particularly useful for discrete random variables and is employed to derive probabilities of various events related to the random variable  $X$ . For the HDWC distribution, the MGF and PGF can be formulated as

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \Pr(X = x; \theta) = \frac{2(1 - \theta)}{1 - \theta e^t} - \frac{1 - \theta^2}{1 - \theta^2 e^t}, \quad (3.1)$$

and

$$\psi_X(t) = \sum_{x=0}^{\infty} z^x \Pr(X = x; \theta) = \frac{2(1 - \theta)}{1 - \theta z} - \frac{1 - \theta^2}{1 - \theta^2 z}. \quad (3.2)$$

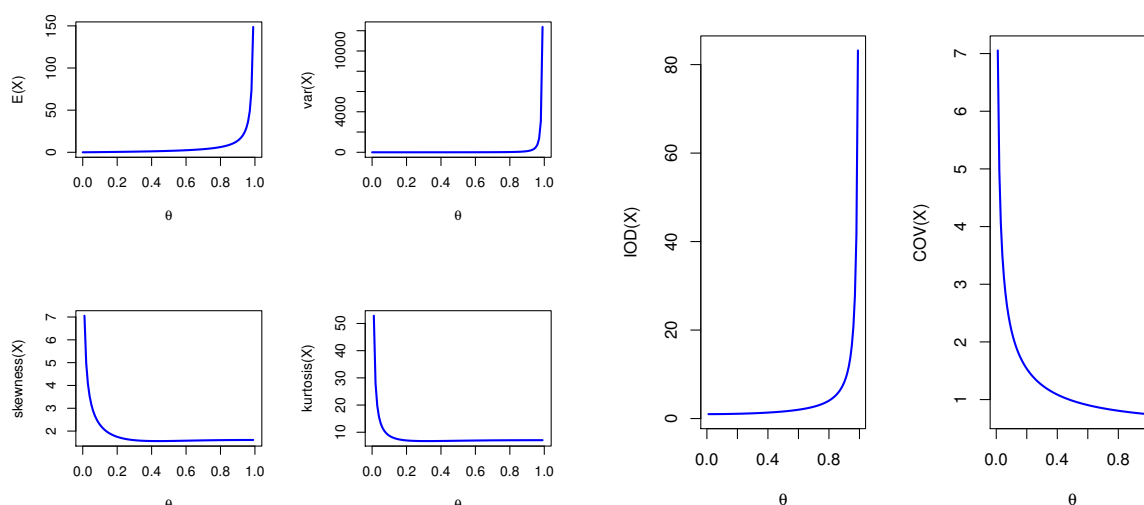
MGFs are extensively employed in statistical analysis to compute moments, crucial for determining central tendencies, variances, index of dispersion and other statistical properties of distributions, while PGFs are utilized in analyzing discrete distributions, particularly in problems involving counting and combinatorial aspects. In reliability engineering, MGFs aid in analyzing component and system lifetimes, predicting failure rates, and assessing system reliability over time. Queueing theory benefits from both MGFs and PGFs, assisting in the analysis of arrival and service times, queue lengths, and overall system performance metrics. Actuarial science utilizes MGFs for modeling insurance risks, mortality rates, and financial aspects related to insurance and pensions. In epidemiology and public health, probability generating functions play a key role in modeling disease spread, calculating infection probabilities, and evaluating the impact of interventions. In machine learning and data science, these functions are integral to probability distributions used in various models, such as the Poisson distribution for count data and the exponential distribution for survival analysis. For the HDWC distribution, the raw moments, say  $E(X^n)$ , can be formulated as

$$E(X^n) = \sum_{x=0}^{\infty} [(x+1)^n - x^n] (1 - F(x; \theta)) = \sum_{x=0}^{\infty} [(x+1)^n - x^n] \theta^{x+1} (2 - \theta^{x+1}). \quad (3.3)$$

Using Eq (3.3), the initial four moments can be represented in closed forms as

$$\begin{aligned} E(X) &= \frac{\theta(\theta + 2)}{1 - \theta^2}, \quad E(X^2) = \frac{\theta(\theta^3 + 6\theta^2 + 5\theta + 2)}{(1 - \theta^2)^2}, \quad E(X^3) = \frac{\theta(\theta^5 + 14\theta^4 + 28\theta^3 + 32\theta^2 + 13\theta + 2)}{(1 - \theta^2)^3}, \\ E(X^4) &= \frac{\theta(\theta^7 + 30\theta^6 + 111\theta^5 + 230\theta^4 + 219\theta^3 + 122\theta^2 + 29\theta + 2)}{(1 - \theta^2)^4}. \end{aligned}$$

They are called “raw” because they are calculated directly from the data without any adjustments or transformations. The skewness, kurtosis, covariance (COV), and index of dispersion (IOD) can be calculated from the first four moments. Figure 3 illustrates several descriptive statistics pertaining to the HDWC distribution.



**Figure 3.** Some descriptive statistics for the HDWC model.

The data depicted in Figure 3 indicates that the HDWC model is adept at analyzing positively skewed data with a leptokurtic shape and can also be used to investigate phenomena associated with overdispersion “ $\text{IOD}(X) > 1$ ”. Moreover, a positive covariance signifies that the variables tend to co-vary in the same direction.

### 3.4. Order statistics

Order statistics refer to the statistical analysis of ranked data points within a dataset. They are particularly useful in understanding the distribution and characteristics of extreme values, such as the minimum and maximum values, as well as percentiles and quartiles. By arranging data points in ascending or descending order, order statistics allow researchers to analyze the spread and central tendencies of a dataset more effectively. For instance, in environmental studies, order statistics can help identify the most extreme weather events, such as the hottest day of the year or the coldest month, which is crucial for risk assessment, infrastructure planning, and climate change analysis. Suppose we have a random sample  $X_1, X_2, \dots, X_n$  of size  $n$  from the CDF of the HDWC model. The CDF of the  $i^{\text{th}}$  order statistic can be expressed as

$$\begin{aligned}
 F_{i:n}(x; \theta) &= \sum_{k=i}^n \binom{n}{k} F(x; \theta)^k (1 - F(x; \theta))^{n-k} = \sum_{k=i}^n \sum_{j=0}^k (-1)^j \binom{n}{k} \binom{k}{j} (1 - F(x; \theta))^{n-k+j} \\
 &= \sum_{k=i}^n \sum_{j=0}^k \sum_{r=0}^{n-k+j} (-1)^{j+r} \binom{n}{k} \binom{k}{j} \binom{n-k+j}{r} 2^{n-k+j-r} \theta^{(x+1)(n-k+j+r)} \\
 &= \sum_{k=i}^n \sum_{j=0}^k \sum_{r=0}^{n-k+j} w_{j,r,k} \theta^{(x+1)(n-k+j+r)},
 \end{aligned}$$

where

$$w_{j,r,k} = (-1)^{j+r} \binom{n}{k} \binom{k}{j} \binom{n-k+j}{r} 2^{n-k+j-r}.$$



The PMF of  $i^{th}$  order statistics for  $x = 1, 2, \dots$  can be formulated as

$$\begin{aligned} \Pr(X_{i:n} = x; \theta) &= F_{i:n}(x; \theta) - F_{i:n}(x-1; \theta) = \sum_{k=i}^n \sum_{j=0}^k \sum_{r=0}^{n-k+j} w_{j,r,k} \theta^{(x+1)(n-k+j+r)} \\ &= \sum_{k=i}^n \sum_{j=0}^k \sum_{r=0}^{n-k+j} w_{j,r,k} \theta^{n-k+j+r} \theta^{x(n-k+j+r)} (\theta^{n-k+j+r} - 1), \end{aligned} \quad (3.4)$$

and

$$\Pr(X_{i:n} = 0; \theta) = F_{i:n}(0; \theta) = \sum_{k=i}^n \binom{n}{k} (1 - \theta^2)^{2k} (1 - (1 - \theta^2)^{n-k}).$$

By referring to Eq (3.4), it is possible to calculate the  $r^{th}$  moments of order statistics, and subsequently, derive the L-moments. L-moments are statistical measures used to describe the shape, location, and scale of a probability distribution. They are particularly useful in situations where traditional moments like mean and variance may not adequately capture the distribution's characteristics, especially in cases of skewed or heavy-tailed data. L-moments are derived from linear combinations of order statistics, providing information about the distribution's central tendency, dispersion, and skewness. They are commonly used in hydrology, meteorology, and other fields where analyzing extreme values and fitting distributions to observed data is important for risk assessment and decision-making.

### 3.5. Various entropies

Residual entropy and cumulative residual entropy are measures used in information theory to quantify the uncertainty or unpredictability of a random variable or system. Residual entropy captures the remaining uncertainty after observing a sequence of events, while cumulative residual entropy accounts for the accumulated uncertainty across multiple observations. These measures are particularly useful in assessing the information content or randomness in data sequences, such as in cryptography, where residual entropy helps evaluate the security of encryption algorithms based on the unpredictability of ciphertexts. The residual entropy of variable  $X$  and its cumulative expression can be formulated, respectively, as follows:

$$\mathcal{E}(X) = - \sum_{x=0}^{\infty} F(x) \log(F(x)),$$

and

$$C\mathcal{E}(X) = - \sum_{x=0}^{\infty} (1 - F(x)) \log(1 - F(x)).$$

By utilizing algebraic techniques with respect to the HDWC model, we can derive

$$\mathcal{E}(X) = \sum_{k=0}^{\infty} \sum_{j=0}^k \sum_{i=0}^j \frac{(-1)^i \binom{k}{j} \binom{j}{i} \theta^{i+k}}{1 - \theta^{i+k}} - \sum_{k=0}^{\infty} \sum_{j=0}^{k+1} \sum_{i=0}^j \frac{(-1)^i \binom{k+1}{j} \binom{j}{i} \theta^{i+k+1}}{1 - \theta^{i+k+1}}, \quad (3.5)$$

and

$$CE(X) = \frac{-\theta(2\theta^2 + 3\theta + 2)\log(\theta)}{(1 - \theta^2)^2} + \sum_{k=0}^{\infty} \sum_{j=0}^{k+1} \frac{(-1)^{k+1} \binom{k+1}{j} \theta^{j+1} (2 - \theta - \theta^{j+2})}{(k+1)(1 - \theta^{j+1})(1 - \theta^{j+2})}. \quad (3.6)$$

On the other hand, Shannon entropy and Rényi entropy are also measures of uncertainty in information theory but differ in their sensitivity to outliers and the concentration of probability mass. Shannon entropy, the most widely used entropy measure, characterizes the average information content in a probability distribution, providing insights into the diversity and spread of data. In contrast, Rényi entropy generalizes Shannon entropy by introducing a parameter that controls the emphasis on rare events, making it more robust to outliers and suitable for analyzing distributions with heavy tails or extreme values. These entropy measures find applications in various fields such as data compression, pattern recognition, and statistical physics, where understanding the information content and structure of data is critical for modeling and decision-making. The Shannon entropy for the HDWC can be formulated as

$$S(X) = -\log(2(1 - \theta)) - \log(\theta) \times \frac{2\theta + 1}{1 - \theta^2} + \sum_{k=0}^{\infty} \frac{(\frac{\theta+1}{2})^{k+1}}{k+1} \left[ \frac{2(1 - \theta)\theta^{k+2}}{1 - \theta^{k+2}} - \frac{2(1 - \theta^2)\theta^{k+3}}{1 - \theta^{k+3}} \right], \quad (3.7)$$

where

$$S(X) = -E[\log(\Pr(X = x; \theta))] = -\log(2(1 - \theta)) - \log(\theta)E(X + 1) + \sum_{k=0}^{\infty} \frac{(\frac{\theta+1}{2})^{k+1} E(\theta^{(X+1)(k+1)})}{k+1},$$

$$\begin{aligned} -\log(\Pr(X = x; \theta)) &= -\log(2(1 - \theta)) - (x + 1)\log(\theta) - \log\left(1 - \frac{1 + \theta}{2}\theta^{x+1}\right) \\ &= -\log(2(1 - \theta)) - (x + 1)\log(\theta) + \sum_{k=0}^{\infty} \frac{(\frac{\theta+1}{2})^{k+1} \theta^{(x+1)(k+1)}}{k+1}. \end{aligned}$$

The Rényi entropy of variable  $X$ , obtained through the use of generalized binomial expansion, is defined as

$$I_R(\gamma) = \frac{1}{1 - \gamma} \log\left(\sum_{x=0}^{\infty} f(x; \theta)^\gamma\right) = \frac{1}{1 - \gamma} \left\{ \gamma \log(2(1 - \theta)) + \log\left(\sum_{i=0}^{\infty} \frac{(-1)^i (\frac{\theta+1}{2})^i \binom{\gamma}{i}}{1 - \theta^{i+\gamma}}\right) \right\}, \quad (3.8)$$

where  $\gamma > 0$  and  $\gamma \neq 1$ . Some numerical entropy of the HDWC model can be listed in Table 1.

**Table 1.** Numerical entropy of the HDWC distribution.

$\gamma \downarrow \theta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	0.8219	1.1588	1.4393	1.7058	1.9796	2.2803	2.6355	3.1010	3.8412
2	0.3778	0.7073	0.9942	1.2568	1.5185	1.8051	2.1485	2.6065	3.3508
5	0.2633	0.5538	0.8552	1.1247	1.3610	1.6250	1.9632	2.4234	3.1698

Table 1 quantifies the uncertainty linked to its parameter combinations. Entropy increases with both  $\gamma$  and  $\theta$ , indicating greater uncertainty. For fixed  $\gamma$ , entropy grows nonlinearly with  $\theta$ , while for fixed  $\theta$ , it decreases as  $\gamma$  increases. Smaller  $\gamma$  values, associated with heavier tails, lead to higher entropy and greater randomness, whereas larger  $\gamma$  values result in a more concentrated distribution and lower entropy. Larger  $\theta$  values amplify entropy, reflecting a broader and less predictable distribution, with entropy rising faster at higher  $\theta$ . The interplay between  $\gamma$  and  $\theta$  shapes the overall randomness. For example, a small  $\gamma$  and large  $\theta$  produce the highest entropy, while the reverse yields the lowest. These insights are valuable for applications in engineering, medicine, and actuarial sciences, where entropy helps assess the suitability of the HDWC distribution for modeling uncertainty. The relationship between entropy and parameters guides parameter selection and model fitting, leveraging the distribution's flexibility to match specific data or phenomena.

### 3.6. Mean residual life and its residual coefficient of variation function

The mean residual life (MRL) and its variance function are important concepts in survival analysis, which deals with the time until an event of interest occurs. In the case of discrete random variables, such as in discrete survival analysis, these concepts are adapted to work with discrete time points. Considering  $F(x; \cdot)$  as the CDF of an element with a finite first moment, where  $X$  denotes the random variable associated with  $F(x; \cdot)$ , in the discrete context, the MRL, say  $\varpi(k; \theta)$ , is defined as follows

$$\varpi(k; \theta) = E(X - k | X \geq k) = \frac{1}{\Pr(X \geq k)} \sum_{x=k}^{\infty} (x - k) \Pr(X = x); \quad k = 0, 1, 2, \dots \quad (3.9)$$

Consider  $X$  to be an HDWC random variable, then the MRL can be formulated as

$$\varpi(k; \theta) = \frac{2\theta^{k+2} - \theta^{2k+2} + 2\theta^{k+1}}{(\theta^2 - 1)(\theta^{2k} - 2\theta^k)}; \quad k = 0, 1, 2, \dots \quad (3.10)$$

The MRL of the HDWC distribution is a decreasing function. The relationship between the HRF and MRL can be established as

$$h(k; \theta) = 1 - \frac{\varpi(k; \theta)}{1 + \varpi(k + 1; \theta)}; \quad k = 0, 1, 2, \dots$$

Thus, the survival function (SF) of the HDWC model can be expressed as

$$S(k; \theta) = \prod_{0 \leq i \leq k} \frac{\varpi(i; \theta)}{1 + \varpi(i + 1; \theta)}; \quad \varpi(0; \theta) = E(X), \quad k = 0, 1, 2, \dots \quad (3.11)$$

The function for the variance of residual life (VRL), denoted as  $\Upsilon_{VRL}(k)$ , is defined as

$$\begin{aligned} \Upsilon_{VRL}(k; \theta) &= \text{Var}(X | X \geq k) = E(X^2 | X \geq k) - [E(X | X \geq k)]^2 \\ &= \frac{2}{2\theta^k - \theta^{2k}} \left[ \frac{[(k-1)\theta^2 - k]\theta^{2k+2}}{(\theta^2 - 1)^2} - \frac{2[(k-1)\theta - k]\theta^{k+1}}{(\theta - 1)^2} \right] - (2k-1)\varpi(k; \theta) - [\varpi(k; \theta)]^2. \end{aligned}$$

The random variable  $X$  exhibits increasing (decreasing) VRL if

$$\Upsilon_{VRL}(k + 1; \theta) \leq (\geq) \varpi(k; \theta) [1 + \varpi(k + 1; \theta)].$$

The residual coefficient of variation (RCOV), denoted as  $\Xi(k; \theta)$ , can be explicitly derived as

$$\Xi(k; \theta) = \sqrt{\Upsilon_{VRL}(k; \theta) / \varpi(k; \theta)}; \quad k = 0, 1, 2, \dots$$

The HRF, MRL, and VRL of the HDWC model are interconnected in the following manner

$$\Upsilon_{VRL}(k+1; \theta) - \Upsilon_{VRL}(k; \theta) = h(k; \theta) [\Upsilon_{VRL}(k+1; \theta) - \varpi(k; \theta)(1 + \varpi(k+1; \theta))].$$

Further, the HRF, MRL, VRL, and RCOV of the HDWC distribution are interconnected as

$$\Upsilon_{VRL}(k+1; \theta) - \Upsilon_{VRL}(k; \theta) = h(k; \theta) [\varpi(k+1; \theta)]^2 \left\{ [\Xi(k+1; \theta)]^2 - \frac{\varpi(k; \theta)(1 + \varpi(k+1; \theta))}{[\varpi(k+1; \theta)]^2} \right\}. \quad (3.12)$$

### 3.7. Stress-strength model

The stress-strength model (SSM) is a statistical framework used to assess reliability and risk in systems involving discrete random variables. It compares the strength of a system or material (represented by a random variable) against the stress it is subjected to, often in engineering, manufacturing, and quality control contexts. For example, in mechanical engineering, this model can evaluate the likelihood of a component failing under varying stress conditions. Similarly, in quality control, it helps determine if a product meets certain strength criteria given manufacturing variations. The model is also applicable in pharmaceuticals to assess the efficacy of drug formulations concerning patient-specific factors. Let  $X_1 \sim \text{HDWC}(\theta_1)$  and  $X_2 \sim \text{HDWC}(\theta_2)$ , then

$$\begin{aligned} \Pr(X_1 \leq X_2) &= \sum_{x=0}^{\infty} \Pr(X_1 \leq X_2 | X_2 = x) \Pr(X_2 = x) = \sum_{x=0}^{\infty} F_{X_1}(x; \theta_1) \Pr(X_2 = x; \theta_2) \\ &= 1 - \frac{2\theta_1(1 - \theta_2)}{1 - \theta_1\theta_2} + \frac{2\theta_1(1 - \theta_2^2)}{1 - \theta_1\theta_2^2} + \frac{2\theta_1^2(1 - \theta_2)}{1 - \theta_2\theta_1^2} - \frac{\theta_1^2(1 - \theta_2^2)}{1 - \theta_1^2\theta_2^2}. \end{aligned} \quad (3.13)$$

By quantifying the relationship between stress and strength through probability distributions, the SSM aids in decision-making regarding design tolerances, product specifications, and risk mitigation strategies.

### 3.8. Premium principles

Insurance premiums are determined using premium principles, which consider the risk levels associated with different events. Over time, various premium principles have been developed. In this section, we introduce some of these principles, assuming that the loss distribution follows the HDWC distribution. Within this context, let  $\lambda \geq 0$  represent the risk loading parameter.

#### 3.8.1. Expected value principle

The expected value principle (EVP) is a cornerstone in insurance and risk management, guiding the calculation of insurance premiums. It asserts that the premium assessed for insurance protection must equate to the anticipated losses' expected value, modified by a risk loading component. The EVP can be articulated as

$$\text{EVP}(\lambda; \cdot) = (1 + \lambda)E(Z),$$

where  $\text{EVP}(\lambda; \cdot)$  represents the insurance premium,  $\lambda$  is the risk loading factor, and  $E(Z)$  denotes the expected value of losses or the distribution's expected value. The term  $(1 + \lambda)$  "Risk Loading" signifies the risk loading factor, included to cover miscellaneous expenses and ensure a profit margin for the insurer. The value of  $\lambda$  varies depending on factors like administrative costs, claims processing, underwriting, and the desired profit level of the insurer. The EVP of the HDWC distribution, denoted as  $\text{EVP}(\lambda; \theta)$ , is defined by the equation:

$$\text{EVP}(\lambda; \theta) = \frac{\theta(1 + \lambda)(\theta + 2)}{1 - \theta^2}. \quad (3.14)$$

The EVP serves as a fundamental concept in insurance pricing, applicable across diverse insurance categories such as property, liability, health, and life insurance. Insurers use the EVP to set premiums based on expected losses, adjusting for risk loading to ensure fairness in pricing. This methodology aligns premiums with the expected costs of coverage, enabling insurers to manage expenses, generate profits, and ensure that policyholders pay appropriate premiums relative to the risks transferred. From an economic standpoint, the EVP promotes risk transfer, enhances efficiency through risk pooling, and minimizes the financial impacts of uncertainty. However, it operates under the assumption of known and accurately estimated loss distributions, overlooking factors such as moral hazard and market competition that can influence pricing and market dynamics.

### 3.8.2. Exponential premium principle

The exponential premium principle (EPP) is a key aspect of insurance pricing strategies, employing an exponential function to calculate premiums. This principle acknowledges that the likelihood of encountering significant loss events escalates quickly as risk levels elevate. The EPP asserts that insurance premiums should escalate exponentially in proportion to the risk level linked to the insured event. Put differently, as the probability of severe losses increases, premiums should surge at an escalating pace to sufficiently address the potential expenses associated with such occurrences. The EPP is determined by solving for EPP in the equation:

$$\varphi(v - \text{EPP}(\lambda; \cdot)) = E(v - Z).$$

Here,  $v$  refers to an individual's wealth, and  $\varphi(t) = -e^{-\lambda t}$  signifies the exponential utility function. As a result, the EPP is derived as:

$$\text{EPP}(\lambda; \cdot) = \frac{1}{\lambda} M_Z(\lambda),$$

where  $M_Z(\lambda)$  is the MGF. The EPP of the HDWC distribution, say  $\text{EPP}(\lambda; \theta)$ , is characterized as

$$\text{EPP}(\lambda; \theta) = \frac{1}{\lambda} \left\{ \frac{2(1 - \theta)}{1 - \theta e^t} - \frac{1 - \theta^2}{1 - \theta^2 e^t} \right\}. \quad (3.15)$$

This principle finds widespread application across different insurance sectors, such as property insurance (e.g., protection against natural disasters), liability insurance (e.g., coverage for high-risk activities), and health insurance (e.g., coverage for pre-existing medical conditions). Within these sectors, premiums are frequently determined using sophisticated actuarial models that integrate the EPP to guarantee sufficient coverage for various risks.

## 4. Various estimation approaches

### 4.1. Maximum likelihood estimation (MLE)

Suppose we have a random sample  $X_1, X_2, \dots, X_k$  obtained from the HDWC model. The log-likelihood, say  $L(\underline{x}|\theta)$ , function for the HDWC distribution can be expressed as follows:

$$L(\underline{x}|\theta) = k \log(1 - \theta) + \log(\theta) \sum_{i=1}^k x_i + \sum_{i=1}^k \log(2 - (1 + \theta)\theta^{x_i}). \quad (4.1)$$

Taking the derivative of the  $L(\underline{x}|\theta)$  with respect to  $\theta$  and equating it to zero, we obtain

$$\frac{\partial L(\underline{x}|\theta)}{\partial \theta} = \frac{-k}{1 - \theta} + \frac{1}{\theta} \sum_{i=1}^k x_i - \sum_{i=1}^k \frac{\theta^{x_i}(1 + (1 + \theta) \log(\theta))}{2 - (1 + \theta)\theta^{x_i}} = 0. \quad (4.2)$$

It is not feasible to find an analytical solution for this equation. Therefore, it necessitates the use of a numerical iterative method such as the Newton-Raphson method within R software or other optimization techniques.

### 4.2. Moment estimation (ME)

If we have a random sample  $X_1, X_2, \dots, X_k$  derived from the HDWC model, subsequently, the moment estimator, referred to as  $\hat{\theta}_M$ , can be formulated as

$$\hat{\theta}_M = \frac{-\bar{X} + \sqrt{\bar{X}^2 + \bar{X}(1 + \bar{X})}}{1 + \bar{X}}, \quad (4.3)$$

where  $\bar{X}$  represents the sample mean of the observed values. Estimation techniques like MLE and ME offer powerful tools for parameter estimation, but they come with notable computational and methodological challenges, especially when implemented in R. MLE, while efficient and consistent for large datasets, often requires iterative numerical methods such as Newton-Raphson, which can face convergence issues, instability due to singular or near-singular Hessians, and challenges with non-convex likelihood surfaces. These issues are amplified in high-dimensional models or with poor initial parameter values. ME, on the other hand, provides a simpler alternative but is less efficient, more prone to bias and variance in small samples, and heavily dependent on the choice of moments, which may not capture the data's complexity.

In R, implementing these methods demands careful handling of numerical precision, algorithmic stability, and computational efficiency. Challenges include the iterative nature of MLE, handling divergence or oscillation in numerical methods, and optimizing performance for large datasets. Users must choose starting values judiciously, apply robust optimization algorithms, and leverage advanced features in R packages like `optim` and `bbmle` to mitigate these issues. By addressing these computational hurdles and understanding the trade-offs of each method, practitioners can effectively apply MLE and ME to a wide range of statistical modeling problems while ensuring reliable and accurate results.

### 4.3. Maximum product of spacings estimator (MPSE)

In this section, we explore the HDWC parameter estimation using the MPSE method with a complete sample. Suppose we have a random sample  $X_1, X_2, \dots, X_k$  obtained from the HDWC distribution. For  $h = 1, 2, \dots, m+1$ , consider the following:

$$W_h(\theta) = F(x_{(h)}|\theta) - F(x_{(h-1)}|\theta),$$

to be the uniform spacings of a random sample from the HDWC model, where  $F(x_{(0)}|\theta) = 0$ ,  $F(x_{(m+1)}|\theta) = 1$ , and  $\sum_{h=1}^{m+1} W_h(\theta) = 1$ . The MPSE of  $\theta$ , say  $\widehat{\theta}_{MPSE}$ , can be derived by maximizing the geometric mean of the spacings

$$V(\theta) = \left[ \prod_{h=1}^{m+1} W_h(\theta) \right]^{\frac{1}{m+1}}, \quad (4.4)$$

with respect to the parameter  $\theta$ .

### 4.4. Percentile estimator (PCE)

Consider  $s_h = h/(k+1)$  to be an unbiased estimator of  $F(x_{(h)}|\theta)$ . Hence, the PCE of the parameter  $\theta$ , denoted by  $\widehat{\theta}_{PC}$ , can be reported by minimizing

$$P(\theta) = \sum_{h=1}^k (x_{(h)} - \varpi(s_h))^2, \quad (4.5)$$

with respect to the parameter  $\theta$ , where  $\varpi(s_h) = F^{-1}(x_{(h)}|\theta)$  is the quantile function of the HDWC model.

### 4.5. Anderson–Darling and right-tail Anderson–Darling estimators

Consider a random sample  $X_1, X_2, \dots, X_k$  drawn from the HDWC model. The Anderson-Darling estimator (ADE) serves as another form of minimum distance estimator. The ADE for the HDWC parameter, denoted as  $\widehat{\theta}_{AD}$ , is obtained by minimizing

$$AD(\theta) = -k - \frac{1}{k} \sum_{h=1}^k (2h-1) [\log F(x_{(h)}|\theta) + \log (1 - F(x_{(h)}|\theta))]. \quad (4.6)$$

Regarding the parameter  $\theta$ , the model undergoes optimization, whereas the right-tail Anderson-Darling estimator (RADE) for the model parameter is obtained by minimizing

$$RAD(\theta) = \frac{k}{2} - 2 \sum_{h=1}^k F(x_{(h:k)}|\theta) - \frac{1}{k} \sum_{h=1}^k (2h-1) [\log (1 - F(x_{(k+1-h:k)}|\theta))], \quad (4.7)$$

with respect to the parameter  $\theta$ .

#### 4.6. Weighted (least-squares) estimators

Let's consider a random sample from the HDWC model, denoted by order statistics  $X_{(1)}, X_{(2)}, \dots, X_{(k)}$ . The least-squares estimator (LSE) of the HDWC parameter, represented as  $\widehat{\theta}_{LS}$ , is derived by solving the nonlinear equation defined as follows:

$$\sum_{h=1}^k \left[ F(x_{(h)}|\theta) - \frac{h}{k+1} \right] \frac{\partial}{\partial \theta} F(x_{(h)}|\theta) = 0, \quad (4.8)$$

with respect to the parameter  $\theta$ . The weighted LSE (WLSE), say  $\widehat{\theta}_{WLS}$ , can be derived by solving the nonlinear equation defined by

$$\sum_{h=1}^k \frac{(k+1)^2 (k+2)}{h(k-h+1)} \left[ F(x_{(h)}|\theta) - \frac{h}{k+1} \right] \frac{\partial}{\partial \theta} F(x_{(h)}|\theta) = 0, \quad (4.9)$$

with respect to the parameter  $\theta$ .

#### 4.7. Cramer-Von-Mises estimator (CVME)

The CVME occurs due to the difference between the estimated CDF and the empirical CDF. Determining the CVME of the HDWC parameter entails solving the nonlinear equation specified as follows:

$$\sum_{h=1}^k \left[ F(x_{(h)}|\theta) - \frac{2h-1}{2k} \right] \frac{\partial}{\partial \theta} F(x_{(h)}|\theta) = 0, \quad (4.10)$$

with respect to the parameter  $\theta$ .

### 5. Simulation ranking techniques: Different estimators

In this section, we evaluate the performance of MPSE, ME, ADE, MLE, LSE, RADE, PCE, CVME, and WLSE with respect to the sample size 'n' using the R software and HDWC parameters. The process of generating a random variable  $X$  from the HDWC distribution starts with generating a value  $Y$  from the continuous distribution. Next, the obtained  $Y$  value is discretized to generate  $X$ , where  $X$  is defined as the largest integer less than or equal to  $Y$ . To replicate this procedure, we conduct Markov Chain Monte Carlo (MCMC) simulations using various schemes. The evaluation is conducted through a simulation study:

- 1) Generate  $N = 10000$  samples of various sizes " $n_i; i = 1, 2, 3, 4, 5$ " from the HDWC model as follows:
  - Scheme I:  $\theta = 0.05 \mid n_1 = 20, n_2 = 100, n_3 = 250, n_4 = 500, n_5 = 750$ .
  - Scheme II:  $\theta = 0.2 \mid n_1 = 20, n_2 = 100, n_3 = 250, n_4 = 500, n_5 = 750$ .
  - Scheme III:  $\theta = 0.5 \mid n_1 = 20, n_2 = 100, n_3 = 250, n_4 = 500, n_5 = 750$ .
  - Scheme VI:  $\theta = 0.8 \mid n_1 = 20, n_2 = 100, n_3 = 250, n_4 = 500, n_5 = 750$ .
- 2) Compute the MPSE, ME, ADE, MLE, LSE, RADE, PCE, CVME, and WLSE for the 10,000 samples, say  $\widehat{\theta}_k$  for  $k = 1, 2, \dots, 10,000$ .



3) Caculate the bias and the mean squared errors (MSE) for  $N = 10,000$  samples as

$$|\text{Bias}(\theta)| = \frac{1}{N} \sum_{k=1}^N |\widehat{\theta}_k - \theta_k|, \quad \text{MSE}(\theta) = \frac{1}{N} \sum_{k=1}^N (\widehat{\theta}_k - \theta_k)^2.$$

The MSE quantifies the average squared deviation between predicted and actual values, where a lower MSE indicates closer predictions to the actual values.

4) The empirical results of simulation are listed in the Tables 2–6.

As the sample size ‘ $n$ ’ increases, the trend observed in Tables 2 to 6 shows a clear reduction in the bias of the parameter  $\theta$ , approaching zero, while the MSE of the HDWC parameter also decreases toward zero. The findings indicate that the derived estimators consistently improve in performance as the sample size increases. Across varying sample sizes, all estimation methods perform well, with Table 6 demonstrating that the ME method outperforms MLE in simulations based on a ranking approach. This ranking process evaluates the methods against criteria such as bias, variance, MSE, efficiency, consistency, and robustness. Simulated data are generated, the estimation methods are applied, and their metrics are compared to determine their effectiveness. The final ranking balances these criteria, providing insights into the most suitable method for the problem at hand.

The choice between MLE and ME ultimately depends on the specific context, including factors like data size, computational complexity, and model requirements. Each method offers distinct advantages: MLE is celebrated for its efficiency, consistency, and precision, particularly in large datasets, making it a powerful choice when computational resources and well-defined likelihood functions are available. On the other hand, ME is a simpler, computationally lighter alternative, particularly advantageous when the likelihood function is difficult to derive or when computational simplicity is a priority. Selecting the appropriate method requires a careful balance between accuracy, simplicity, and the resources available for analysis.

**Table 2.** Simulation outcomes for scheme I.

$n$		MPSE	RADE	ME	MLE	ADE	PCE	LSE	WLSE	CVME
20	Bias	0.38582 <sub>3</sub>	0.39488 <sub>5</sub>	0.35948 <sub>2</sub>	0.34783 <sub>1</sub>	0.42751 <sub>9</sub>	0.41768 <sub>8</sub>	0.40982 <sub>7</sub>	0.38894 <sub>4</sub>	0.39987 <sub>6</sub>
	MSE	0.28498 <sub>5</sub>	0.30487 <sub>6</sub>	0.25384 <sub>2</sub>	0.21049 <sub>1</sub>	0.32847 <sub>9</sub>	0.32264 <sub>8</sub>	0.31085 <sub>7</sub>	0.27648 <sub>3</sub>	0.28471 <sub>4</sub>
	Sum of Ranks	8 <sub>4</sub>	11 <sub>6</sub>	4 <sub>2</sub>	2 <sub>1</sub>	18 <sub>9</sub>	16 <sub>8</sub>	14 <sub>7</sub>	7 <sub>3</sub>	10 <sub>5</sub>
100	Bias	0.24375 <sub>3</sub>	0.26634 <sub>5</sub>	0.21353 <sub>2</sub>	0.20947 <sub>1</sub>	0.30875 <sub>9</sub>	0.28896 <sub>8</sub>	0.27958 <sub>7</sub>	0.26295 <sub>4</sub>	0.27085 <sub>6</sub>
	MSE	0.17749 <sub>6</sub>	0.19487 <sub>7</sub>	0.14375 <sub>1</sub>	0.15485 <sub>3</sub>	0.20487 <sub>9</sub>	0.19947 <sub>8</sub>	0.16485 <sub>5</sub>	0.15326 <sub>2</sub>	0.16183 <sub>4</sub>
	Sum of Ranks	9 <sub>4</sub>	12 <sub>6,5</sub>	3 <sub>1</sub>	4 <sub>2</sub>	18 <sub>8</sub>	16 <sub>9</sub>	12 <sub>6,5</sub>	6 <sub>3</sub>	10 <sub>5</sub>
250	Bias	0.10384 <sub>3</sub>	0.132757 <sub>7</sub>	0.08864 <sub>2</sub>	0.04274 <sub>1</sub>	0.15364 <sub>9</sub>	0.14489 <sub>8</sub>	0.13086 <sub>6</sub>	0.12885 <sub>4</sub>	0.12947 <sub>5</sub>
	MSE	0.05464 <sub>5</sub>	0.09948 <sub>7</sub>	0.00375 <sub>1</sub>	0.00948 <sub>2</sub>	0.11094 <sub>9</sub>	0.10745 <sub>8</sub>	0.08474 <sub>6</sub>	0.01834 <sub>4</sub>	0.01147 <sub>3</sub>
	Sum of Ranks	8 <sub>4</sub>	14 <sub>7</sub>	3 <sub>1,5</sub>	3 <sub>1,5</sub>	18 <sub>9</sub>	16 <sub>8</sub>	12 <sub>6</sub>	8 <sub>4</sub>	8 <sub>4</sub>
500	Bias	0.00568 <sub>1</sub>	0.018382 <sub>6</sub>	0.00928 <sub>3</sub>	0.00814 <sub>2</sub>	0.07447 <sub>9</sub>	0.06648 <sub>8</sub>	0.05376 <sub>7</sub>	0.01286 <sub>5</sub>	0.01208 <sub>4</sub>
	MSE	0.00083 <sub>3</sub>	0.00209 <sub>6</sub>	0.00028 <sub>2</sub>	0.00014 <sub>1</sub>	0.00847 <sub>8</sub>	0.00958 <sub>9</sub>	0.00229 <sub>7</sub>	0.00084 <sub>4</sub>	0.00099 <sub>5</sub>
	Sum of Ranks	4 <sub>2</sub>	12 <sub>6</sub>	5 <sub>3</sub>	3 <sub>1</sub>	17 <sub>8,5</sub>	17 <sub>8,5</sub>	14 <sub>7</sub>	9 <sub>4,5</sub>	9 <sub>4,5</sub>
750	Bias	0.00058 <sub>6</sub>	0.00026 <sub>4</sub>	0.00004 <sub>2</sub>	0.00001 <sub>1</sub>	0.00283 <sub>9</sub>	0.00098 <sub>8</sub>	0.00072 <sub>7</sub>	0.00018 <sub>3</sub>	0.00029 <sub>5</sub>
	MSE	0.00008 <sub>5</sub>	0.00049 <sub>6</sub>	0.00001 <sub>2</sub>	0.00000 <sub>1</sub>	0.00095 <sub>8</sub>	0.00109 <sub>9</sub>	0.00064 <sub>7</sub>	0.00006 <sub>3</sub>	0.00007 <sub>4</sub>
	Sum of Ranks	11 <sub>6</sub>	10 <sub>5</sub>	4 <sub>2</sub>	2 <sub>1</sub>	17 <sub>8,5</sub>	17 <sub>8,5</sub>	14 <sub>7</sub>	6 <sub>3</sub>	9 <sub>4</sub>

**Table 3.** Simulation outcomes for scheme II.

$n$	Est.	MPSE	RADE	ME	MLE	ADE	PCE	LSE	WLSE	CVME
20	Bias	0.89294 <sub>3</sub>	0.96487 <sub>6</sub>	0.81893 <sub>1</sub>	0.85737 <sub>2</sub>	0.97736 <sub>7</sub>	1.18367 <sub>9</sub>	0.99477 <sub>8</sub>	0.91736 <sub>5</sub>	0.91074 <sub>4</sub>
	MSE	0.51748 <sub>2</sub>	0.66531 <sub>6</sub>	0.48384 <sub>1</sub>	0.54762 <sub>3</sub>	0.68932 <sub>7</sub>	0.70947 <sub>8</sub>	0.80476 <sub>9</sub>	0.55087 <sub>4</sub>	0.56487 <sub>5</sub>
Sum of Ranks		5 <sub>2.5</sub>	12 <sub>6</sub>	2 <sub>1</sub>	5 <sub>2.5</sub>	14 <sub>7</sub>	17 <sub>8.5</sub>	17 <sub>8.5</sub>	9 <sub>4.5</sub>	9 <sub>4.5</sub>
100	Bias	0.50376 <sub>3</sub>	0.55382 <sub>4</sub>	0.46486 <sub>1</sub>	0.47597 <sub>2</sub>	0.56883 <sub>5</sub>	0.72414 <sub>9</sub>	0.68468 <sub>7</sub>	0.69756 <sub>8</sub>	0.58975 <sub>6</sub>
	MSE	0.24366 <sub>3</sub>	0.29048 <sub>5</sub>	0.21034 <sub>1</sub>	0.22894 <sub>2</sub>	0.28486 <sub>4</sub>	0.43786 <sub>9</sub>	0.40873 <sub>8</sub>	0.39749 <sub>7</sub>	0.31039 <sub>6</sub>
Sum of Ranks		6 <sub>3</sub>	9 <sub>4.5</sub>	2 <sub>1</sub>	4 <sub>2</sub>	9 <sub>4.5</sub>	18 <sub>9</sub>	15 <sub>7.5</sub>	15 <sub>7.5</sub>	12 <sub>6</sub>
250	Bias	0.29476 <sub>3</sub>	0.38681 <sub>7</sub>	0.26884 <sub>2</sub>	0.23374 <sub>1</sub>	0.34583 <sub>5</sub>	0.42811 <sub>9</sub>	0.38846 <sub>8</sub>	0.36486 <sub>6</sub>	0.30847 <sub>4</sub>
	MSE	0.12943 <sub>3</sub>	0.17103 <sub>6</sub>	0.10038 <sub>2</sub>	0.08365 <sub>1</sub>	0.16386 <sub>5</sub>	0.21038 <sub>9</sub>	0.19229 <sub>8</sub>	0.18837 <sub>7</sub>	0.13984 <sub>4</sub>
Sum of Ranks		6 <sub>3</sub>	13 <sub>6.5</sub>	4 <sub>2</sub>	2 <sub>1</sub>	10 <sub>5</sub>	18 <sub>9</sub>	16 <sub>8</sub>	13 <sub>6.5</sub>	8 <sub>4</sub>
500	Bias	0.13874 <sub>5</sub>	0.15386 <sub>8</sub>	0.08467 <sub>2</sub>	0.03846 <sub>1</sub>	0.12047 <sub>3</sub>	0.14977 <sub>7</sub>	0.14284 <sub>6</sub>	0.15841 <sub>9</sub>	0.12298 <sub>4</sub>
	MSE	0.00686 <sub>6</sub>	0.01993 <sub>9</sub>	0.00091 <sub>2</sub>	0.00023 <sub>1</sub>	0.00193 <sub>3</sub>	0.00738 <sub>7</sub>	0.00563 <sub>5</sub>	0.01826 <sub>8</sub>	0.00278 <sub>4</sub>
Sum of Ranks		9 <sub>5</sub>	17 <sub>8.5</sub>	4 <sub>2</sub>	2 <sub>1</sub>	6 <sub>3</sub>	14 <sub>7</sub>	11 <sub>6</sub>	17 <sub>8.5</sub>	8 <sub>4</sub>
750	Bias	0.00246 <sub>5</sub>	0.00841 <sub>6</sub>	0.00014 <sub>2</sub>	0.00008 <sub>1</sub>	0.00088 <sub>4</sub>	0.02985 <sub>8</sub>	0.03429 <sub>9</sub>	0.02846 <sub>7</sub>	0.00074 <sub>3</sub>
	MSE	0.00029 <sub>5</sub>	0.00058 <sub>6</sub>	0.00003 <sub>3</sub>	0.00000 <sub>1</sub>	0.00009 <sub>4</sub>	0.00098 <sub>9</sub>	0.00064 <sub>7</sub>	0.00076 <sub>8</sub>	0.00002 <sub>2</sub>
Sum of Ranks		10 <sub>5</sub>	12 <sub>6</sub>	5 <sub>2.5</sub>	2 <sub>1</sub>	8 <sub>4</sub>	17 <sub>9</sub>	16 <sub>8</sub>	15 <sub>7</sub>	5 <sub>2.5</sub>

**Table 4.** Simulation outcomes for scheme III.

$n$	Est.	MPSE	RADE	ME	MLE	ADE	PCE	LSE	WLSE	CVME
20	Bias	0.65092 <sub>8</sub>	0.67391 <sub>9</sub>	0.48764 <sub>1</sub>	0.50743 <sub>2</sub>	0.59579 <sub>6</sub>	0.62874 <sub>7</sub>	0.55762 <sub>5</sub>	0.53756 <sub>4</sub>	0.53731 <sub>3</sub>
	MSE	0.43754 <sub>8</sub>	0.42857 <sub>7</sub>	0.28275 <sub>1</sub>	0.30935 <sub>2</sub>	0.38576 <sub>6</sub>	0.43965 <sub>9</sub>	0.33096 <sub>4</sub>	0.34867 <sub>5</sub>	0.32956 <sub>3</sub>
Sum of Ranks		16 <sub>8</sub>	16 <sub>8</sub>	2 <sub>1</sub>	4 <sub>2</sub>	12 <sub>6</sub>	16 <sub>8</sub>	9 <sub>4.5</sub>	9 <sub>4.5</sub>	6 <sub>3</sub>
100	Bias	0.42648 <sub>9</sub>	0.40575 <sub>7</sub>	0.31045 <sub>1</sub>	0.35068 <sub>3</sub>	0.39591 <sub>6</sub>	0.41049 <sub>8</sub>	0.37098 <sub>5</sub>	0.31105 <sub>2</sub>	0.35092 <sub>4</sub>
	MSE	0.28846 <sub>9</sub>	0.27498 <sub>8</sub>	0.18466 <sub>2</sub>	0.23475 <sub>5</sub>	0.25486 <sub>6</sub>	0.27443 <sub>7</sub>	0.22068 <sub>3</sub>	0.16586 <sub>1</sub>	0.22461 <sub>4</sub>
Sum of Ranks		18 <sub>9</sub>	15 <sub>7.5</sub>	3 <sub>1.5</sub>	8 <sub>4</sub>	12 <sub>6</sub>	15 <sub>7.5</sub>	8 <sub>4</sub>	3 <sub>1.5</sub>	8 <sub>4</sub>
250	Bias	0.24960 <sub>9</sub>	0.23487 <sub>7</sub>	0.18465 <sub>1</sub>	0.20948 <sub>3</sub>	0.24048 <sub>8</sub>	0.23086 <sub>6</sub>	0.23058 <sub>5</sub>	0.23029 <sub>4</sub>	0.20085 <sub>2</sub>
	MSE	0.12968 <sub>7</sub>	0.11068 <sub>2</sub>	0.08563 <sub>1</sub>	0.12572 <sub>5</sub>	0.13546 <sub>9</sub>	0.12985 <sub>8</sub>	0.12576 <sub>6</sub>	0.12058 <sub>4</sub>	0.11673 <sub>3</sub>
Sum of Ranks		16 <sub>8</sub>	9 <sub>5</sub>	2 <sub>1</sub>	8 <sub>3.5</sub>	17 <sub>9</sub>	14 <sub>7</sub>	11 <sub>6</sub>	8 <sub>3.5</sub>	5 <sub>2</sub>
500	Bias	0.10857 <sub>8</sub>	0.06856 <sub>4</sub>	0.02865 <sub>1</sub>	0.05978 <sub>3</sub>	0.11048 <sub>9</sub>	0.09586 <sub>6</sub>	0.10295 <sub>7</sub>	0.08485 <sub>5</sub>	0.05548 <sub>2</sub>
	MSE	0.00857 <sub>8</sub>	0.00028 <sub>1</sub>	0.00085 <sub>5</sub>	0.00043 <sub>3</sub>	0.00958 <sub>9</sub>	0.00205 <sub>6</sub>	0.00745 <sub>7</sub>	0.00030 <sub>2</sub>	0.00078 <sub>4</sub>
Sum of Ranks		16 <sub>8</sub>	5 <sub>1</sub>	6 <sub>3</sub>	6 <sub>3</sub>	18 <sub>9</sub>	12 <sub>6</sub>	14 <sub>7</sub>	7 <sub>5</sub>	6 <sub>3</sub>
750	Bias	0.00385 <sub>8</sub>	0.00058 <sub>4</sub>	0.00004 <sub>1</sub>	0.00008 <sub>2</sub>	0.00492 <sub>9</sub>	0.00374 <sub>7</sub>	0.00109 <sub>6</sub>	0.00064 <sub>5</sub>	0.00023 <sub>3</sub>
	MSE	0.00009 <sub>8</sub>	0.00008 <sub>6</sub>	0.00000 <sub>1</sub>	0.00000 <sub>2</sub>	0.00009 <sub>9</sub>	0.00008 <sub>7</sub>	0.00007 <sub>4</sub>	0.00008 <sub>5</sub>	0.00006 <sub>3</sub>
Sum of Ranks		16 <sub>7</sub>	10 <sub>5</sub>	2 <sub>1</sub>	4 <sub>2</sub>	18 <sub>9</sub>	14 <sub>8</sub>	10 <sub>5</sub>	10 <sub>5</sub>	6 <sub>3</sub>

**Table 5.** Simulation outcomes for scheme IV.

$n$	Est.	MPSE	RADE	ME	MLE	ADE	PCE	LSE	WLSE	CVME
50	Bias	1.25755 <sub>9</sub>	1.11048 <sub>8</sub>	0.89564 <sub>1</sub>	0.93756 <sub>3</sub>	1.08575 <sub>6</sub>	1.10498 <sub>5</sub>	1.10948 <sub>7</sub>	0.93325 <sub>2</sub>	0.98566 <sub>4</sub>
	MSE	0.54736 <sub>9</sub>	0.52846 <sub>8</sub>	0.41735 <sub>1</sub>	0.45638 <sub>2</sub>	0.47209 <sub>5</sub>	0.47974 <sub>6</sub>	0.46737 <sub>4</sub>	0.45975 <sub>3</sub>	0.50884 <sub>7</sub>
Sum of Ranks		18 <sub>9</sub>	16 <sub>8</sub>	2 <sub>1</sub>	5 <sub>2.5</sub>	11 <sub>5.5</sub>	11 <sub>5.5</sub>	11 <sub>5.5</sub>	5 <sub>2.5</sub>	11 <sub>5.5</sub>
200	Bias	0.74975 <sub>9</sub>	0.68597 <sub>5</sub>	0.51748 <sub>1</sub>	0.57795 <sub>3</sub>	0.70058 <sub>6</sub>	0.72058 <sub>8</sub>	0.71885 <sub>7</sub>	0.55069 <sub>2</sub>	0.67597 <sub>4</sub>
	MSE	0.35538 <sub>9</sub>	0.32694 <sub>4</sub>	0.26638 <sub>2</sub>	0.29407 <sub>3</sub>	0.32098 <sub>4</sub>	0.34373 <sub>8</sub>	0.33876 <sub>7</sub>	0.25487 <sub>1</sub>	0.32873 <sub>6</sub>
Sum of Ranks		18 <sub>9</sub>	9 <sub>4</sub>	3 <sub>1.5</sub>	6 <sub>3</sub>	10 <sub>5.5</sub>	16 <sub>8</sub>	14 <sub>7</sub>	3 <sub>1.5</sub>	10 <sub>5.5</sub>
400	Bias	0.43306 <sub>9</sub>	0.40476 <sub>7</sub>	0.29876 <sub>1</sub>	0.32058 <sub>2</sub>	0.41058 <sub>8</sub>	0.39859 <sub>6</sub>	0.38659 <sub>5</sub>	0.34058 <sub>3</sub>	0.38579 <sub>4</sub>
	MSE	0.23846 <sub>9</sub>	0.21946 <sub>7</sub>	0.14483 <sub>1</sub>	0.16578 <sub>2</sub>	0.22434 <sub>8</sub>	0.20538 <sub>5</sub>	0.20873 <sub>6</sub>	0.18365 <sub>3</sub>	0.19836 <sub>4</sub>
Sum of Ranks		18 <sub>9</sub>	14 <sub>7</sub>	2 <sub>1</sub>	4 <sub>2</sub>	16 <sub>8</sub>	11 <sub>5.5</sub>	11 <sub>5.5</sub>	6 <sub>3</sub>	8 <sub>4</sub>
500	Bias	0.28859 <sub>9</sub>	0.24437 <sub>8</sub>	0.10387 <sub>1</sub>	0.14096 <sub>2</sub>	0.23059 <sub>6</sub>	0.23309 <sub>7</sub>	0.21095 <sub>5</sub>	0.16584 <sub>3</sub>	0.19946 <sub>4</sub>
	MSE	0.14826 <sub>9</sub>	0.12098 <sub>8</sub>	0.01537 <sub>1</sub>	0.09873 <sub>3</sub>	0.11846 <sub>7</sub>	0.11746 <sub>6</sub>	0.10538 <sub>5</sub>	0.10048 <sub>4</sub>	0.08463 <sub>2</sub>
Sum of Ranks		18 <sub>9</sub>	16 <sub>8</sub>	2 <sub>1</sub>	5 <sub>2</sub>	13 <sub>6.5</sub>	13 <sub>6.5</sub>	10 <sub>5</sub>	7 <sub>4</sub>	6 <sub>3</sub>
750	Bias	0.08756 <sub>9</sub>	0.01876 <sub>6</sub>	0.00046 <sub>1</sub>	0.00395 <sub>3</sub>	0.01958 <sub>7</sub>	0.02968 <sub>8</sub>	0.00869 <sub>4</sub>	0.00096 <sub>2</sub>	0.00881 <sub>5</sub>
	MSE	0.00875 <sub>9</sub>	0.00134 <sub>8</sub>	0.00002 <sub>1</sub>	0.00009 <sub>3</sub>	0.00084 <sub>7</sub>	0.00074 <sub>5</sub>	0.00083 <sub>6</sub>	0.00008 <sub>2</sub>	0.00017 <sub>4</sub>
Sum of Ranks		18 <sub>9</sub>	14 <sub>7.5</sub>	2 <sub>1</sub>	6 <sub>3</sub>	14 <sub>7.5</sub>	13 <sub>6</sub>	10 <sub>5</sub>	4 <sub>2</sub>	9 <sub>4</sub>

**Table 6.** Ranking of estimation methods based on simulation results.

	<i>n</i>	MPSE	RADE	ME	MLE	ADE	PCE	LSE	WLSE	CVME
Schema I	20	4	6	2	1	9	8	7	3	5
	100	4	6.5	1	2	8	9	6.5	3	5
	250	4	7	1.5	1.5	9	8	4	4	4
	500	2	6	3	1	8.5	8.5	4.5	4.5	4.5
	750	6	5	2	1	8.5	8.5	3	3	4
Schema II	20	2.5	6	1	2.5	7	8.5	8.5	4.5	4.5
	100	3	4.5	1	2	4.5	9	7.5	7.5	6
	250	3	6.5	2	1	5	9	8	6.5	4
	500	5	8.5	2	1	3	7	6	8.5	4
	750	5	6	2.5	1	4	9	8	7	2.5
Schema III	20	8	8	1	2	6	8	4.5	4.5	3
	100	9	7.5	1.5	4	6	7.5	4	1.5	4
	250	8	5	1	3.5	9	7	6	3.5	2
	500	8	1	3	3	9	6	7	5	3
	750	7	5	1	2	9	8	5	5	3
Schema IV	20	9	8	1	2.5	5.5	5.5	5.5	2.5	5.5
	100	9	4	1.5	3	5.5	8	7	1.5	5.5
	250	9	7	1	2	8	5.5	5.5	3	4
	500	9	8	1	2	6.5	6.5	5	4	3
	750	9	7.5	1	3	7.5	6	5	2	4
Sum of Ranks		123.5	123	31	41	138.5	152.5	117.5	84	80.5
Overall Rank		7	6	1	2	8	9	5	4	3

## 6. Data analysis

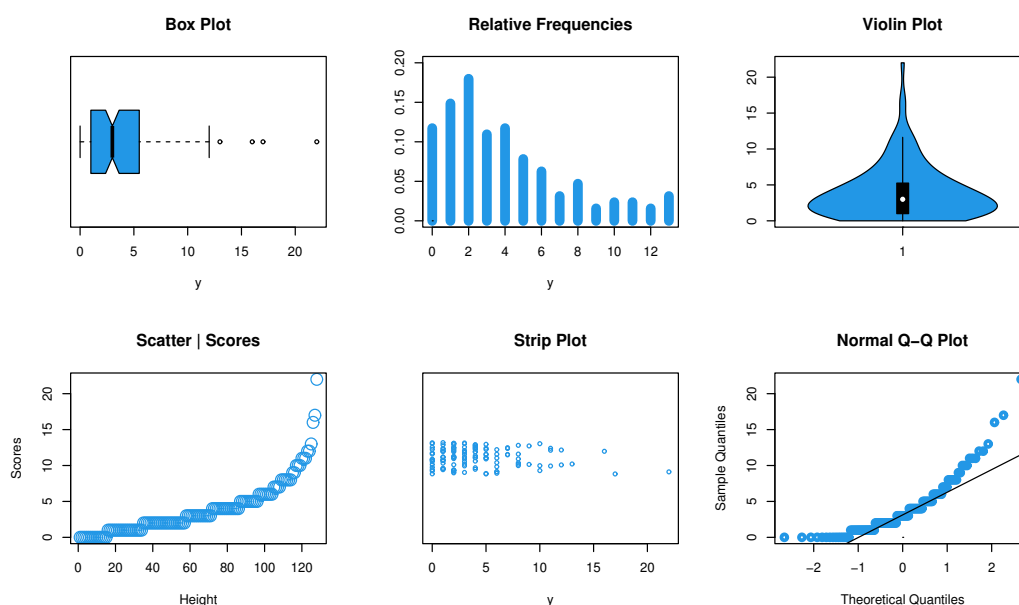
In this section, we showcase the versatility of the HDWC distribution in accurately modeling a variety of datasets from different domains. We assess the fitted distributions using the Chi-square ( $\chi^2$ ) test and its corresponding P-value in dataset I. In datasets II and III, the tested models are compared using the Kolmogorov-Smirnov (KS) test with its corresponding P-value and some criteria, namely, the maximized log-likelihood ( $l$ ), Akaike information criterion (AIC), corrected AIC (CAIC), Bayesian IC (BIC), and Hannan-Quinn IC (HQIC). Since there is a limited number of frequencies for each observation in datasets II and III, the Pearson's Chi-square statistic cannot be employed for an inference test. Therefore, the KS measure is adequate in this case. The codes can be listed in the following link: <https://docs.google.com/document/d/13XpIsgrCGryY9Ma1777NQ46fDpz32Ufc/edit>. We will compare the performance of the HDWC distribution against other competing distributions listed in Table 7.

**Table 7.** The competitive models of the HDWC distribution.

Distribution	Abbreviation	Author(s)
Discrete Inverted Nadarajah-Haghighi	DINH	Singh et al. [15]
Geometric	Geo	-
Negative Binomial (1 Parameter)	NBI	-
Discrete Rayleigh	DR	Roy [16]
Discrete Inverse Rayleigh	DIR	Hussain and Ahmad [17]
Poisson	Poi	Poisson [18]
Discrete Burr-Hatke	DBH	El-Morshedy et al. [19]
Discrete Pareto	DPa	Krishna and Pundir [20]

### 6.1. Dataset I

The dataset provided represents the count of computer breakdowns occurring over a period of 128 consecutive weeks of operation (Hand et al. [21]). Nonparametric plots are suitable for the initial visualization of this dataset, and the corresponding results can be shown in Figure 4. It is observed that there are some extreme values, displaying a right-skewed distribution. Table 8 presents the MLEs for the parameter(s),  $-l$ , and Chi-square test values, accompanied by their corresponding P-values, for all competing distributions in dataset I.

**Figure 4.** Nonparametric plots for dataset I.

**Table 8.** The MLEs,  $-l$ ,  $\chi^2$ , and P-value for all competing distributions in dataset I.

$X$	Obs. Fre.	HDWC	DINH	Geo	NBI	DR	DIR	Poi	DBH	DPa
0	15	10.339	12.401	25.527	25.520	3.636	3.237	2.310	65.853	47.806
1	19	20.099	30.161	20.434	20.432	10.302	47.809	9.272	21.915	19.190
2	23	20.893	19.935	16.360	16.358	15.306	34.021	18.613	10.931	10.760
3	14	18.288	12.782	13.096	13.097	18.041	16.649	24.912	6.538	7.021
4	15	14.798	8.721	10.491	10.486	18.440	8.773	25.014	4.342	5.001
5	10	11.467	6.291	8.385	8.395	16.918	5.079	20.085	3.088	3.774
6	8	8.657	4.732	6.720	6.721	14.172	3.174	13.440	2.304	2.967
7	4	6.426	3.693	5.381	5.381	10.944	2.106	7.713	1.782	2.404
8	6	4.717	2.950	4.312	4.308	7.839	1.466	3.870	1.417	1.994
9	2	3.436	2.412	3.456	3.449	5.228	1.059	1.736	1.151	1.686
10	3	2.490	2.013	2.753	2.762	3.256	0.789	0.690	0.953	1.447
11	3	1.798	1.691	2.205	2.211	1.897	0.604	0.253	0.800	1.258
12	2	1.295	1.460	1.766	1.770	1.036	0.472	0.080	0.680	1.106
+13	4	3.297	18.758	7.114	7.110	0.985	2.762	0.006	6.246	21.586
Total	128	128	128	128	128	128	128	128	128	128
$-l$		317.707	331.931	320.703	320.703	347.148	356.525	384.974	379.346	369.766
$MLE_\theta$		0.716	1.244	0.801	0.895	0.972	0.025	4.016	0.971	0.509
$MLE_q$		—	1.634	—	—	—	—	—	—	—
$\chi^2$		4.950	20.199	11.651	10.684	49.897	50.533	88.995	143.174	94.971
D.F		8	6	9	9	8	5	6	5	6
P-value		0.763	0.003	0.234	0.298	0	0	0	0	0

Of all the distributions tested, the HDWC distribution shows the best performance as the most suitable distribution for fitting dataset I. The estimated PMFs of all the distributions tested for dataset I are illustrated in Figure 5. From this figure, the same results are those derived from the Table 8.

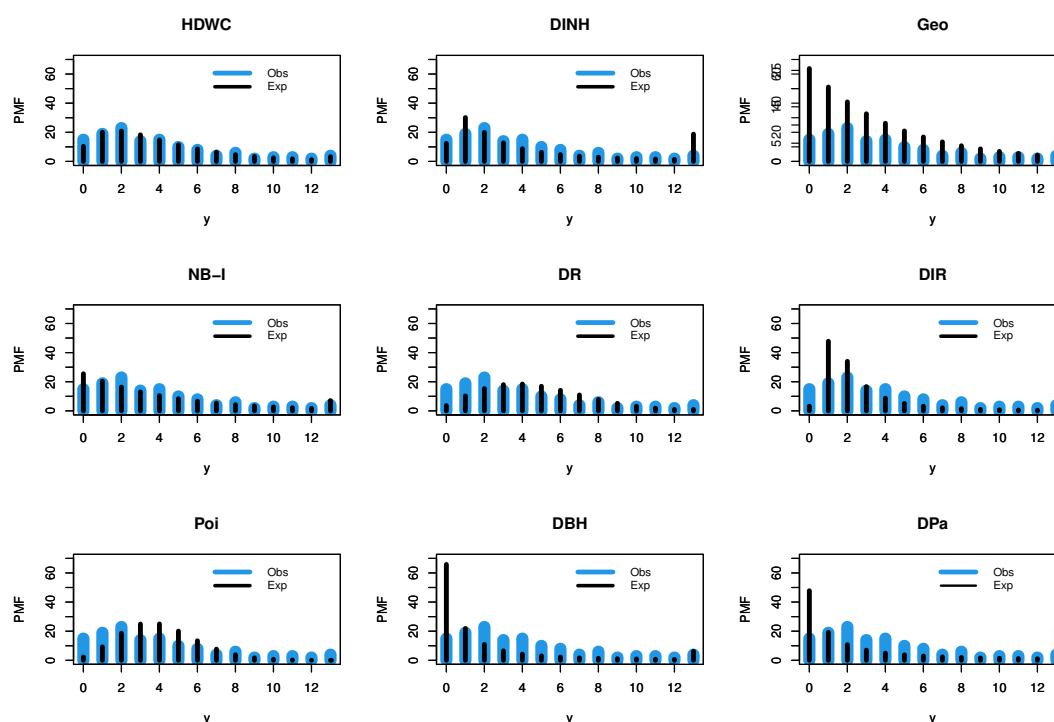
**Figure 5.** The estimated PMFs for dataset I.

Table 9 lists different estimators for dataset I, and it was noted that the MLE and ME are the best estimators for the dataset I. In addition, the CVME, LSE, and WLSE techniques work quite well for modeling this data.

**Table 9.** Different estimators for dataset I.

	MPSE	ADE	ME	MLE	LSE	RADE	PCE	CVME	WLSE
$\theta$	0.8615	0.8623	0.7173	0.7158	0.7386	0.8611	0.8660	0.7385	0.7387
$\chi^2$	147.0785	149.5152	4.9860	4.950	8.3453	145.6854	161.6198	8.3215	8.3522
D.F	11	11	8	8	9	10	11	9	9
P.value	0	0	0.7591	0.763	0.4999	0	0	0.5021	0.4991

Table 10 presents numerical summaries of the empirical descriptive statistics (EDS).

**Table 10.** Numerical summaries of the EDS for dataset I.

Approach	$E(X)$	$Var(X)$	$IOD(X)$	$Sk(X)$	$Ku(X)$
ME	4.0148	11.4015	2.8399	1.5941	36.1645
MLE	4.0281	11.4683	2.8471	1.5942	36.2005
LSE	4.4507	13.6949	3.0769	1.5965	37.2589
CVME	4.4485	13.6827	3.0758	1.5965	37.2537
WLSE	4.4529	13.7071	3.0782	1.5965	37.2640

The characteristics of dataset I clearly show a right-skewed distribution with leptokurtic features and overdispersion.

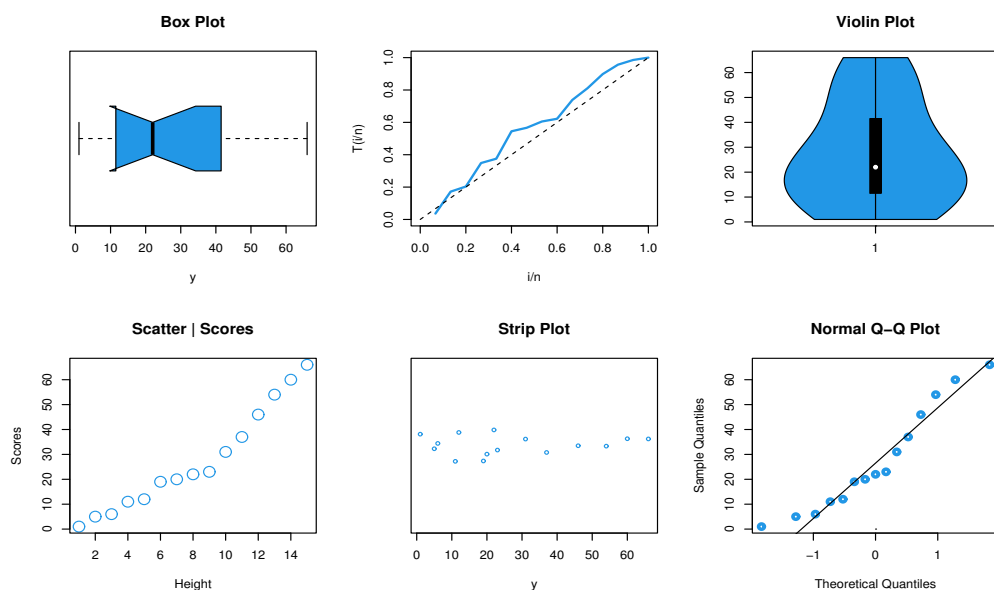
## 6.2. Dataset II

The second dataset comprises the failure times for a sample of 15 electronic components subjected to an accelerated life test (refer to Lawless [22]). Nonparametric plots for dataset II are illustrated in Figure 6. It is noted that there are no extreme values, yet the distribution remains right-skewed.

The MLEs and goodness of fit (GOF) measures for this data are reported in Table 11.

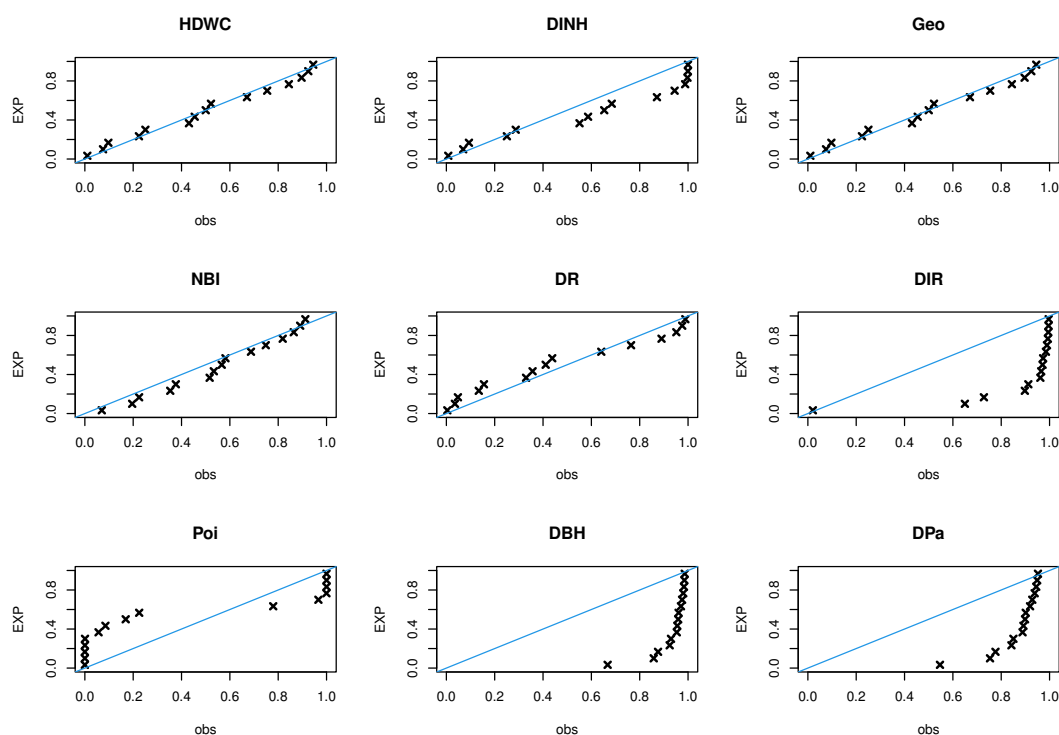
**Table 11.** The MLEs and GOF measures for dataset II.

Statistic	HDWC	DINH	Geo	NBI	DR	DIR	Poi	DBH	DPa
$MLE_{\theta}$	0.948	0.578	0.965	0.982	0.999	$1.8 \times 10^{-7}$	27.533	0.999	0.720
$MLE_q$	—	0.193	—	—	—	—	—	—	—
$-l$	64.770	67.879	65.002	65.000	66.394	89.096	151.206	91.368	77.402
AIC	131.540	139.758	132.002	132.000	134.788	180.192	304.413	184.737	156.805
CAIC	131.848	140.758	132.310	132.308	135.096	180.499	304.721	185.045	157.112
BIC	132.248	141.174	132.710	132.708	135.496	180.899	305.121	185.445	157.513
HQIC	131.533	139.743	131.994	131.992	134.781	180.184	304.405	184.729	156.797
KS	0.112	0.207	0.176	0.177	0.216	0.698	0.381	0.791	0.405
P-value	0.981	0.481	0.673	0.675	0.433	< 0.0001	0.025	< 0.0001	0.009



**Figure 6.** Nonparametric plots for dataset II.

As observed, at a significance level of 0.05, the HDWC, NBI, Geo, DINH, and DR models (for full names, refer to Table 7) perform effectively in modeling dataset II, with the HDWC distribution being the most flexible. Figure 7 displays the probability-probability (P-P) plots for dataset II, which support the empirical findings presented in Table 11.



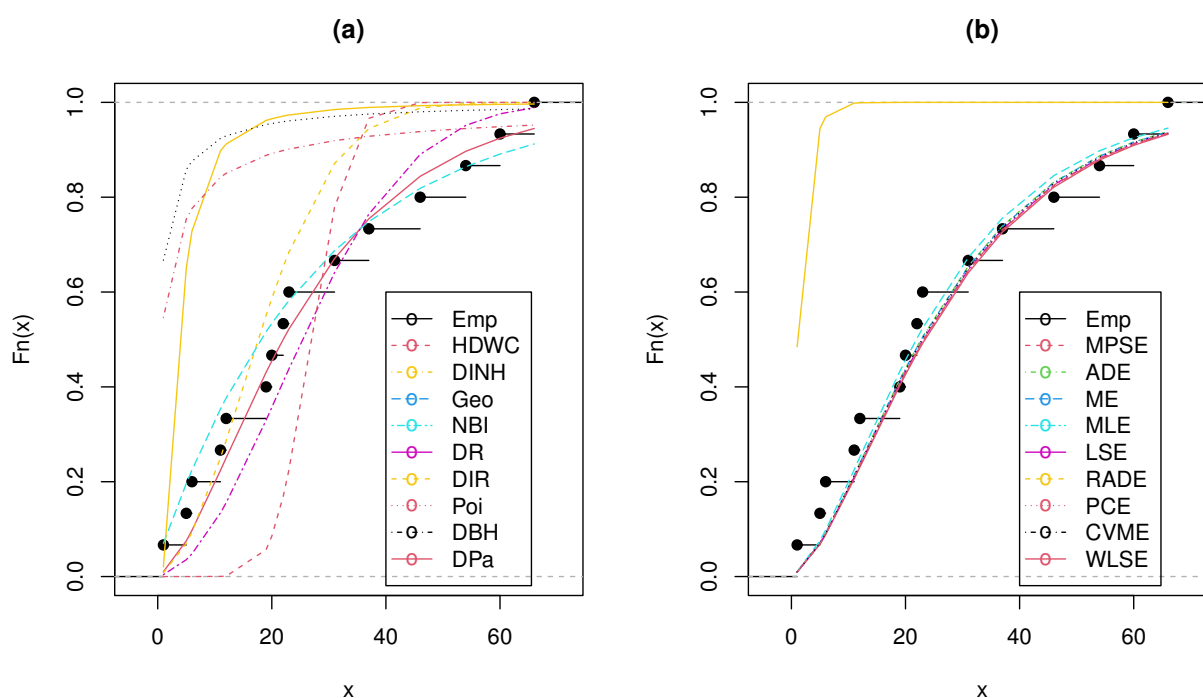
**Figure 7.** The P-P plots for dataset II.

Table 12 lists various estimators for dataset II, and it was noted that all techniques work quite well for modeling this data except the RADE.

**Table 12.** Different estimators for dataset II.

	MPSE	ADE	ME	MLE	LSE	RADE	PCE	CVME	WLSE
$\theta$	0.9509	0.9496	0.9479	0.9478	0.9503	0.5512	0.9504	0.9499	0.9506
KS	0.1119	0.1079	0.1120	0.1119	0.1073	0.8781	0.1102	0.1077	0.11087
P-Value	0.9810	0.9869	0.9810	0.9809	0.9875	0.0000	0.9838	0.9871	0.9827

The empirical CDFs plots for dataset II for all competing distributions are given in Figure 8 (left panel) and the empirical CDFs plots for dataset II for the HDWC using the different estimation methods are reported in Figure 8 (right panel) which prove the empirical results mentioned in Tables 11 and 12.



**Figure 8.** The empirical CDFs plots for dataset II.

Table 13 presents numerical summaries of the EDS.

Dataset II clearly exhibits a right-skewed distribution with leptokurtic characteristics and overdispersion.

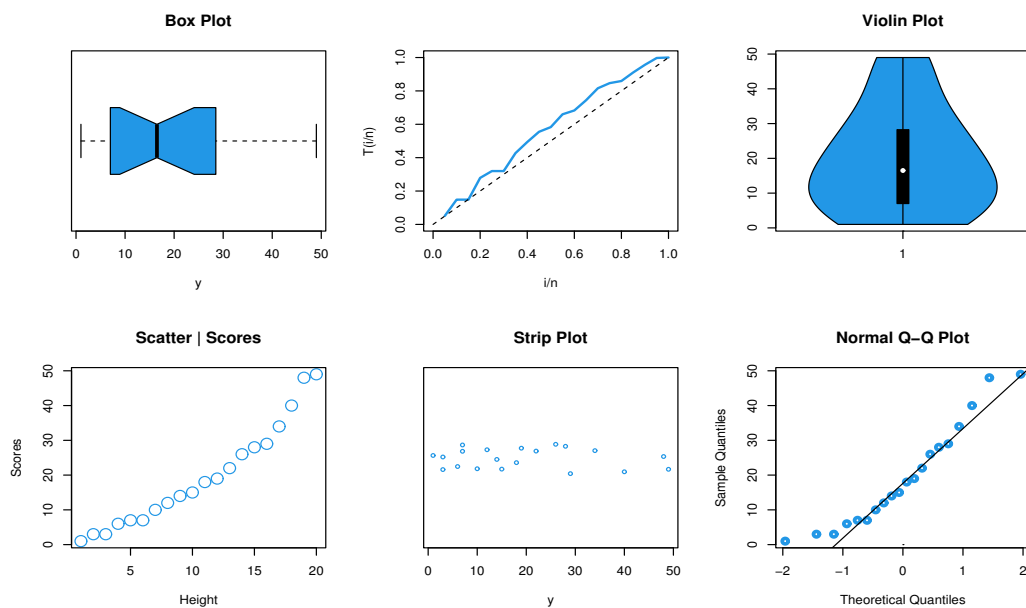


**Table 13.** The EDS for dataset II.

Approach	$E(X)$	$Var(X)$	$IOD(X)$	$Sk(X)$	$Ku(X)$
MPSE	29.2936	493.2271	16.8373	1.6096	47.9050
ADE	28.5054	467.4808	16.3997	1.6095	47.8417
ME	27.5341	436.7002	15.8603	1.6095	47.7587
MLE	27.5895	438.4270	15.8911	1.6095	47.7636
LSE	28.9247	481.0908	16.6325	1.6096	47.8757
PCE	28.9856	483.0828	16.6663	1.6096	47.8807
CVME	28.6837	473.2432	16.4987	1.6095	47.8562
WLSE	29.1080	487.1033	16.7343	1.6096	47.8904

### 6.3. Dataset III

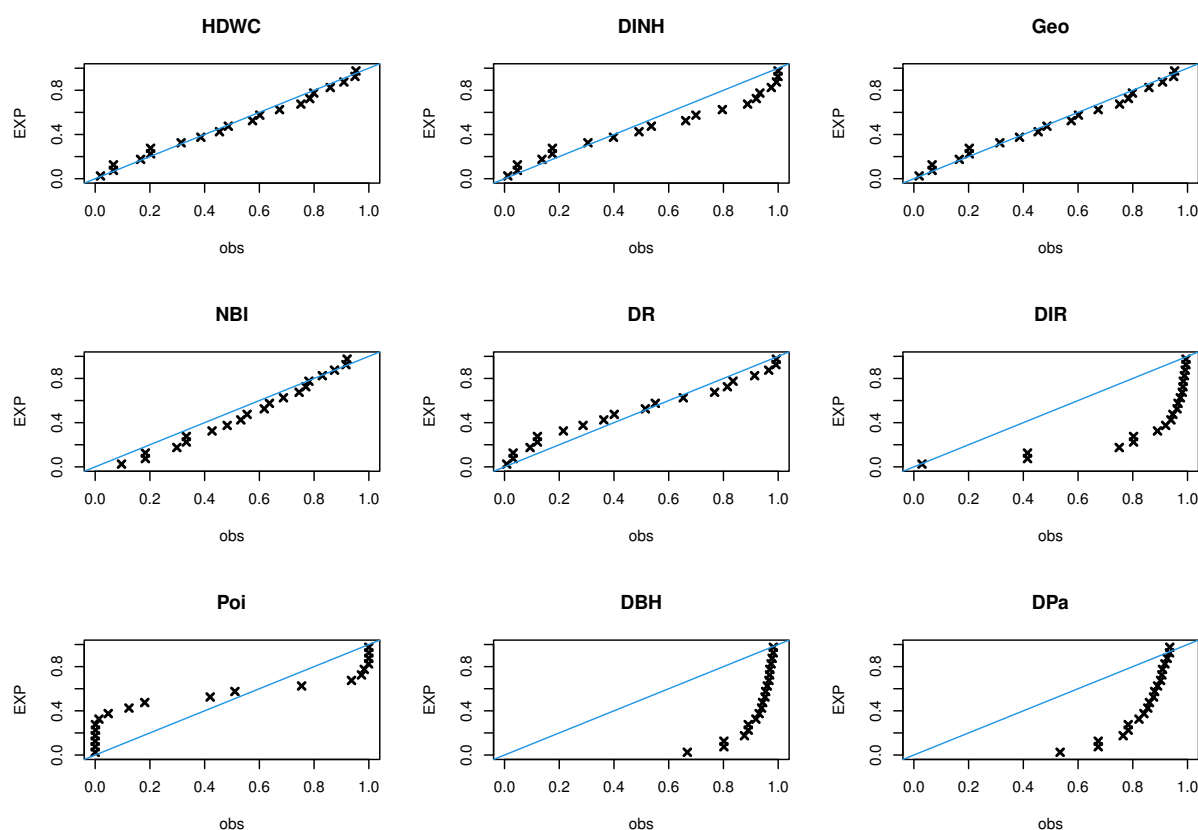
This dataset includes leukemia remission times (in weeks) for 20 patients, as discussed by Damien and Walker [23] in relation to the concept of discretization. Nonparametric plots for dataset III are displayed in Figure 9.

**Figure 9.** Nonparametric plots for dataset III.**Table 14.** The GOF test for dataset III.

Statistic	HDWC	DINH	Geo	NBI	DR	DIR	Poi	DBH	DPa
$MLE_{\theta}$	0.928	0.737	0.951	0.975	0.998	$7.82 \times 10^{-7}$	19.550	0.998	0.696
$MLE_q$	—	14.798	—	—	—	—	—	—	—
$-L$	79.053	82.818	79.962	79.963	81.175	101.987	152.718	110.283	95.448
AIC	160.107	169.635	161.925	161.926	164.351	205.975	307.436	222.565	192.896
CAIC	160.329	170.341	162.147	162.148	164.572	206.197	307.658	222.787	193.118
BIC	161.102	171.627	162.921	162.922	165.346	206.973	308.432	223.561	193.892
HQIC	160.301	170.024	162.119	162.120	164.544	206.169	307.630	222.759	193.090
KS	0.103	0.189	0.145	144	0.199	0.556	0.352	0.751	0.392
P-value	0.984	0.467	0.796	0.795	0.401	< 0.001	0.014	< 0.001	0.004

The MLEs and the GOF test and measures for this data are reported in Table 14.

As can be seen, based on significance level 0.05, all tested distributions work quite well for analyzing dataset III except DIR, DBH, DPa, and Poi, but the HDWC distribution is the best (For full names, refer to Table 7). Figure 10 presents the P-P plots for dataset III, reinforcing the empirical results shown in Table 14.



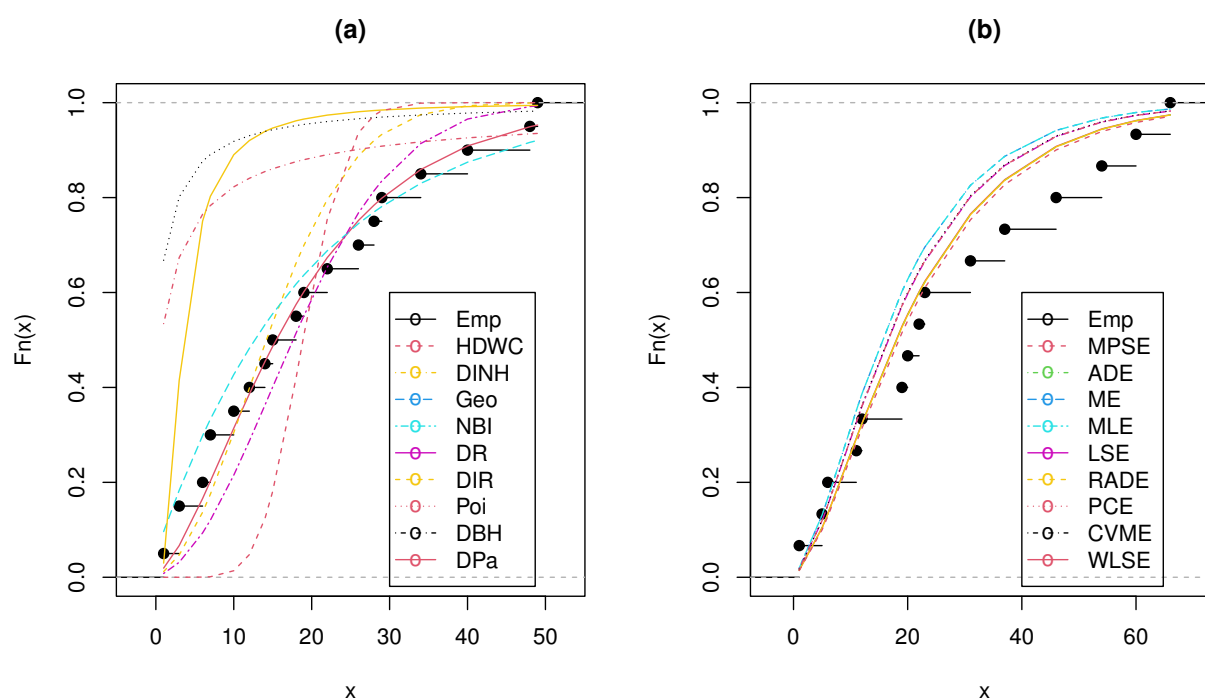
**Figure 10.** The P-P plots for dataset III.

Table 15 provides a list of various estimators for dataset III, noting that all methods perform flexibility in modeling this data.

**Table 15.** Different estimators for dataset III.

	MPSE	ADE	ME	MLE	LSE	RADE	PCE	CVME	WLSE
$\theta$	0.9371	0.9373	0.9279	0.9279	0.9319	0.9373	0.9387	0.9315	0.9318
KS	0.0763	0.0769	0.1026	0.1026	0.1137	0.0769	0.0828	0.1145	0.1138
P-Value	0.9996	0.9995	0.9843	0.9843	0.9530	0.9995	0.9986	0.9524	0.9529

Figure 11 (left panel) displays the empirical CDF plots for dataset III across all competing distributions, while Figure 11 (right panel) presents the empirical CDF plots for the HDWC distribution using various estimation methods. Together, these figures support the empirical findings discussed in Tables 14 and 15.



**Figure 11.** The empirical CDFs plots for dataset III.

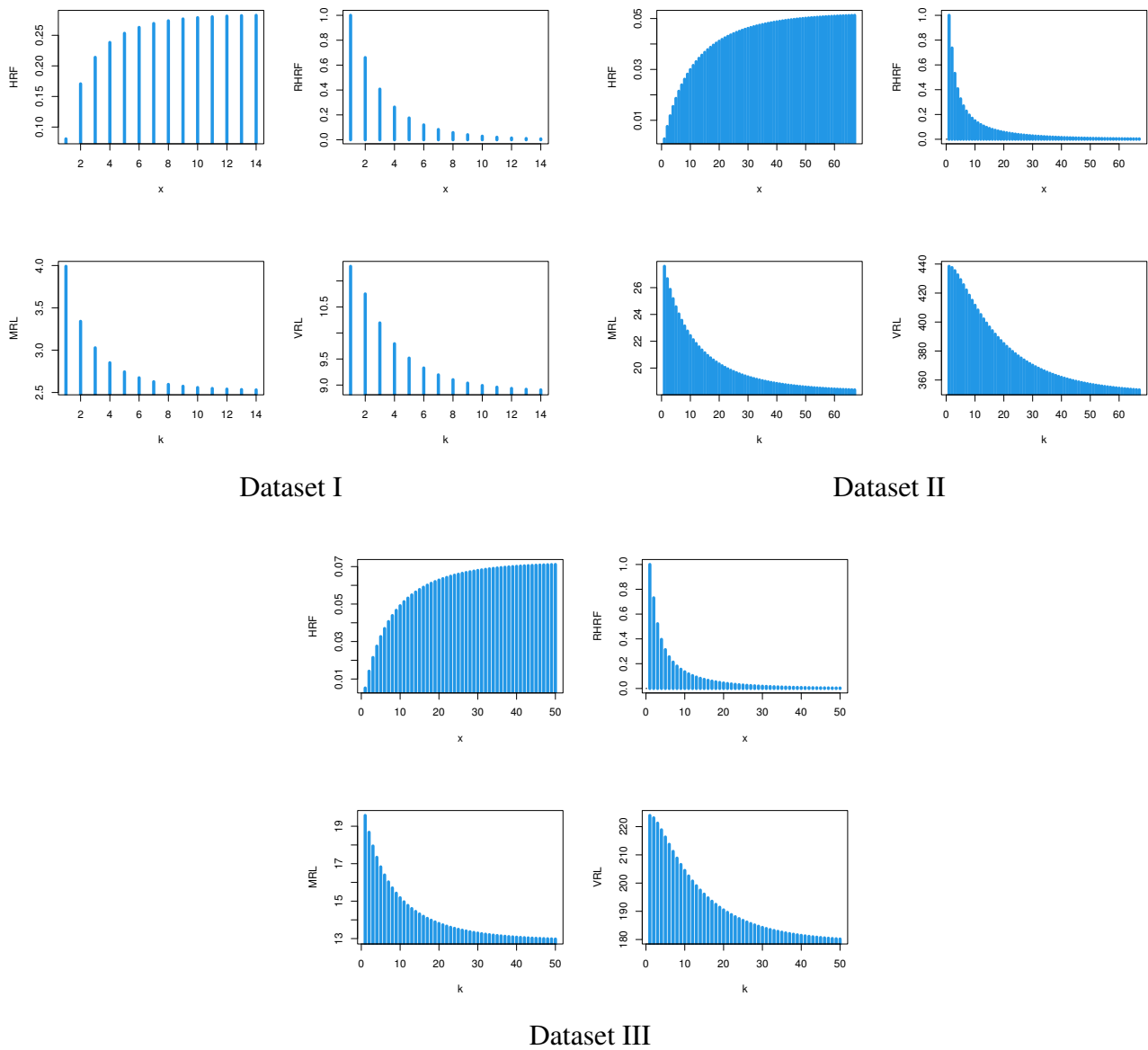
Table 16 provides numerical summaries of the EDS for dataset III.

Dataset III distinctly displays a right-skewed distribution with leptokurtic traits and overdispersion.

**Table 16.** The EDS for dataset III.

Approach	$E(X)$	$Var(X)$	$IOD(X)$	$Sk(X)$	$Ku(X)$
MPSE	22.5892	296.2575	13.1149	1.6093	47.2310
ADE	22.6653	298.2128	13.1572	1.6093	47.2409
ME	19.5450	223.3083	11.4252	1.6091	46.7802
MLE	19.5739	223.9526	11.4413	1.6091	46.7851
LSE	20.7676	251.3672	12.1038	1.6092	46.9764
RADE	22.6653	298.2128	13.1572	1.6093	47.2409
PCE	23.2119	312.4469	13.4606	1.6093	47.3094
CVME	20.6389	248.3358	12.0323	1.6092	46.9568
WLSE	20.7353	250.6043	12.0858	1.6092	46.9715

#### 6.4. Reliability analysis of systems and treatments: Insights from HRF, MRL, RHRF, and VRL trends



**Figure 12.** The reliability properties of datasets.

The analysis of the three datasets provides insights into their reliability characteristics from different perspectives. Dataset I, representing the count of computer breakdowns over 128 consecutive weeks, reveals an increasing HRF, indicating a growing likelihood of breakdowns as time progresses, likely due to aging or wear-and-tear effects. The MRL decreases as more breakdowns occur, suggesting shorter expected operational times before subsequent failures. Similarly, the RHRF and VRL also decrease, reinforcing the notion of declining system reliability over time and highlighting the importance of proactive maintenance strategies. Dataset II, involving failure times for 15 electronic components subjected to accelerated life tests, also shows an increasing HRF, suggesting higher failure probabilities as time advances under stress conditions. The decreasing MRL

highlights the reduced lifespan of components as failures occur, while the declining RHRF and VRL confirm the ongoing degradation. These observations emphasize the necessity of robust design, quality control, and testing to enhance component reliability in challenging environments. Dataset III, which contains leukemia remission times for 20 patients, demonstrates a similar reliability trend. The increasing HRF suggests that the risk of remission ending grows over time, potentially due to disease progression or diminishing treatment effectiveness. The decreasing MRL indicates shorter expected remission durations, while the declining RHRF and VRL highlight the reduced stability of remission periods. These findings underscore the importance of tailored treatment plans and continuous monitoring to improve patient outcomes. Figure 12 visually summarizes these findings, demonstrating consistent patterns across the datasets, all of which reflect progressively worsening reliability over time, calling for targeted strategies like preventive maintenance, rigorous testing, and personalized interventions to enhance performance and outcomes.

The reliability analysis of these datasets provides valuable insights into the progression of failures and the potential for improving performance and resilience. By understanding these patterns, researchers and practitioners can develop effective approaches tailored to the unique challenges presented by engineering systems, electronic components, and medical treatment plans.

## 7. Conclusions and future directions

The article introduced a discrete distribution with a single parameter, developed using a weighted combining discretization approach, referred to as the HDWC distribution. This model offers a comprehensive set of statistical properties, all derived and expressed in closed forms. It has proven to be particularly effective for modeling datasets that exhibit right-skewness and leptokurtic shapes, making it well-suited for real-world data with heavy tails and extreme values. The HDWC distribution is particularly valuable in situations where the hazard rate function increases, especially when dealing with outliers, providing a more robust framework compared to traditional models. The article explored multiple estimation methods for the HDWC model, including MPSE, ME, ADE, MLE, LSE, RADE, PCE, CVME, and WLSE, and found that all methods were successful in estimating the model's parameter. Among these, the ME approach demonstrated the best performance. To further validate the effectiveness of the HDWC distribution, three real-world datasets were analyzed, showing that the HDWC distribution outperformed other competing distributions, such as the geometric, Poisson, and negative binomial models, in all aspects of the analysis. The article also outlined several promising directions for future research. These include extending the model to handle time-dependent stochastic processes for applications such as survival analysis and event prediction, adapting the model for spatially correlated data in fields like epidemiology and environmental studies, and incorporating Bayesian methods to improve parameter estimation, especially for small datasets. Additionally, the potential for combining the HDWC model with machine learning techniques or exploring hybrid models could further enhance its flexibility and predictive power, opening up new avenues for its application in complex data modeling.

## Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

## Acknowledgments

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2024/01/31935).

## Data Availability

The data that supports the findings of this study are available within the article.

## Conflict of interest

There are no conflicts of interest.

## References

1. T. Nakagawa, S. Osaki, The discrete Weibull distribution, *IEEE Trans. Reliab.*, **24** (1975), 300–301. <https://doi.org/10.1109/TR.1975.5214915>
2. K. B. Kulasekera, D. W. Tonkyn, A new discrete distribution, with applications to survival, dispersal and dispersion, *Commun. Stat.- Simul. Comput.*, **21** (1992), 499–518. <https://doi.org/10.1080/03610919208813032>
3. H. Sato, M. Ikota, A. Sugimoto, H. Masuda, A new defect distribution metrology with a consistent discrete exponential formula and its applications, *IEEE Trans. Semicond. Manuf.*, **12** (1999), 409–418. <https://doi.org/10.1109/66.806118>
4. J. D. Smith, A review of Finn, Fischer, and Handler (Eds.), collaborative/therapeutic assessment: A casebook and guide, *JPA*, **95** (2012), 234–235. <https://doi.org/10.1080/00223891.2012.730086>
5. M. Roederer, A. Treister, W. Moore, L. A. Herzenberg, Probability binning comparison: A metric for quantitating univariate distribution differences, *Cytometry*, **45** (2001), 37–46. [https://doi.org/10.1002/1097-0320\(20010901\)45:1<37::AID-CYTO1142>3.0.CO;2-E](https://doi.org/10.1002/1097-0320(20010901)45:1<37::AID-CYTO1142>3.0.CO;2-E)
6. A. Barbiero, A. Hitaj, Discrete approximations of continuous probability distributions obtained by minimizing Cramer-von Mises-type distances, *Stat. Papers*, **64** (2023), 1669–1697. <https://doi.org/10.1007/s00362-022-01356-2>
7. T. Ghosh, D. Roy, N. K. Chandra, Reliability approximation through the discretization of random variables using reversed hazard rate function, *Int. J. Math. Comput. Stat. Nat. Phys. Eng.*, **7** (2013), 96–100.
8. S. Chakraborty, Generating discrete analogues of continuous probability distributions-A survey of methods and constructions, *J. Stat. Distrib. Appl.*, **2** (2015), 6. <https://doi.org/10.1186/s40488-015-0028-6>
9. S. Kotsiantis, D. Kanellopoulos, Discretization techniques: A recent survey, *GESTS Int. Trans. Comput. Sci. Eng.*, **32** (2006), 47–58.
10. G. Casella, R. L. Berger, *Statistical Inference Vol. 70*, Duxbury Press, 1990. Available from: <https://philpapers.org/rec/CASSIV>.

11. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006. Available from: <https://link.springer.com/book/9780387310732#bibliographic-information>.
12. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin, *Bayesian Data Analysis*, CRC Press, 2013. <https://doi.org/10.1201/b16018>
13. R. E. Barlow, Statistical theory of reliability and life testing, 1975. <https://cir.nii.ac.jp/crid/1571980074720917504>
14. C. D. Lai, M. Xie, *Stochastic Ageing and Dependence for Reliability*, Springer, 2006. <https://doi.org/10.1007/0-387-34232-X>
15. B. Singh, R. P. Singh, A. S. Nayal, A. Tyagi, Discrete inverted Nadarajah-Haghighi distribution: Properties and classical estimation with application to complete and censored data, *Stat. Optim. Inf. Comput.*, **10** (2022), 1293–1313. <https://doi.org/10.19139/soic-2310-5070-1365>
16. D. Roy, Discrete rayleigh distribution, *IEEE Trans. Reliab.*, **53** (2004), 255–260. <https://doi.org/10.1109/TR.2004.829161>
17. T. Hussain, M. Ahmad, Discrete inverse Rayleigh distribution, *Pak. J. Stat.*, **30** (2014).
18. S. D. Poisson, *Probabilité des Jugements en Matière Criminelle et en Matière Civile, Précédées des Règles Générales du Calcul des Probabilités*, Paris, France: Bachelier, 1837.
19. M. El-Morshedy, M. S. Eliwa, E. Altun, Discrete Burr-Hatke distribution with properties, estimation methods and regression model, *IEEE Access*, **8** (2020), 74359–74370. <https://doi.org/10.1109/ACCESS.2020.2988431>
20. H. Krishna, P. S. Pundir, Discrete Burr and discrete Pareto distributions, *Stat. Methodol.*, **6** (2009), 177–188. <https://doi.org/10.1016/j.stamet.2008.07.001>
21. D. J. Hand, F. Daly, K. J. McConway, A. D. Lunn, E. O. Ostrowski, *A Hand Book of Small Data Sets*, Chapman and Hall/CRC, 1993. <https://doi.org/10.1201/9780429246579>
22. J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, 2011.
23. P. Damien, S. Walker, A Bayesian non-parametric comparison of two treatments, *Scand. J. Stat.*, **29** (2002), 51–56. <https://doi.org/10.1111/1467-9469.00891>



AIMS Press

© 2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)