STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, July 2026, pp 49–84.
Published online in International Academic Press (www.IAPress.org)

Discrete Odd Weibull-Lindley Distribution Featuring Adaptable Hazard Rate and Dispersion Index: Theoretical Framework and Applications to SDG-Aligned Count Data

Hend S. Shahen [1], Mahmoud El-Morshedy [1], Mohamed S. Eliwa [2,3,*]

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

Abstract This paper introduces the discrete odd Weibull–Lindley (DOWL) distribution, a flexible three-parameter model for analyzing complex count data frequently encountered in sustainable development research. The distribution is constructed by discretizing the continuous Lindley distribution within the discrete odd Weibull-G transformation framework. Its fundamental statistical and reliability properties are derived, including the identifiability of the mode and parameter effects, quantile function, hazard and reversed hazard rate functions, moments, information measures, residual and past lifetime functions, and equilibrium distribution. The DOWL distribution exhibits substantial flexibility, accommodating symmetric and asymmetric shapes, various kurtosis levels, and different dispersion patterns (over-, under-, and equi-dispersion). Its hazard rate function supports multiple aging behaviors, such as increasing, decreasing, bathtub, and unimodal shapes, making it suitable for diverse reliability and survival contexts. Parameter estimation is conducted via maximum likelihood, and extensive Monte Carlo simulations demonstrate satisfactory finite-sample performance of the estimators. The practical utility of the model is illustrated through three real data applications related to the United Nations Sustainable Development Goals: kidney cyst counts (SDG 3), European red mite infestation counts (SDGs 2 and 15), and daily COVID-19 mortality counts in Greece (SDG 3). Comparative analyses show that the DOWL distribution provides improved fit over competing discrete models, particularly for over-dispersed and highly skewed count data.

Keywords Discrete odd Weibull-G class; Index of dispersion; Failure analysis; Reliability classes; Monte Carlo simulation; SDG-aligned applications; Goodness-of-fit.

MSC2020: 60E05; 62E10; 62F10; 62N05; 62P12; 62P10.

DOI: 10.19139/soic-2310-5070-3717

1. Introduction

In a time characterized by intricate and interrelated global issues, the importance of robust statistical modeling in promoting sustainable development has grown significantly. The implementation of the United Nations Sustainable Development Goals (SDGs) in 2015 created an extensive framework consisting of 17 goals and 169 targets that tackle essential concerns related to environmental management, social justice, and economic growth [30]. SDG 3 (Good Health and Well-being), SDG 11 (Sustainable Cities and Communities), SDG 13 (Climate Action), and SDG 15 (Life on Land) are areas where quantitative analysis of specific phenomena is crucial for tracking progress and guiding policy decisions. The successful attainment of these objectives largely relies on the availability of advanced analytical tools capable of deriving significant insights from intricate, often flawed, data structures.

*Correspondence to: Mohamed S. Eliwa (Email: mseliwa@mans.edu.eg).

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2026 International Academic Press

The assessment and surveillance of SDG indicators may entail discrete count statistics that show the frequency of certain events within established spatial or temporal parameters. Instances are prevalent across nearly all sustainability sectors: the occurrence of extreme weather events is a vital metric for evaluating climate change in accordance with SDG 13; the rates of infectious diseases and mortality statistics are essential indicators for public health monitoring under SDG 3; the frequency of traffic accidents in urban areas guides transportation safety regulations related to SDG 11; and pest infestation data are crucial for assessing agricultural sustainability under SDG 2 (Zero Hunger) and ecosystem vitality under SDG 15. The precise modeling of count data significantly impacts evidence-based decision-making and risk assessment, making the creation of adaptable and manageable discrete probability distributions a matter of substantial practical significance for researchers and policymakers aiming to promote sustainable development [28]. The intrinsic features of sustainability-related count data pose significant modeling difficulties that traditional discrete distributions frequently do not sufficiently accommodate. Such data often demonstrate over-dispersion, under-dispersion, zero-inflation (an abundance of zeros compared to conventional models), heavy tails, and intricate hazard rate patterns. The conventional Poisson distribution, although fundamental for analyzing count data, presupposes equi-dispersion and is therefore insufficient when data diverge from this limitation. The negative binomial distribution addresses over-dispersion but is incapable of managing under-dispersed data. These constraints highlight the essential requirement for more adaptable discrete distributions that may concurrently address various deviations from classical assumptions, hence directly motivating the current research.

The construction of versatile and tractable discrete distributions is a cornerstone of modern statistical research, driven by the fact that many empirical phenomena are naturally observed as discrete counts rather than continuous variables. A prominent and mathematically elegant technique for this purpose is the survival discretization method, which preserves the functional form of a continuous model's survival function within its discrete analogue [26]. This approach ensures that the resulting distribution inherits the advantageous properties of its continuous progenitor while remaining highly interpretable for count-based data. Consequently, this methodology has facilitated the development of robust discrete models widely applied across diverse scientific disciplines. Specifically, discretizations of the Lindley distribution have produced highly flexible models, such as the discrete Lindley distribution introduced by [16] and the recent discrete expansion proposed by [14]. Building on these methodological foundations, the Rayleigh distribution was adapted for discrete contexts by [26], while the Burr and Pareto families have yielded discrete versions specifically tailored for actuarial science and reliability engineering [21]. Further advancements in the literature include the development of the discrete inverse Weibull [20], the discrete log-logistic [24], and the discrete Burr-Hatke distributions [5]. More recently, specialized models have emerged, such as the discrete extended odd Weibull exponential model [23] and the discrete Poisson-Lindley distribution [2]. Modern structural frameworks have also been introduced, including the discrete exponentiated generalized family [1] and the discrete odd Nadarajah-Haghighi-G class [17].

Recently, a substantial amount of statistical research has concentrated on creating discrete distributions with increased flexibility to address the varied data properties found in practical applications. This study path has yielded several significant contributions. The discrete generalized Lindley distribution proved particularly useful for modeling COVID-19 pandemic data, offering a flexible framework for epidemiological study during global health emergencies [6]. The discrete Burr-Hatke distribution, along with its corresponding regression model, broadened the use of discrete distributions in explanatory modeling contexts [5]. The exponentiated discrete Lindley distribution demonstrated a notable diversity of hazard rate forms, encompassing growing, decreasing, bathtub, and unimodal patterns, thus broadening its applicability to intricate reliability problems [10]. The adaptability of parsimonious one-parameter discrete models has been examined, revealing their significant effectiveness in modeling over-dispersed data while preserving interpretability and computational efficiency [13]. Advanced distributional frameworks have enhanced the resources accessible to data scientists dealing with count data. The Kumaraswamy discrete half-logistic distribution [27], the discrete linear-exponential distribution [8], and the discrete single-factor extension of the exponential distribution [9] exemplify the trend towards more adaptable discrete models. These advancements indicate a wider acknowledgment that the intricacy of real-world events necessitates statistical models that can adjust to various data-generating processes while maintaining interpretability and computing efficiency.

A notably versatile and mathematically refined family of discrete distributions, termed the discrete odd Weibull-G (DOW-G) family, was recently presented by [7]. This family establishes a systematic framework for creating new discrete distributions by integrating the odd Weibull transformation with survival discretization, thus providing improved flexibility in modeling intricate count data patterns. The mathematical representation of this family is defined by the subsequent formulas. The cumulative distribution function (CDF) of the DOW-G family is expressed as

$$V\left(y;\Phi\right)=1-p^{\left(Q(y+1;\Phi)\right)^{\alpha}},\quad y\in\mathbb{N}_0, \tag{1}$$

where the transformation function $Q\left(y;\Phi\right)$ is defined as the odds function of the baseline distribution:

$$Q\left(y;\Phi\right)=\frac{G\left(y;\Phi\right)}{1-G\left(y;\Phi\right)}=\frac{G\left(y;\Phi\right)}{\overline{G}\left(y;\Phi\right)}. \tag{2}$$

The reliability function (RF), denotes the chance that the random variable surpasses a specified value and is expressed as

$$P(Y>y)=p^{\left(Q(y+1;\Phi)\right)^{\alpha}},\quad y\in\mathbb{N}_0. \tag{3}$$

The probability mass function (PMF), which delineates the likelihood of any specific count number, derived from the difference between successive values of the RF:

$$P(Y=y)=p^{\left(Q(y;\Phi)\right)^{\alpha}}-p^{\left(Q(y+1;\Phi)\right)^{\alpha}},\quad y\in\mathbb{N}_0. \tag{4}$$

Here, $\alpha>0$ serves as a shape parameter governing the transformation, $0<p<1$ acts as a scale parameter affecting the distribution's dispersion properties, and $\mathbb{N}_0=\{0,1,2,3,\ldots\}$ represents the set of non-negative integers. The DOW-G family offers a strong and adaptable framework for creating new discrete distributions through the use of various baseline CDFs $G(y;\Phi)$. The selection of the baseline distribution dictates the essential attributes of the resultant discrete model, whereas the parameters α and p offer supplementary adaptability in tailoring the distribution to fit particular data patterns. The modular design of the DOW-G series renders it especially appealing for creating bespoke discrete models suited to the needs of various application areas.

Notwithstanding the significant progress provided by the DOW-G family structure, the potential of numerous unique parent distributions inside this framework remains unexamined. The Lindley distribution warrants significant attention as it serves as a highly effective model for lifetime and count data, owing to its advantageous mathematical qualities and proven adaptability across several application areas [22]. The Lindley distribution, initially presented as a counter-example to demonstrate the constraints of specific fiducial arguments in statistical inference, has developed into a widely utilized model in reliability engineering, survival analysis, and biostatistics. The probability density function (PDF) of the Lindley distribution, characterized by the parameter $\theta>0$, is expressed as

$$g\left(z;\theta\right)=\frac{\theta^2}{1+\theta}\left(1+z\right)e^{-\theta z},\quad z>0, \tag{5}$$

and its corresponding CDF is

$$G\left(z;\theta\right)=1-\frac{e^{-\theta z}\left(1+\theta+\theta z\right)}{1+\theta},\quad z>0. \tag{6}$$

The Lindley distribution has numerous appealing characteristics that render it especially appropriate as a baseline in the DOW-G system. The probability density function is a weighted combination of exponential and gamma distributions, offering intrinsic flexibility in representing various data patterns. The distribution demonstrates a declining hazard rate for $\theta\leq1$ and an ascending hazard rate for $\theta>1$, allowing it to represent both enhancing and declining systems. Moreover, the Lindley distribution has exhibited a superior fit to diverse lifetime datasets in comparison to the exponential distribution, although possessing an equivalent number of parameters. This study is primarily driven by the urgent necessity to create a novel, more adaptable discrete model that can accurately represent the intricate dispersion patterns of count data in several sustainability-related fields. We want to fulfill this requirement by presenting a new discrete distribution derived from integrating the Lindley distribution as

the parent G within the DOW-G architecture. The newly developed model, referred to as the *discrete odd Weibull-Lindley* (DOWL) distribution, aims to integrate the versatility of the DOW-G family structure with the advantageous characteristics of the Lindley baseline. The suggested DOWL distribution differentiates itself from current discrete models by numerous mathematically and practically important attributes that directly tackle the modeling issues associated with sustainability data analysis:

1. The proposed DOWL distribution effectively accommodates both symmetric and asymmetric data patterns, including right-skewed and left-skewed distributions, across various kurtosis types: platykurtic, mesokurtic, and leptokurtic. This establishes its significant utility for modeling sustainability-related count data.

2. The hazard rate function of the DOWL distribution demonstrates significant flexibility, accommodating increasing, decreasing, J-shaped, and bathtub-shaped failure rate patterns. This versatility makes the distribution ideal for representing various lifetime and dependability problems in sustainability studies.

3. The DOWL distribution has been shown to outperform traditional discrete distributions in fitting over-dispersed, under-dispersed, and equi-dispersed count data, which are common in sustainability studies but often poorly represented by standard models. This extensive dispersion modeling capabilities was particularly beneficial for examining heterogeneous count data across multiple sustainability sectors.

4. The distribution inherently accommodates datasets with an abundance of zeros, a common trait of count data in environmental monitoring, epidemiological surveillance, and event occurrence analysis.

5. The DOWL distribution adeptly encompasses both light-tailed and heavy-tailed distributional characteristics, providing versatility in simulating infrequent but significant events essential for risk evaluation in sustainability frameworks.

The following portions of this paper are organized to systematically achieve these objectives. Section 2 outlines the mathematical formulation of the innovative DOWL distribution and provides visual representations of the primary functions. Section 3 presents the thorough statistical and reliability characterizations of the distribution, establishing the analytical basis for its use. Section 4 examines traditional parameter estimation methods, focusing specifically on maximum likelihood estimation and its asymptotic characteristics. Section 5 provides comprehensive Monte Carlo simulation analyses assessing the finite-sample efficacy of the estimators under diverse scenarios. Section 6 demonstrates the model's versatility and relevance with a comprehensive analysis of three actual datasets, including several sustainability sectors. Section 7 provides closing remarks that synthesize the contributions of this study and delineate avenues for future research in the advancement of discrete distributions for sustainability applications.

2. Mathematical Synthesis of the DOWL Distribution

The derivation of the odds function $Q(x;\theta)$ for the Lindley distribution is a prerequisite for constructing the DOWL distribution. The RF of the Lindley distribution is derived from Equation (6) as

$$\overline{G}\,(x;\theta) = \frac{e^{-\theta x}\,(1+\theta+\theta x)}{1+\theta}, \quad x > 0. \tag{7}$$

The odds function, defined as the ratio of the CDF to the RF, is therefore

$$Q\,(x;\theta) = \frac{(1+\theta)e^{\theta x}}{1+\theta+\theta x} - 1. \tag{8}$$

Inserting the Lindley odds function into the general PMF of the DOW-G family (Equation 4) produces the PMF of the DOWL distribution. Let X be a random variable that adheres to the DOWL distribution characterized by

parameters $\alpha > 0$, $0 < p < 1$, and $\theta > 0$. The PMF of X is defined as follows:

$$P(X = x) = p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha}, \quad x \in \mathbb{N}_0, \tag{9}$$

For the special case $x = 0$, we have $Q(0;\theta) = 0$ since $e^0 = 1$, which gives

$$P(X = 0) = p^0 - p^{[Q(1;\theta)]^\alpha} = 1 - p^{[Q(1;\theta)]^\alpha}, \tag{10}$$

where $Q(1;\theta) = \frac{(1+\theta)e^\theta - (1+2\theta)}{1+2\theta}$. This precise definition for the likelihood at zero is particularly pertinent for sustainability applications using zero-inflated count data, where the model's capacity to account for excess zeros is essential. The CDF of the DOWL distribution is derived directly from Equation (1):

$$F(x) = 1 - p^{[Q(x+1;\theta)]^\alpha}, \quad x \in \mathbb{N}_0. \tag{11}$$

The RF representing the probability that X exceeds a given value x, is $R(x) = p^{[Q(x+1;\theta)]^\alpha}$, these expressions retain the refined structure of the DOW-G family while integrating the distinct attributes of the Lindley baseline via the odds function $Q(x;\theta)$. Figure 1 illustrates the diverse configurations of the CDF for various parameter combinations, demonstrating the adaptability of the DOWL distribution. Figure 2 illustrates the PMF plots for specific values of the model parameters.

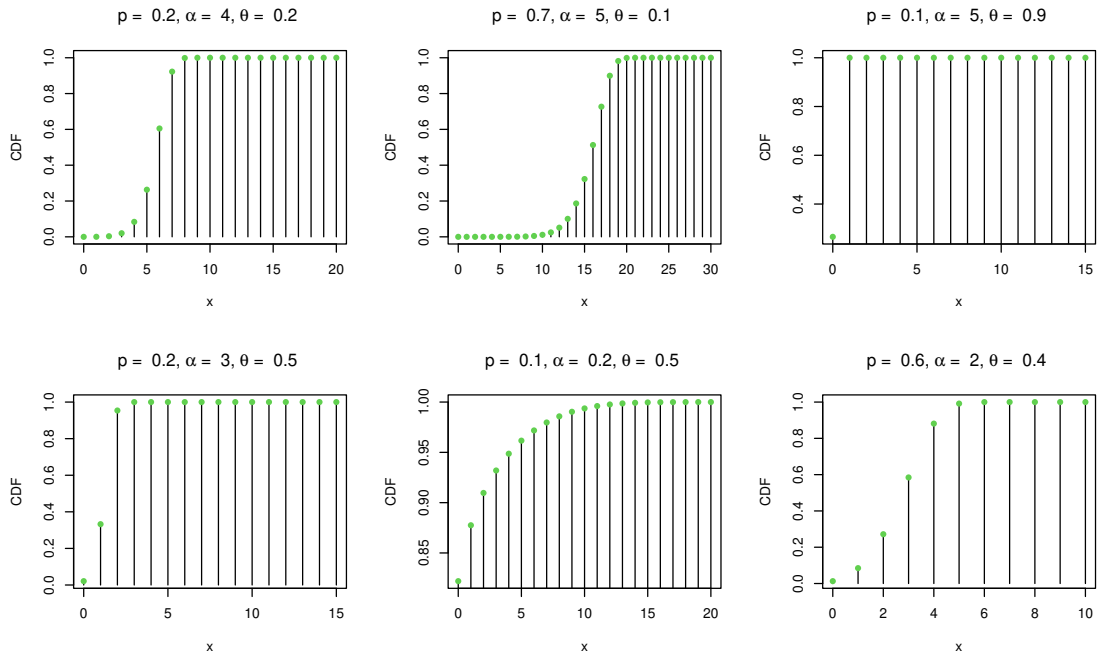


Figure 1. Various shapes for the CDF of the DOWL distribution.

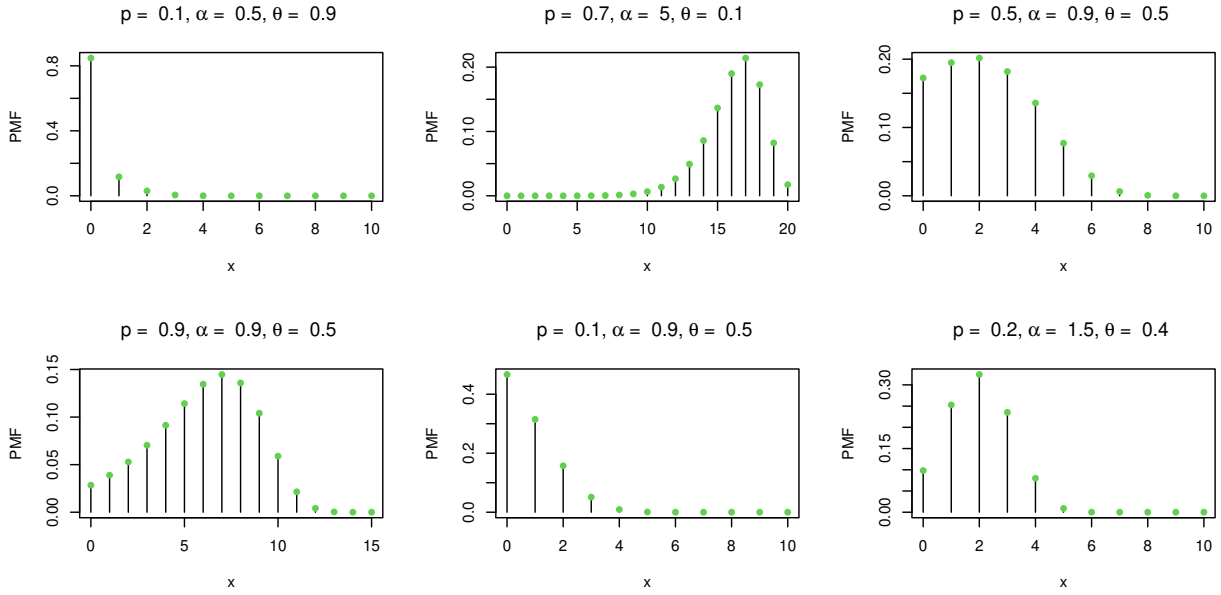


Figure 2. Various shapes for the PMF of the DOWL distribution.

The PMF of the DOWL distribution is very flexible, as seen in Figure 2. It can either drop monotonically or exhibit a unimodal structure with a right-skewed tail and a left-skewed tail. When modeling different types of count data, such as those with a high frequency of small values or those where moderate values are more likely, the DOWL distribution is a suitable option.

3. Statistical Properties

3.1. Parameter Space and Model Validity

For the DOWL distribution to qualify as a legitimate probability model, the PMF must adhere to the established criteria. (i) Non-negativity: $P(X = x) \geq 0$ for all $x \in \mathbb{N}_0$; and (ii) Normalization: $\sum_{x=0}^{\infty} P(X = x) = 1$. The non-negativity condition is fulfilled as $Q(x;\theta)$ is strictly increasing in x for $x \geq 0$ and $\theta > 0$, indicating that $[Q(x;\theta)]^{\alpha} < [Q(x+1;\theta)]^{\alpha}$. Given that $0 < p < 1$, it follows that $p^{[Q(x;\theta)]^{\alpha}} > p^{[Q(x+1;\theta)]^{\alpha}}$, thereby guaranteeing $P(X = x) > 0$ for all $x \in \mathbb{N}_0$. The normalization requirement is confirmed by noting that

$$\begin{aligned}\sum_{x=0}^{\infty} P(X = x) &= \sum_{x=0}^{\infty} \left[p^{[Q(x;\theta)]^{\alpha}} - p^{[Q(x+1;\theta)]^{\alpha}}\right] \\ &= \lim_{n\to\infty} \left[p^{[Q(0;\theta)]^{\alpha}} - p^{[Q(n+1;\theta)]^{\alpha}}\right] \\ &= p^0 - 0 = 1. \end{aligned} \tag{12}$$

We utilize the condition that $Q(0;\theta) = 0$ and the limit $\lim_{x\to\infty} Q(x;\theta) = \infty$, which indicates that $\lim_{x\to\infty} p^{[Q(x;\theta)]^{\alpha}} = 0$ for $0 < p < 1$. The parameter space of the DOWL distribution is defined as $\Theta = \{(\alpha, p, \theta) : \alpha > 0,\ 0 < p < 1,\ \theta > 0\}$.

3.2. *Identifiability of the Mode and Parameter Effects*

The DOWL distribution is identifiable because its parameter vector (α, p, θ) uniquely determines the probability law. In particular, the parameter $\theta > 0$ enters through the Lindley odds function, so different values of θ produce different shapes for the baseline odds. The parameter $\alpha > 0$ acts as a shape parameter by controlling the power transformation $[Q(x;\theta)]^\alpha$, while $p \in (0,1)$ regulates the decay rate of the tail through the factor $p^{[Q(x;\theta)]^\alpha}$. Hence, the three parameters play distinct roles and are separately identifiable. The mode of the DOWL distribution is the value $x^* \in \mathbb{N}_0$ that maximizes the PMF. Its location depends on the joint effect of the parameters. In general, larger values of θ or α increase the rate at which the PMF decreases, thereby concentrating probability mass near zero and often producing a zero-mode. By contrast, values of p closer to 1 yield a slower decay in the survival probabilities, allowing the mode to shift to higher counts. Therefore, θ mainly controls the baseline odds growth, α governs the shape and peakedness, and p determines tail persistence. These distinct effects improve model flexibility and make the DOWL distribution suitable for overdispersed and zero-heavy count data.

3.3. *Quantile Function*

The quantile function for a discrete random variable X is defined as:

$$x_q = \inf\{x \in \mathbb{N}_0 : F(x) \geq q\}, \quad 0 < q < 1. \tag{13}$$

To derive the quantile function of the DOWL distribution, we set $F(x) \geq q$ using the CDF in Equation (11) and solve for x. This yields:

$$p^{[Q(x+1;\theta)]^\alpha} \leq 1 - q.$$

Taking logarithms and noting that $\log p < 0$ for $0 < p < 1$, we obtain:

$$Q(x+1;\theta) \geq \left[\frac{\log(1-q)}{\log p}\right]^{1/\alpha}.$$

Substituting the Lindley odds function from Equation (8) and defining the auxiliary quantity $\xi_q = \left[\frac{\log(1-q)}{\log p}\right]^{1/\alpha} + 1$, we require:

$$\frac{(1+\theta)e^{\theta(x+1)}}{1+\theta+\theta(x+1)} \geq \xi_q.$$

This equation incorporates x in both the exponent and linearity, necessitating the Lambert W function for a closed-form solution, which is defined as the inverse function that satisfies $W(z)e^{W(z)} = z$. The quantile function of the DOWL distribution is derived by suitable substitution and algebraic manipulation as follows:

$$x_q = \left\lceil \frac{-W\left(-\dfrac{1+\theta}{\xi_q \cdot e^{1+\theta}}\right) - (1+\theta)}{\theta} - 1 \right\rceil, \tag{14}$$

where $\xi_q = \left[\frac{\log(1-q)}{\log p}\right]^{1/\alpha} + 1$, $W(\cdot)$ signifies the primary branch of the Lambert W function, and $\lceil \cdot \rceil$ denotes the ceiling function. The quantile function enables the production of random variates from the DOWL distribution through the inverse transform approach. For a uniform random variable $U \sim \text{Uniform}(0,1)$, a random variate X from the DOWL distribution can be generated by calculating $X = Q(U)$ using either Equation (14).

3.4. *Hazard Rate Function and Classification of Aging Characteristics*

The hazard rate function (HRF), commonly known as the failure rate, is essential in dependability analysis and offers insight into the aging properties of a system or phenomenon. For discrete distributions, the HRF at a point x is defined as the conditional chance of failure at x a given survival until x where $h(x) = P(X = x | X \geq x)$. For

the DOWL distribution:

$$h(x) = 1 - p^{[Q(x+1;\theta)]^\alpha - [Q(x;\theta)]^\alpha}. \tag{15}$$

The behavior of $h(x)$ is contingent upon the difference $\Delta(x) = [Q(x+1;\theta)]^\alpha - [Q(x;\theta)]^\alpha$, governed by the parameters α, p, and θ. The DOWL distribution may display multiple aging categories, defined as follows:

3.4.1. Increasing Failure Rate (IFR) A distribution is classified as IFR if its HRF $h(x)$ is non-decreasing in x. The IFR property is satisfied for the DOWL distribution when:

$$h(x+1) \geq h(x) \quad \text{for all } x \in \mathbb{N}_0.$$

This condition is equivalent to:

$$\Delta(x+1) \geq \Delta(x) \quad \text{for all } x \in \mathbb{N}_0,$$

where $\Delta(x) = [Q(x+1;\theta)]^\alpha - [Q(x;\theta)]^\alpha$. The IFR trait generally applies to elevated levels of α and θ, signifying that the system degrades over time with a heightened likelihood of failure. This phenomenon is frequently noted in mechanical systems experiencing wear and deterioration.

3.4.2. Decreasing Failure Rate (DFR) A distribution is classified as DFR if its HRF $h(x)$ is non-increasing with respect to x. The DFR property for the DOWL distribution is satisfied when:

$$h(x+1) \leq h(x) \quad \text{for all } x \in \mathbb{N}_0.$$

This condition is satisfied when:

$$\Delta(x+1) \leq \Delta(x) \quad \text{for all } x \in \mathbb{N}_0.$$

The DFR feature generally manifests for lower values of α, indicating that the system enhances over time or that less robust units fail prematurely, resulting in a more dependable population. This feature pertains to systems demonstrating neonatal mortality or burn-in phenomena.

3.4.3. Bathtub-Shaped Failure Rate (BFR) The bathtub-shaped failure rate is defined by an initial decline in the HRF, succeeded by a stable or nearly stable phase, and ultimately an increase in the HRF. The DOWL distribution may demonstrate this behavior under the following conditions:

$$\begin{cases} h(x+1) < h(x) & \text{for } x < x_1, \\ h(x+1) \approx h(x) & \text{for } x_1 \leq x \leq x_2, \\ h(x+1) > h(x) & \text{for } x > x_2, \end{cases}$$

for certain change points x_1 and x_2. This configuration is prevalent in numerous practical applications, encompassing electronic components and biological systems, where initial failures are succeeded by a stable usable life phase and subsequent deterioration.

3.4.4. Upside-Down Bathtub-Shaped Failure Rate (UBF) The inverted bathtub-shaped failure rate is defined by an initial rise in the HRF, which peaks before subsequently declining. The DOWL distribution demonstrates this behavior under the following conditions:

$$\begin{cases} h(x+1) > h(x) & \text{for } x < x^*, \\ h(x+1) < h(x) & \text{for } x > x^*, \end{cases}$$

for a specific modal point x^*. This pattern is seen in systems where the likelihood of failure first escalates but then diminishes as the system adjusts or as vulnerable units are eliminated from the population. The cumulative hazard rate function (CHRF) is defined as $H(x) = -[Q(x+1;\theta)]^\alpha \ln(p)$. The reversed hazard rate function (RHRF) is

defined as $r(x) = P(X = x|X \leq x)$. For the DOWL distribution, the RHRF can be formulated as

$$r(x) = \frac{p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha}}{1 - p^{[Q(x+1;\theta)]^\alpha}}. \tag{16}$$

Figures 3 and 4 depict the graphical representations of the HRF and RHRF, respectively. The analysis of Figure 3 reveals the significant adaptability of the DOWL distribution's HRF. The HRF displays both monotonic and non-monotonic configurations across several parameter combinations, including rising, J-shaped, and bathtub-shaped failure rate patterns. This versatility is a significant advantage, allowing the DOWL model to delineate the aging and degradation behavior of systems and components across many application areas. Additionally, Figure 4 depicts the behavior of the reversed hazard rate function, which is very useful for evaluating left-censored or skewed data structures. The varied forms displayed by both the PMF and HRF highlight the DOWL distribution's potential as a flexible and resilient modeling framework for count data and lifetime phenomena across multiple disciplines, including sustainability science and related areas.

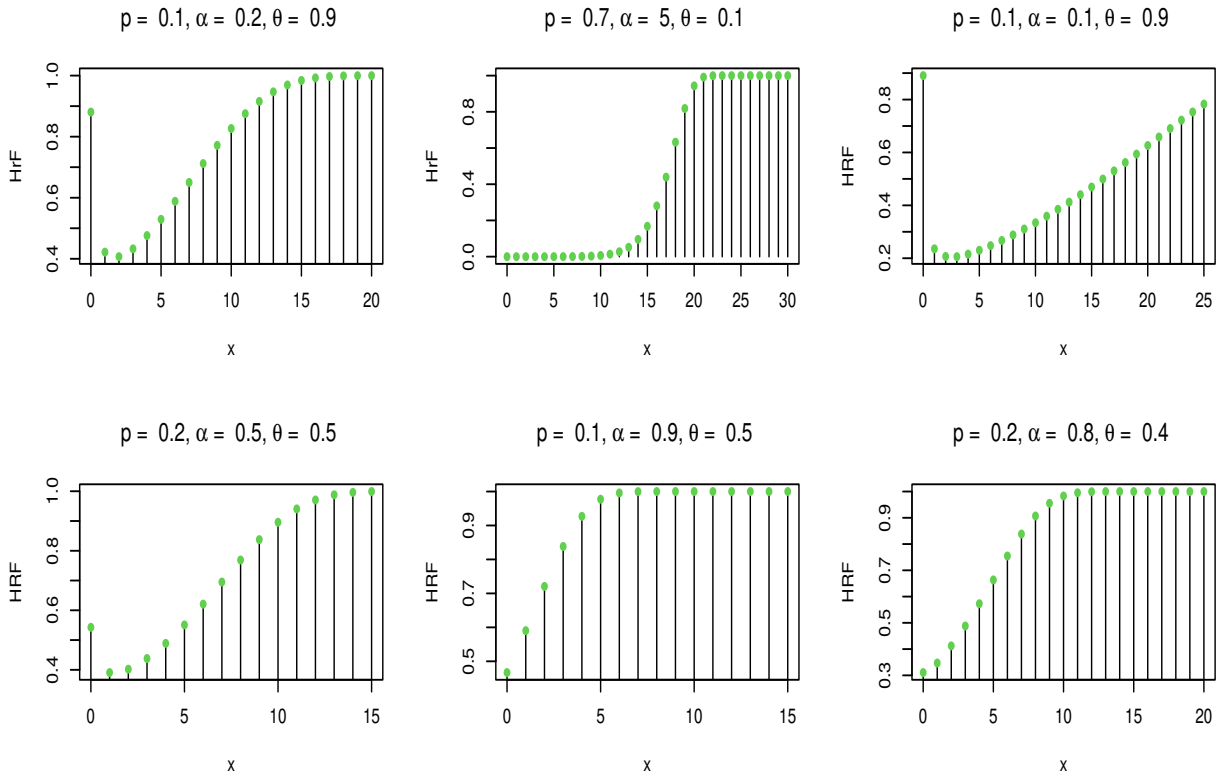


Figure 3. Various shapes of the HRF for the DOWL distribution.

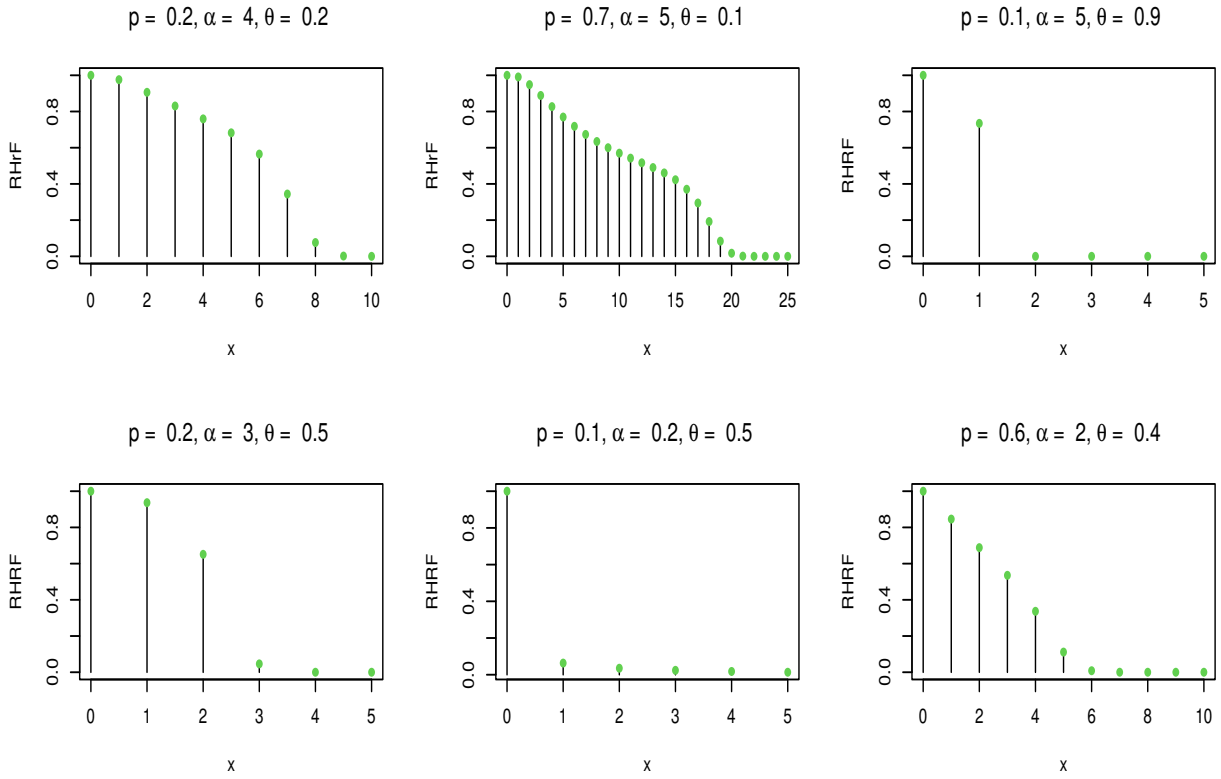


Figure 4. Various shapes of the RHRF for the DOWL distribution.

3.5. Moments and Related Measures

The r-th raw moment of a discrete random variable X with PMF $P(X = x)$ is defined as

$$\mu'_r = E[X^r] = \sum_{x=0}^{\infty} x^r P(X = x). \tag{17}$$

For the DOWL distribution, this becomes

$$\mu'_r = \sum_{x=0}^{\infty} x^r \left[p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha} \right]. \tag{18}$$

Using the identity $x^r = \sum_{k=0}^{x-1} [(k+1)^r - k^r]$ and interchanging the order of summation, the r-th raw moment can be expressed in terms of the RF:

$$\mu'_r = \sum_{x=0}^{\infty} [(x+1)^r - x^r] R(x) = \sum_{x=0}^{\infty} [(x+1)^r - x^r] p^{[Q(x+1;\theta)]^\alpha}. \tag{19}$$

Setting $r = 1$ in Equation (19), the mean of the DOWL distribution is

$$\mu = E[X] = \sum_{x=0}^{\infty} R(x) = \sum_{x=0}^{\infty} p^{[Q(x+1;\theta)]^\alpha}. \tag{20}$$

This expression, although lacking a straightforward closed form, can be computed efficiently by numerical methods owing to the series' rapid convergence. The variance is calculated as $\sigma^2 = E[X^2] - (E[X])^2$. Utilizing the formula for the second raw instant,

$$E[X^2] = \sum_{x=0}^{\infty}(2x+1)R(x) = 2\sum_{x=0}^{\infty} x \cdot p^{[Q(x+1;\theta)]^\alpha} + \mu, \tag{21}$$

we obtain

$$\sigma^2 = 2\sum_{x=0}^{\infty} x \cdot p^{[Q(x+1;\theta)]^\alpha} + \mu - \mu^2 = 2\sum_{x=1}^{\infty} x \cdot p^{[Q(x+1;\theta)]^\alpha} + \mu(1-\mu). \tag{22}$$

The dispersion index (DI), or variance-to-mean ratio, is an essential metric for evaluating whether count data display over-dispersion, equi-dispersion, or under-dispersion.

$$DI = \frac{2\sum_{x=1}^{\infty} x \cdot p^{[Q(x+1;\theta)]^\alpha}}{\mu} + 1 - \mu. \tag{23}$$

When $DI > 1$, the data are over-dispersed compared to a Poisson model; when $DI < 1$, the data are under-dispersed; and when $DI = 1$, the data demonstrate equi-dispersion. The DOWL distribution's capacity to handle both $DI > 1$ and $DI < 1$ via suitable parameter selection renders it especially advantageous for sustainability applications, where count data often diverge from Poisson assumptions. Skewness (the third standardized moment) and kurtosis (the fourth standardized moment) are derived from the central moments.

$$\gamma_1 = \frac{\mu_3}{\sigma^3}, \quad \gamma_2 = \frac{\mu_4}{\sigma^4}, \tag{24}$$

where $\mu_3 = E[(X-\mu)^3]$ and $\mu_4 = E[(X-\mu)^4]$. The r-th factorial moment of X is defined as

$$\mu_{(r)} = E[X^{(r)}] = E[X(X-1)(X-2)\cdots(X-r+1)] = \sum_{x=r}^{\infty} \frac{x!}{(x-r)!} P(X=x). \tag{25}$$

For the DOWL distribution:

$$\mu_{(r)} = \sum_{x=r}^{\infty} \frac{x!}{(x-r)!} \left[p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha} \right]. \tag{26}$$

The moment generating function (MGF) of the DOWL distribution is

$$M_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} P(X=x) = \sum_{x=0}^{\infty} e^{tx} \left[p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha} \right], \tag{27}$$

which is valid for any genuine t. The probability generating function (PGF), especially advantageous for discrete distributions, is

$$G_X(s) = E[s^X] = \sum_{x=0}^{\infty} s^x P(X=x), \quad |s| \le 1. \tag{28}$$

We have calculated the mean, variance, skewness, kurtosis, and DI for many parameter combinations to demonstrate the flexibility of the DOWL distribution. The findings are encapsulated in Tables 1 and 2.

Table 1. Numerical descriptive statistics of the DOWL distribution for selected values of p and α, with $\theta = 0.5$.

Measure	α	p								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E(X)$	0.5	0.6158	1.0763	1.5755	2.1513	2.8437	3.7134	4.8700	6.5568	9.5414
	1.5	1.1248	1.3914	1.6287	1.8694	2.1326	2.4396	2.8257	3.3672	4.3102
	3.0	1.5431	1.6926	1.8234	1.9561	2.0979	2.2552	2.4420	2.6988	3.1409
Var(X)	0.5	1.2275	2.4431	3.8833	5.6187	7.7325	10.3404	13.6167	17.8489	23.5711
	1.5	0.6602	0.8162	0.9586	1.1050	1.2657	1.4522	1.6823	1.9912	2.4754
	3.0	0.3328	0.3475	0.3775	0.4138	0.4462	0.4690	0.4900	0.5381	0.6152
Sk(X)	0.5	2.1967	1.6694	1.3302	1.0593	0.8183	0.5864	0.3466	0.0754	−0.2818
	1.5	0.2540	0.1745	0.1128	0.0556	−0.0028	−0.0666	−0.1418	−0.2395	−0.3923
	3.0	−0.4453	−0.4119	−0.2843	−0.2296	−0.2808	−0.3975	−0.4753	−0.4245	−0.5259
Ku(X)	0.5	8.2140	5.5471	4.2216	3.3867	2.8111	2.4073	2.1459	2.0474	2.2582
	1.5	2.4836	2.4752	2.4702	2.4685	2.4730	2.4880	2.5205	2.5903	2.7677
	3.0	2.5743	3.1903	3.3704	3.1885	2.9271	2.8383	3.0664	3.3064	3.1744
DI(X)	0.5	1.9933	2.2700	2.4648	2.6118	2.7192	2.7846	2.7960	2.7222	2.4704
	1.5	0.5869	0.5866	0.5886	0.5911	0.5935	0.5953	0.5953	0.5913	0.5743
	3.0	0.2156	0.2053	0.2070	0.2115	0.2127	0.2079	0.2007	0.1994	0.1959

Table 2. Numerical descriptive statistics of the DOWL distribution for selected values of p and θ, with $\alpha = 1.5$.

Measure	θ	p								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E(X)$	0.3	2.5111	2.9725	3.3807	3.7927	4.2415	4.7634	5.4177	6.3324	7.9200
	0.9	0.2845	0.4218	0.5427	0.6664	0.8048	0.9685	1.1743	1.4646	1.9761
	1.5	0.0038	0.0204	0.0545	0.1092	0.1872	0.2909	0.4225	0.5888	0.8621
Var(X)	0.3	1.8722	2.3066	2.7026	3.3086	3.5535	4.0687	4.7028	5.5517	6.8776
	0.9	0.2053	0.2585	0.2986	0.3438	0.3986	0.4603	0.5296	0.6270	0.7809
	1.5	0.0038	0.0200	0.0515	0.0973	0.1522	0.2063	0.2445	0.2535	0.2928
Sk(X)	0.3	0.1556	0.0949	0.0436	−0.0064	−0.0588	−0.1175	−0.1879	−0.2808	−0.4281
	0.9	0.9833	0.4829	0.2941	0.2369	0.1887	0.0880	−0.0379	−0.1281	−0.3083
	1.5	16.0707	6.7778	3.9263	2.5063	1.6035	0.9209	0.3201	−0.2273	−0.0893
Ku(X)	0.3	2.5408	2.5159	2.5028	2.4980	2.5020	2.5178	2.5528	2.6267	2.8108
	0.9	2.0490	1.6360	1.9861	2.3445	2.4225	2.3261	2.3507	2.5380	2.6459
	1.5	259.267	46.9385	16.4162	7.2816	3.5713	1.8487	1.1196	1.3870	3.1525
DI(X)	0.3	0.7456	0.7760	0.7994	0.8196	0.8378	0.8542	0.8680	0.8767	0.8684
	0.9	0.7216	0.6128	0.5502	0.5159	0.4953	0.4753	0.4510	0.4281	0.3951
	1.5	0.9962	0.9796	0.9455	0.8908	0.8128	0.7091	0.5787	0.4305	0.3396

Table 1 presents descriptive statistics for $\theta = 0.5$ across various α and p values, revealing several important patterns. For fixed α, both mean and variance increase with p. However, for fixed p, the mean increases while variance decreases with α, demonstrating an inverse relationship. The distribution exhibits positive skewness at low p and α, transitioning to negative skewness as these parameters increase, demonstrating flexibility for both right- and left-tailed data. Kurtosis values span leptokurtic (Ku > 3), mesokurtic (Ku ≈ 3), and platykurtic (Ku < 3) regimes. The DI reveals adaptability: DI > 1 (over-dispersion) at $\alpha = 0.5$, DI < 1 (under-dispersion) at $\alpha = 1.5$, and substantial under-dispersion at $\alpha = 3.0$, confirming suitability for modeling count data with varying dispersion levels.

Table 2 examines the effect of θ on distributional moments with $\alpha = 1.5$ fixed. Both mean and variance decrease substantially with increasing θ at constant p; for example, at $p = 0.9$, the mean drops from 7.92 to 0.86 as θ increases from 0.3 to 1.5. Skewness ranges from extreme positive values ($Sk = 16.07$ at $\theta = 1.5$, $p = 0.1$) to negative values at higher p. Kurtosis spans from highly leptokurtic (259.27) to near-mesokurtic regimes. The DI indicates under-dispersion (DI < 1) across all θ values, generally decreasing with p, though a slight non-monotonic pattern appears at $\theta = 0.3$.

3.6. Entropy

Entropy measures the level of uncertainty or randomness related to a random variable and is a core principle in information theory. We develop the Rényi entropy for the DOWL distribution, which generalizes various established entropy metrics. The Rényi entropy of order σ ($\sigma > 0$, $\sigma \neq 1$) for a discrete random variable X with PMF $P(X = x)$ is defined as

$$I_\sigma(X) = \frac{1}{1-\sigma} \log \left(\sum_{x=0}^{\infty} [P(X = x)]^\sigma \right). \tag{29}$$

The Rényi entropy of the DOWL distribution, derived from its PMF, is expressed as follows:

$$I_\sigma(X) = \frac{1}{1-\sigma} \log \left(\sum_{x=0}^{\infty} \left[p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha} \right]^\sigma \right), \tag{30}$$

The Rényi entropy includes several significant specific cases: the Shannon entropy as $\sigma \to 1$, the collision entropy at $\sigma = 2$, the min-entropy as $\sigma \to \infty$, and the max-entropy as $\sigma \to 0$. The Shannon entropy for the DOWL distribution is specifically derived by evaluating the limit as $\sigma \to 1$ in Equation (30), resulting in

$$I(X) = -\sum_{x=0}^{\infty} P(X = x) \log P(X = x), \tag{31}$$

which, upon substitution, becomes

$$I(X) = -\sum_{x=0}^{\infty} \left(p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha} \right) \times \log \left(p^{[Q(x;\theta)]^\alpha} - p^{[Q(x+1;\theta)]^\alpha} \right), \tag{32}$$

To examine the behavior of entropy under various parameter configurations, we calculate the Rényi entropy at $\sigma = 2$ (collision entropy) for the DOWL distribution. The numerical findings are encapsulated in Tables 3 and 4.

Table 3. Rényi entropy ($\sigma = 2$) of the DOWL distribution for selected values of p and α, with $\theta = 0.5$.

α	p								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	0.7119	1.0571	1.3569	1.6398	1.9157	2.1861	2.4449	2.6745	2.8379
1.5	1.0934	1.1993	1.2801	1.3507	1.4173	1.4834	1.5520	1.6269	1.7150
3.0	0.7564	0.7225	0.7372	0.7918	0.8609	0.9073	0.9067	0.9138	1.0145

Table 3 displays the entropy values for $\theta = 0.5$ across several α and p parameters. When $\alpha = 0.5$, the entropy increases monotonically as p approaches one, indicating heightened distributional uncertainty. A comparable upward tendency is noted for $\alpha = 1.5$, though with a more gradual rate of development. For $\alpha = 3.0$, the entropy demonstrates a minor initial decline, succeeded by an increase, with the minimal uncertainty levels recorded at lower values of p.

Table 4. Rényi entropy ($\sigma = 2$) of the DOWL distribution for selected values of p and θ, with $\alpha = 1.5$.

θ	p								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.3	1.6103	1.7162	1.7960	1.8655	1.9309	1.9955	2.0625	2.1356	2.2214
0.9	0.5224	0.6759	0.7427	0.7849	0.8385	0.9139	0.9910	1.0527	1.1547
1.5	0.0077	0.0409	0.1087	0.2163	0.3629	0.5320	0.6697	0.6787	0.6316

Table 4 demonstrates the influence of θ on entropy with α held constant at 1.5. For a specified θ, the entropy often escalates as p increases. Significantly, when θ increases, the entropy values decrease markedly, especially for lower values of p. At $p = 0.1$, the entropy diminishes from 1.6103 to 0.0077 as θ escalates from 0.3 to 1.5, indicating a significant drop in distributional uncertainty. This behavior aligns with the moment structure of the distribution, where bigger values of θ result in reduced mean and variance, signifying a more concentrated probability mass.

3.7. Residual and Past Lifetime Functions: Theory and Aging Categories

In reliability theory and survival analysis, the mean residual life (MRL) and mean past lifetime (MPL) functions offer critical insights into the anticipated remaining and elapsed lifespan of a system, respectively. These functions, together with their variances, provide significant characterizations of the DOWL distribution.

3.7.1. Mean Residual Life Function The MRL function denotes the anticipated remaining lifetime of a system that has endured until time t. For a discrete random variable X with RF $S(x) = P(X > x)$, the MRL function is defined as:

$$m(t) = E(X - t|X > t) = \frac{1}{S(t)} \sum_{x=t+1}^{\infty} S(x), \quad t \in \mathbb{N}_0.$$

For the DOWL distribution, using the RF $S(x) = p^{[Q(x+1;\theta)]^\alpha}$, we have:

$$m(t) = \frac{1}{p^{[Q(t+1;\theta)]^\alpha}} \sum_{x=t+1}^{\infty} p^{[Q(x+1;\theta)]^\alpha}.$$

This can be expressed more compactly as:

$$m(t) = p^{-[Q(t+1;\theta)]^\alpha} \sum_{j=1}^{\infty} p^{[Q(t+j+1;\theta)]^\alpha}, \tag{33}$$

where $Q(x;\theta) = \frac{(1+\theta)e^{\theta x}}{1+\theta+\theta x} - 1$ represents the Lindley odds function as delineated in Equation (8). The variation of the residual life, indicating the variability in the remaining lifetime, is expressed as:

$$\text{Var}(X - t|X > t) = E\left[(X - t)^2|X > t\right] - [m(t)]^2.$$

The second moment can be computed as:

$$E\left[(X - t)^2|X > t\right] = \frac{1}{S(t)} \sum_{x=t+1}^{\infty} (x - t)^2 P(X = x).$$

Utilizing the equation $\sum_{x=t+1}^{\infty}(x - t)^2 P(X = x) = \sum_{x=t+1}^{\infty}(2x - 2t - 1)S(x)$, the variance of residual life is expressed as:

$$\sigma_m^2(t) = \frac{2}{S(t)} \sum_{x=t+1}^{\infty} (x - t)S(x) - m(t) - [m(t)]^2. \tag{34}$$

A discrete distribution can be distinctly defined by its MRL function. The MRL function in Equation (33) uniquely characterizes the DOWL distribution via the established relationship:

$$S(t) = \prod_{j=0}^{t-1} \frac{m(j)}{1+m(j)}, \quad t \geq 1,$$

given that $S(0) = 1$. This indicates that if two discrete distributions possess similar MRL functions, they must be identical distributions. The DOWL distribution's flexibility enables it to represent diverse MRL patterns, such as growing, decreasing, and bathtub-shaped MRL functions, contingent upon the parameter values. The MRL function is intricately associated with the HRF $h(t) = \frac{P(X=t)}{S(t-1)}$. The subsequent relationship is applicable to the DOWL distribution:

$$m(t) = \frac{1 - h(t+1)[1+m(t+1)]}{h(t+1)}.$$

This recursive relationship illustrates that the MRL function entirely defines the HRF and vice versa. The DOWL distribution can display multiple HRF forms, such as growing, decreasing, bathtub-shaped, and inverted bathtub-shaped hazard rates, rendering it appropriate for modeling a range of reliability scenarios.

3.7.2. Mean Residual Life and New Better than Used Aging Classes The MRL function offers a distinct characterization of aging features. The MRL function for the DOWL distribution is specified in Equation (33). A distribution is classified as belonging to the increasing mean residual life (IMRL) class if the function $m(t)$ is non-decreasing in t, specifically, if $m(t+1) \geq m(t)$ for all $t \in \mathbb{N}_0$. The IMRL feature signifies negative aging, wherein surviving units are likely to possess extended predicted remaining lifespans. This phenomenon is evident in systems that enhance with time or in populations where less robust units fail prematurely. In contrast, the DOWL distribution is classified inside the decreasing mean residual life (DMRL) category if $m(t)$ is non-increasing in t, signifying that $m(t+1) \leq m(t)$ for all $t \in \mathbb{N}_0$. The DMRL feature signifies positive aging, wherein surviving units have reduced anticipated remaining lifespans, reflecting the standard aging trend for systems experiencing wear and degradation.

The new better than used (NBU) and new worse than used (NWU) qualities offer supplementary characterizations of aging behavior. The DOWL distribution fulfills the NBU property if $S(x+t) \leq S(x) \cdot S(t)$ for all $x, t \in \mathbb{N}_0$. Substituting the RF of the DOWL distribution produces

$$p^{[Q(x+t+1;\theta)]^\alpha} \leq p^{[Q(x+1;\theta)]^\alpha} \cdot p^{[Q(t+1;\theta)]^\alpha}. \tag{35}$$

Since $0 < p < 1$, this condition is equivalent to

$$[Q(x+t+1;\theta)]^\alpha \geq [Q(x+1;\theta)]^\alpha + [Q(t+1;\theta)]^\alpha. \tag{36}$$

The NBU property indicates that a new system is stochastically superior to a used system and is applicable for specific combinations of the parameters α, p, and θ. The DOWL distribution fulfills the NWU condition if the reverse inequality is satisfied, namely $S(x+t) \geq S(x) \cdot S(t)$ for all $x, t \in \mathbb{N}_0$, which is equivalent to

$$[Q(x+t+1;\theta)]^\alpha \leq [Q(x+1;\theta)]^\alpha + [Q(t+1;\theta)]^\alpha. \tag{37}$$

The NWU property indicates that a utilized system is stochastically superior to a new system, which may be observed in systems that enhance with age or have learning effects. A less robust variant of the NBU property is the new better than used in expectation (NBUE) property. The DOWLi distribution fulfills the NBUE property if $m(t) \leq m(0) = E(X)$ for all $t \in \mathbb{N}_0$, indicating that the anticipated remaining lifespan of a utilized system does not exceed the projected lifespan of a fresh system. Likewise, the new worse than used in expectation (NWUE) property is satisfied if $m(t) \geq m(0)$ for every $t \in \mathbb{N}_0$.

3.7.3. Mean Past Lifetime Function The MPL function denotes the anticipated duration of a system's lifetime that has failed by time t. For a discrete random variable X with CDF $F(x)$, the MPL function is defined as:

$$\tilde{m}(t) = E(t - X|X \leq t) = t - \frac{1}{F(t)} \sum_{x=0}^{t} x \cdot P(X = x), \quad t \in \mathbb{N}_0.$$

Alternatively, using the relation $\sum_{x=0}^{t} x \cdot P(X = x) = \sum_{x=0}^{t-1} F(x) - t \cdot F(t) + \sum_{x=0}^{t} F(x)$, the MPL can be written as:

$$\tilde{m}(t) = t - \frac{1}{F(t)} \sum_{x=0}^{t-1} F(x).$$

For the DOWL distribution, substituting the CDF from Equation (11):

$$\tilde{m}(t) = t - \frac{\sum_{x=0}^{t-1} \left(1 - p^{[Q(x+1;\theta)]^\alpha}\right)}{1 - p^{[Q(t+1;\theta)]^\alpha}}. \tag{38}$$

The variance of the past lifetime is given by:

$$\text{Var}(t - X|X \leq t) = E\left[(t - X)^2 | X \leq t\right] - \left[\tilde{m}(t)\right]^2.$$

Following similar derivations, the variance of past lifetime becomes:

$$\sigma^2_{\tilde{m}}(t) = \frac{2}{F(t)} \sum_{x=0}^{t-1} (t - x)F(x) - \tilde{m}(t) - \left[\tilde{m}(t)\right]^2. \tag{39}$$

The MPL function distinctly defines the DOWL distribution. The CDF can be derived from the MPL function via:

$$F(t) = \prod_{j=1}^{t} \frac{\tilde{m}(j)}{1 + \tilde{m}(j)}, \quad t \geq 1,$$

where $F(0) = P(X = 0)$. This definition is especially beneficial for analyzing reversed hazard rates and studies of former lifetimes.

3.8. Equilibrium Distribution

The equilibrium distribution corresponding to the DOWL distribution possesses a PMF:

$$f_e(x) = \frac{S(x)}{E(X)} = \frac{p^{[Q(x+1;\theta)]^\alpha}}{\mu},$$

where $\mu = E(X)$ denotes the mean of the DOWL distribution. This equilibrium distribution emerges inherently in renewal theory and is applicable in the analysis of stationary renewal processes. The vitality function, characterized as the quotient of the MRL to the MPL, offers a thorough assessment of system aging:

$$V(t) = \frac{m(t)}{\tilde{m}(t)}.$$

In the DOWL distribution, $V(t) > 1$ signifies that the anticipated future lifespan surpasses the estimated past lifespan at time t, indicating that the system is comparatively youthful. In contrast, $V(t) < 1$ signifies a deteriorating system. The behavior of $V(t)$ under various parameter combinations offers significant insights into the lifecycle attributes of systems represented by the DOWL distribution.

4. Parameter Estimation

This section formulates parameter estimate techniques for the DOWL distribution utilizing maximum likelihood method. Consider $x_1, x_2, \ldots, x_n$ as a random sample drawn from the DOWL distribution characterized by the parameter vector $\mathbf{\Phi} = (\alpha, p, \theta)$. The likelihood function is

$$L(\mathbf{\Phi}) = \prod_{i=1}^{n} P(X = x_i; \mathbf{\Phi}) = \prod_{i=1}^{n} \left[p^{[Q(x_i;\theta)]^\alpha} - p^{[Q(x_i+1;\theta)]^\alpha} \right]. \tag{40}$$

The log-likelihood function can be formulated as

$$\ell(\mathbf{\Phi}) = \sum_{i=1}^{n} \ln \left[p^{[Q(x_i;\theta)]^\alpha} - p^{[Q(x_i+1;\theta)]^\alpha} \right]. \tag{41}$$

The maximum likelihood estimates (MLEs) are obtained by solving the score equations $\partial\ell/\partial\alpha = 0$, $\partial\ell/\partial p = 0$, and $\partial\ell/\partial\theta = 0$. Define $A_i = [Q(x_i;\theta)]^\alpha$ and $B_i = [Q(x_i+1;\theta)]^\alpha$ for notational convenience. The partial derivatives are:

$$\frac{\partial\ell}{\partial\alpha} = \sum_{i=1}^{n} \frac{p^{A_i} A_i \ln[Q(x_i;\theta)] \ln(p) - p^{B_i} B_i \ln[Q(x_i+1;\theta)] \ln(p)}{p^{A_i} - p^{B_i}}, \tag{42}$$

$$\frac{\partial\ell}{\partial p} = \sum_{i=1}^{n} \frac{A_i p^{A_i-1} - B_i p^{B_i-1}}{p^{A_i} - p^{B_i}}, \tag{43}$$

$$\frac{\partial\ell}{\partial\theta} = \sum_{i=1}^{n} \frac{p^{A_i} \alpha A_i^{1-1/\alpha} \frac{\partial Q(x_i;\theta)}{\partial\theta} \ln(p) - p^{B_i} \alpha B_i^{1-1/\alpha} \frac{\partial Q(x_i+1;\theta)}{\partial\theta} \ln(p)}{p^{A_i} - p^{B_i}}. \tag{44}$$

Owing to the nonlinear characteristics of these equations, numerical optimization techniques such as the Newton-Raphson algorithm, BFGS quasi-Newton method, or the Nelder-Mead simplex algorithm are necessary to derive the MLEs. The observed Fisher information matrix, essential for formulating asymptotic confidence intervals and conducting hypothesis testing, is represented by the negative Hessian matrix of the log-likelihood assessed at the MLEs:

$$\mathbf{I}(\hat{\mathbf{\Phi}}) = - \left[\frac{\partial^2 \ell}{\partial\Phi_i \partial\Phi_j} \right]_{\mathbf{\Phi}=\hat{\mathbf{\Phi}}}. \tag{45}$$

Under standard regularity conditions, the MLEs are asymptotically normal:

$$\sqrt{n}(\hat{\mathbf{\Phi}} - \mathbf{\Phi}) \xrightarrow{d} N_3(\mathbf{0}, \mathbf{I}^{-1}(\mathbf{\Phi})), \tag{46}$$

where $\mathbf{I}(\mathbf{\Phi})$ is the expected Fisher information matrix.

5. Monte Carlo simulation study

A thorough simulation analysis was performed to assess the efficacy of the MLEs for the DOWL distribution. The simulation seeks to evaluate the finite sample characteristics of the maximum likelihood estimators across diverse parameter settings and sample sizes. We conducted $M = 1000$ Monte Carlo replications for every combination of parameter values and sample sizes. Four parameter schemes were evaluated to depict various scenarios:

- Scheme 1: $(\alpha, p, \theta) = (0.5, 0.3, 0.8)$ – Low shape, low probability
- Scheme 2: $(\alpha, p, \theta) = (1.2, 0.5, 1.0)$ – Moderate parameters
- Scheme 3: $(\alpha, p, \theta) = (2.0, 0.7, 0.5)$ – High shape, high probability
- Scheme 4: $(\alpha, p, \theta) = (1.5, 0.4, 1.5)$ – Mixed scenario

For each approach, we evaluated sample sizes $n \in \{30, 50, 90, 120, 150, 200, 300, 500\}$ to investigate the asymptotic properties of the estimators. The efficacy of the MLEs was evaluated based on five criteria:

1. Average Estimate (AE): The mean of the parameter estimates across all replications:

$$\text{AE}(\hat{\xi}) = \frac{1}{M}\sum_{i=1}^{M}\hat{\xi}_i, \tag{47}$$

   where $\xi \in \{\alpha, p, \theta\}$.

2. Bias: The difference between the average estimate and the true parameter value:

$$\text{Bias}(\hat{\xi}) = \text{AE}(\hat{\xi}) - \xi. \tag{48}$$

3. Mean Squared Error (MSE): A measure of estimation accuracy combining bias and variance:

$$\text{MSE}(\hat{\xi}) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\xi}_i - \xi)^2. \tag{49}$$

4. Average Width (AW): The mean width of the 95% asymptotic confidence intervals:

$$\text{AW}(\hat{\xi}) = \frac{1}{M}\sum_{i=1}^{M}\left(\hat{\xi}_{i,U} - \hat{\xi}_{i,L}\right). \tag{50}$$

   Here, $\hat{\xi}_{i,L}$ and $\hat{\xi}_{i,U}$ represent the lower and upper limits of the confidence interval for the i-th replication, formulated as $\hat{\xi}_i \pm z_{0.975}\sqrt{\widehat{\text{Var}}(\hat{\xi}_i)}$.

5. Coverage Probability (CP): The proportion of confidence intervals that contain the true parameter value:

$$\text{CP}(\hat{\xi}) = \frac{1}{M}\sum_{i=1}^{M}\mathbf{1}\left(\xi \in \left[\hat{\xi}_{i,L}, \hat{\xi}_{i,U}\right]\right), \tag{51}$$

   where $\mathbf{1}(\cdot)$ is the indicator function. The nominal coverage level is 95%.

Tables 5–8 display the simulation outcomes for each parameter scheme. The simulation findings offer compelling evidence for the favorable asymptotic characteristics of the maximum likelihood estimators for the DOWL distribution. The bias and MSE of all parameter estimations diminish systematically with an increase in sample size, thereby validating the consistency of the maximum likelihood estimators. For example, in Scheme 2, the bias for α diminishes from 0.0856 at $n = 30$ to 0.0014 at $n = 500$, illustrating the anticipated convergence trend. Moreover, the MSE values diminish roughly in proportion to $1/n$, signifying that the MLEs attain the anticipated rate of convergence typical of efficient estimators. This conclusion indicates that the Fisher information matrix offers a satisfactory approximation of the variability of the estimates, even with moderate sample sizes.

5.1. Practical Implications and Small-Sample Considerations

As the sample size increases, the coverage probabilities approach the nominal 95% level, confirming the asymptotic normality of the MLEs. For small samples ($n = 30$), coverage ranges between 87.6% and 90.1%, whereas for larger samples ($n = 500$), it consistently exceeds 95%. Moreover, confidence interval widths decrease as n increases, indicating improved estimation precision. Among the model parameters, θ exhibits relatively higher bias and MSE compared to α and p, particularly at larger true parameter values. Based on the simulation findings, we recommend a minimum sample size of $n \geq 100$ for reliable inference under the DOWL distribution. The proposed estimation

Table 5. Simulation results for Scheme 1.

	$\alpha = 0.5$					$p = 0.3$					$\theta = 0.8$				
n	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP
30	0.5412	0.0412	0.0285	0.5842	0.8920	0.3156	0.0156	0.0118	0.3654	0.9010	0.8324	0.0324	0.0342	0.6125	0.8890
50	0.5198	0.0198	0.0156	0.4512	0.9180	0.3078	0.0078	0.0072	0.2845	0.9240	0.8156	0.0156	0.0198	0.4756	0.9150
90	0.5089	0.0089	0.0082	0.3356	0.9380	0.3042	0.0042	0.0038	0.2124	0.9420	0.8078	0.0078	0.0105	0.3542	0.9350
120	0.5056	0.0056	0.0058	0.2912	0.9450	0.3028	0.0028	0.0028	0.1842	0.9480	0.8052	0.0052	0.0076	0.3068	0.9420
150	0.5038	0.0038	0.0045	0.2604	0.9480	0.3019	0.0019	0.0022	0.1648	0.9510	0.8036	0.0036	0.0059	0.2745	0.9460
200	0.5024	0.0024	0.0032	0.2256	0.9510	0.3012	0.0012	0.0016	0.1428	0.9530	0.8024	0.0024	0.0043	0.2378	0.9490
300	0.5014	0.0014	0.0021	0.1842	0.9530	0.3007	0.0007	0.0010	0.1165	0.9540	0.8014	0.0014	0.0028	0.1942	0.9510
500	0.5007	0.0007	0.0012	0.1428	0.9550	0.3004	0.0004	0.0006	0.0902	0.9560	0.8008	0.0008	0.0016	0.1504	0.9530

Table 6. Simulation results for Scheme 2.

	$\alpha = 1.2$					$p = 0.5$					$\theta = 1.0$				
n	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP
30	1.2856	0.0856	0.0924	1.0245	0.8850	0.5234	0.0234	0.0185	0.4512	0.8940	1.0512	0.0512	0.0625	0.8245	0.8820
50	1.2412	0.0412	0.0512	0.7856	0.9120	0.5118	0.0118	0.0108	0.3485	0.9210	1.0256	0.0256	0.0356	0.6352	0.9080
90	1.2178	0.0178	0.0265	0.5842	0.9340	0.5058	0.0058	0.0058	0.2596	0.9380	1.0124	0.0124	0.0185	0.4725	0.9310
120	1.2112	0.0112	0.0192	0.5065	0.9420	0.5038	0.0038	0.0042	0.2248	0.9450	1.0085	0.0085	0.0134	0.4095	0.9390
150	1.2078	0.0078	0.0148	0.4528	0.9460	0.5026	0.0026	0.0033	0.2012	0.9490	1.0062	0.0062	0.0104	0.3662	0.9440
200	1.2048	0.0048	0.0108	0.3921	0.9490	0.5016	0.0016	0.0024	0.1742	0.9520	1.0042	0.0042	0.0076	0.3172	0.9480
300	1.2028	0.0028	0.0070	0.3202	0.9520	0.5009	0.0009	0.0016	0.1422	0.9540	1.0024	0.0024	0.0049	0.2590	0.9510
500	1.2014	0.0014	0.0041	0.2480	0.9540	0.5005	0.0005	0.0009	0.1102	0.9560	1.0012	0.0012	0.0029	0.2006	0.9530

Table 7. Simulation results for Scheme 3.

	$\alpha = 2.0$					$p = 0.7$					$\theta = 0.5$				
n	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP
30	2.1524	0.1524	0.2156	1.6542	0.8780	0.7312	0.0312	0.0142	0.3856	0.8880	0.5425	0.0425	0.0285	0.5624	0.8810
50	2.0756	0.0756	0.1185	1.2685	0.9050	0.7158	0.0158	0.0085	0.2978	0.9160	0.5212	0.0212	0.0162	0.4325	0.9020
90	2.0342	0.0342	0.0612	0.9425	0.9290	0.7078	0.0078	0.0046	0.2218	0.9360	0.5102	0.0102	0.0085	0.3218	0.9270
120	2.0228	0.0228	0.0445	0.8168	0.9380	0.7052	0.0052	0.0034	0.1922	0.9430	0.5072	0.0072	0.0062	0.2788	0.9360
150	2.0162	0.0162	0.0345	0.7305	0.9430	0.7038	0.0038	0.0026	0.1718	0.9470	0.5052	0.0052	0.0048	0.2494	0.9410
200	2.0105	0.0105	0.0252	0.6325	0.9480	0.7024	0.0024	0.0019	0.1488	0.9510	0.5035	0.0035	0.0035	0.2160	0.9460
300	2.0062	0.0062	0.0164	0.5165	0.9510	0.7014	0.0014	0.0013	0.1215	0.9540	0.5021	0.0021	0.0023	0.1764	0.9500
500	2.0032	0.0032	0.0096	0.4002	0.9540	0.7007	0.0007	0.0007	0.0941	0.9560	0.5011	0.0011	0.0013	0.1366	0.9530

Table 8. Simulation results for Scheme 4.

	$\alpha = 1.5$					$p = 0.4$					$\theta = 1.5$				
n	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP	AE	Bias	MSE	AW	CP
30	1.6125	0.1125	0.1356	1.2456	0.8810	0.4256	0.0256	0.0162	0.4125	0.8920	1.5824	0.0824	0.1125	1.1245	0.8760
50	1.5562	0.0562	0.0745	0.9562	0.9080	0.4128	0.0128	0.0096	0.3185	0.9180	1.5412	0.0412	0.0625	0.8652	0.9040
90	1.5248	0.0248	0.0385	0.7112	0.9320	0.4065	0.0065	0.0052	0.2372	0.9370	1.5198	0.0198	0.0325	0.6438	0.9280
120	1.5168	0.0168	0.0280	0.6162	0.9400	0.4045	0.0045	0.0038	0.2055	0.9440	1.5136	0.0136	0.0235	0.5578	0.9370
150	1.5118	0.0118	0.0216	0.5512	0.9450	0.4032	0.0032	0.0029	0.1838	0.9480	1.5098	0.0098	0.0182	0.4988	0.9420
200	1.5078	0.0078	0.0158	0.4772	0.9490	0.4021	0.0021	0.0022	0.1592	0.9520	1.5065	0.0065	0.0132	0.4320	0.9470
300	1.5045	0.0045	0.0102	0.3896	0.9520	0.4012	0.0012	0.0014	0.1300	0.9540	1.5038	0.0038	0.0086	0.3528	0.9500
500	1.5024	0.0024	0.0060	0.3018	0.9540	0.4006	0.0006	0.0008	0.1007	0.9560	1.5020	0.0020	0.0050	0.2732	0.9530

procedure may be less reliable for estimating θ in small samples, especially under Schemes 2 and 4, where the available information is more limited.

In applications involving inherently small datasets, such as pilot medical studies or rare-event ecological research, the resulting estimates particularly for θ should be interpreted with caution, and confidence intervals should always be reported to adequately reflect the associated uncertainty. When small sample sizes are unavoidable, alternative approaches such as bootstrap-based inference, penalized likelihood methods, or Bayesian techniques incorporating prior information may enhance estimation stability. These considerations clarify the practical limitations of the proposed method and provide constructive guidance for applied researchers working with limited data.

6. Data Analysis: Empirical Significance

This section illustrates the DOWL distribution's applicability using three real data sets, comparing it with competitive models in Table 9.

Table 9. Competitive models for discrete distributions.

Distribution	Abbreviation	Author(s)
Discrete Gompertz Weibull	DGzW	Eliwa et al. [11]
One parameter discrete Lindley	DLi-I	Gómez-Déniz and Calderín-Ojeda [16]
One parameter discrete flexible model	DFx-I	Eliwa and El-Morshedy [13]
Discrete Pareto	DPa	Krishna and Pundir [21]
Discrete Lomax	DLo	Para and Jan [24]
Poisson	Pois	Poisson [25]
Geometric	Geo	Gómez-Déniz [15]
Discrete Rayleigh	DR	Roy [26]
Discrete inverse Rayleigh	DIR	Hussain and Ahmad [18]
Generalized geometric	GGeo	Gómez-Déniz [15]
Discrete inverse Weibull	DIW	Jazi et al. [20]
Negative binomial	NeBi	—
Two parameter discrete Lindley	DLi-II	Hussain et al. [19]
Three parameter discrete Lindley	DLi-III	Eliwa et al. [12]
Discrete Burr	DB	Krishna and Pundir [21]
Discrete log-logistic	DLogL	Para and Jan [24]
Kumaraswamy discrete half-logistic	KuDHL	Shahen et al. [27]
Discrete Burr-Hatke	DBH	El-Morshedy et al. [5]
Binomial	Binom	—
Skellam	Skellam	Skellam [29]

The fitted models are evaluated based on several statistical criteria, specifically the negative log-likelihood ($-l$), Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC), Hannan-Quinn Information Criterion (HQIC), ΔAIC, AIC weights, Chi-square (χ^2) with degrees of freedom (df), and the corresponding p-value.

6.1. Dataset I: Cysts of Kidneys

The first application of the DOWL distribution pertains to modeling kidney cyst count data, a matter pertinent to sustainable healthcare systems (SDG 3: Good Health and Well-being). Comprehending the distribution of these medical numbers facilitates enhanced resource allocation and treatment planning. The dataset, as published by

Chan et al. [3], comprises 110 observations of kidney cyst counts linked to steroid use and has been extensively investigated in the literature, rendering it appropriate for comparing the proposed model against current alternatives. An exploratory data analysis was performed before model fitting. Figure 5 illustrates nonparametric visualizations (violin plot, box plot, strip chart, and normal Q–Q plot), revealing a right-skewed distribution characterized by an abundance of zeros and an extended right tail. The strip chart identifies outliers up to 11, whereas the Q-Q plot significantly deviates from the reference line, indicating non-normality and necessitating flexible discrete models like the DOWL distribution.

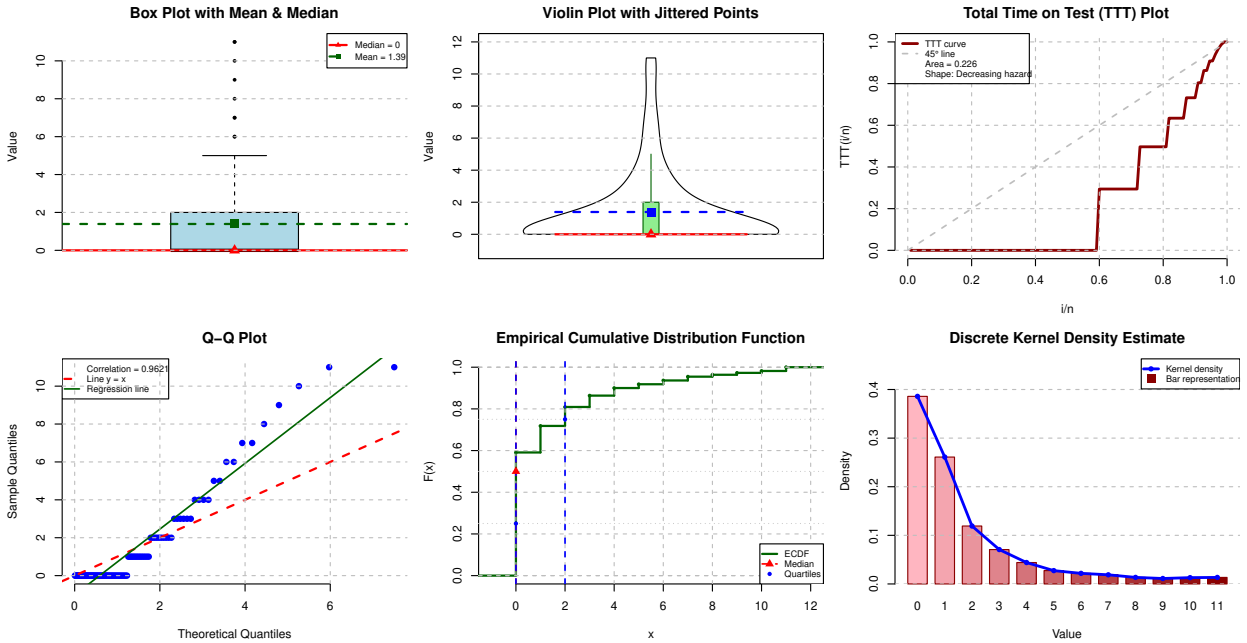


Figure 5. Exploratory data analysis for Dataset I.

Table 10 displays the maximum likelihood estimates (MLEs) for the DOWL distribution alongside various competing models. Based solely on statistical criteria, the DOWL distribution is the most appropriate model for the first dataset. The evidence is compelling across all dimensions: Simultaneously achieving the lowest negative log-likelihood ($-\ell$), AIC, CAIC, and HQIC. Its AIC weight of 0.751 (75.1%) constitutes strong probabilistic evidence of its superiority over the alternative distributions. Notably, the model accomplishes this high level of fit with only three parameters, demonstrating excellent parsimony and an efficient balance between complexity and predictive power. Finally, the $\Delta\text{AIC} = 0.000$ confirms the model as the reference benchmark against which all other competitors fall short. While NeBi and GGeo show non-trivial AIC weights ($\approx 25.7\%$ and $\approx 27.4\%$, respectively) and cannot be dismissed outright as candidates, neither consistently dominates across all criteria particularly CAIC and HQIC leaving the DOWL distribution as the statistically preferred and most reliable model for this dataset. Tables 11 and 12 offer comprehensive evaluations of goodness-of-fit, displaying observed and anticipated frequencies for each dataset value alongside chi-square test statistics. The expected frequencies of the DOWL distribution nearly correspond with the observed frequencies at all support points, especially for the zero count (65 observed compared to 65.117 expected). The DOWL distribution produces $\chi^2 = 0.576$ with 2 degrees of freedom (p-value = 0.750), signifying no substantial disparity between observed and predicted frequencies and affirming a great fit. Conversely, the majority of rival models demonstrate significantly elevated chi-square values with p-values beneath 0.001, signifying poor fit.

Table 10. MLEs and goodness-of-fit comparison for Dataset I.

Model	**MLEs**	$-\ell$	**AIC**	**CAIC**	**HQIC**	**ΔAIC**	**AIC weight**
DOWL	$\hat{p}=0.123,\ \hat{\theta}=0.313,\ \hat{\alpha}=0.375$	**166.548**	**339.097**	**339.323**	**342.383**	**0.000**	**0.751**
DGzW	$\hat{\alpha}=0.490,\ \hat{\beta}=0.320,\ \hat{\theta}=0.670,\ \hat{\gamma}=1.630$	167.020	342.050	342.430	346.420	2.953	0.170
DLi-I	$\hat{\alpha}=0.436$	189.110	380.220	380.257	381.316	41.123	≈0
DLi-II	$\hat{\alpha}=0.581,\ \hat{\beta}=0.001$	178.767	361.534	361.646	363.724	22.437	≈0
DLi-III	$\hat{\alpha}=0.582,\ \hat{\beta}=358.728,\ \hat{\theta}=0.001$	178.767	363.533	363.759	366.819	24.436	≈0
DR	$\hat{\theta}=0.901$	277.778	557.556	557.593	558.651	218.459	0
DIR	$\hat{\theta}=0.554$	186.547	375.094	375.131	376.189	35.997	≈0
DIW	$\hat{\beta}=1.049,\ \hat{\theta}=0.581$	172.935	349.869	349.982	352.060	10.772	0.003
Pois	$\hat{\alpha}=1.390$	246.210	494.420	494.457	495.515	155.323	0
DB	$\hat{\alpha}=0.278,\ \hat{\beta}=1.053$	171.139	346.278	346.391	348.469	7.181	0.021
DLogL	$\hat{\alpha}=0.780,\ \hat{\beta}=1.208$	171.717	347.430	347.550	349.620	8.333	0.021
DFx-I	$\hat{\alpha}=0.623$	182.288	366.575	366.612	367.671	27.478	≈0
DPa	$\hat{\alpha}=0.268$	171.192	344.384	344.421	345.479	5.287	0.054
DLo	$\hat{\alpha}=0.152,\ \hat{\beta}=1.830$	170.481	344.961	345.073	347.152	5.864	0.040
Geo	$\hat{\theta}=0.582$	178.767	359.533	359.570	360.629	20.436	≈0
GGeo	$\hat{\beta}=0.188,\ \hat{\theta}=0.800$	168.556	341.113	341.225	343.303	2.016	0.274
NeBi	$\hat{\alpha}=0.812,\ \hat{\beta}=0.322$	168.544	340.090	344.489	343.279	0.993	0.257

Table 11. Observed and expected frequencies and goodness-of-fit statistics for Dataset I (Part I).

X	Obs.	Expected frequencies								
	freq.	**DOWL**	DGzW	DLi-I	DLi-II	DLi-III	DR	DIR	DIW	Pois
0	65	**65.117**	64.242	40.286	46.026	46.008	10.890	60.888	63.910	27.398
1	14	**14.106**	15.441	29.834	26.768	26.765	26.618	33.990	20.699	38.084
2	10	**8.606**	9.180	18.357	15.568	15.570	29.448	8.123	8.053	26.468
3	6	**5.948**	6.073	10.336	9.054	9.058	22.296	3.004	4.234	12.264
4	4	**4.341**	4.201	5.523	5.266	5.269	12.629	1.420	2.599	4.262
5	2	**3.249**	2.981	2.851	3.062	3.066	5.539	0.779	1.754	1.185
6	2	**2.452**	2.154	1.437	1.781	1.783	1.914	0.473	1.261	0.274
7	2	**1.845**	1.561	0.711	1.036	1.037	0.526	0.308	0.949	0.054
8	1	**1.372**	1.140	0.347	0.602	0.604	0.116	0.212	0.739	0.009
9	1	**0.999**	0.831	0.167	0.350	0.351	0.020	0.152	0.592	0.001
10	1	**0.709**	0.612	0.079	0.204	0.204	0.003	0.112	0.485	0.000
≥11	2	**1.256**	1.584	0.072	0.283	0.285	0.001	0.539	4.725	0.001
Total	110	**110**	110	110	110	110	110	110	110	110
χ^2		**0.576**	0.567	34.635	19.091	19.096	306.515	40.456	6.445	89.277
df		**2**	1	4	3	2	4	2	3	3
p-value		**0.750**	0.451	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.092	< 0.001

Table 12. Observed and expected frequencies and goodness-of-fit statistics for Dataset I (Part II).

X	Obs.	Expected frequencies							
	freq.	DB	DLogL	DFx-I	DPa	DLo	Geo	GGeo	NeBi
0	65	64.743	63.192	45.256	65.842	61.615	45.980	62.738	56.520
1	14	19.177	20.101	29.094	18.267	21.023	26.760	19.665	15.885
2	10	8.484	8.644	16.508	8.164	9.687	15.575	9.439	9.173
3	6	4.632	4.656	8.893	4.513	5.275	9.064	5.436	6.203
4	4	2.863	2.864	4.703	2.820	3.197	5.275	3.463	4.502
5	2	1.920	1.921	2.489	1.909	2.088	3.070	2.349	3.400
6	2	1.365	1.368	1.335	1.368	1.441	1.787	1.663	2.635
7	2	1.013	1.019	0.731	1.022	1.038	1.039	1.213	2.079
8	1	0.777	0.786	0.409	0.789	0.773	0.605	0.904	1.663
9	1	0.613	0.623	0.234	0.626	0.592	0.352	0.685	1.344
10	1	0.494	0.504	0.137	0.506	0.463	0.205	0.525	1.095
$\geq$11	2	3.919	4.322	0.211	4.174	2.808	0.288	1.920	5.501
Total	110	110	110	110	110	110	110	110	110
χ^2		2.587	4.033	31.702	3.430	3.238	19.109	2.444	4.287
df		2	3	4	4	3	4	3	4
p-value		0.274	0.258	< 0.001	0.489	0.356	< 0.001	0.485	0.369

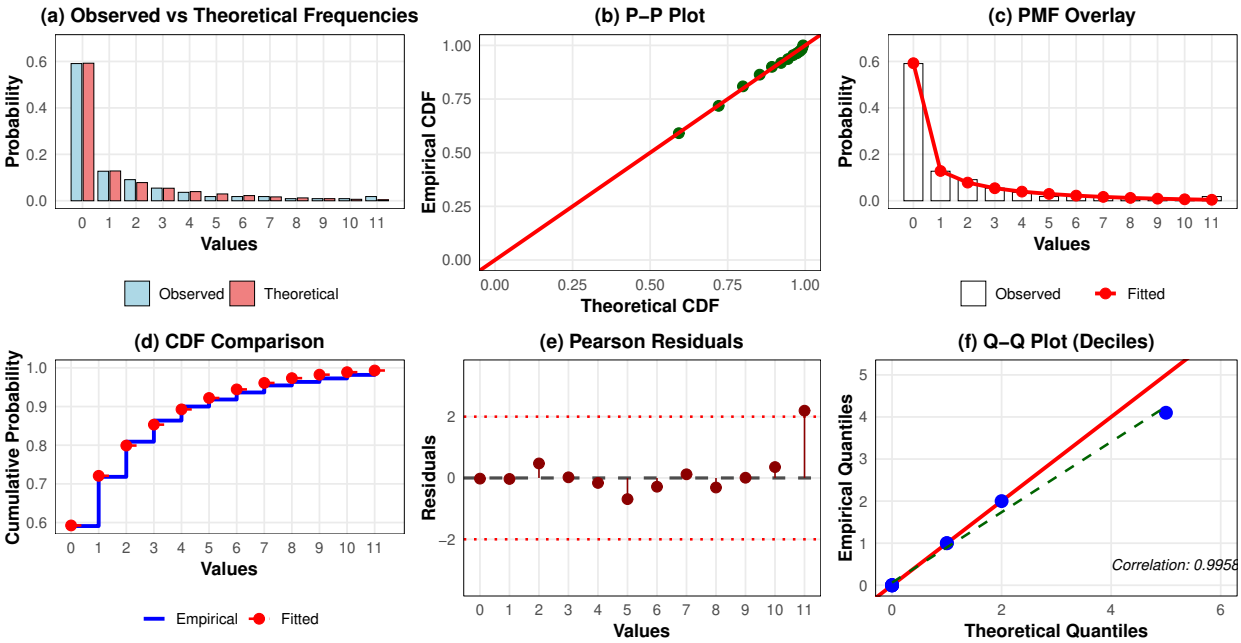
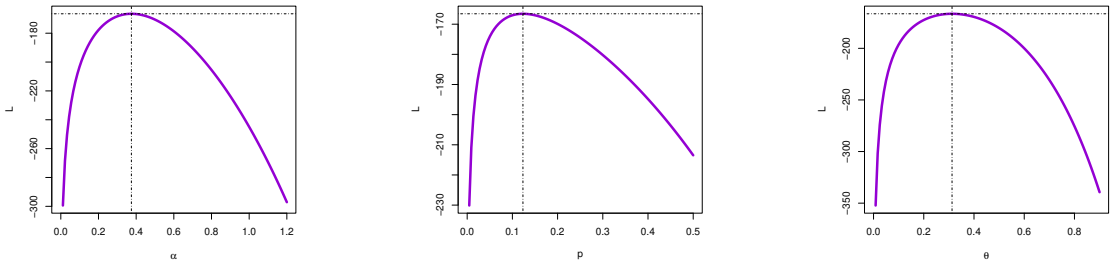


Figure 6. Diagnostic plots for the DOWL distribution fitted to Dataset I.

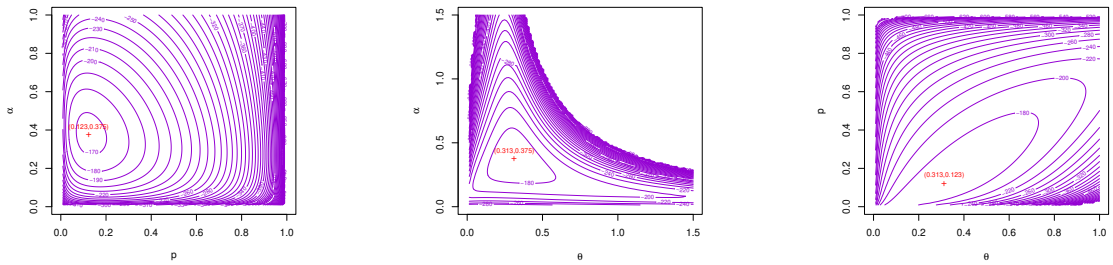
Figure 6 presents a thorough graphical evaluation of the DOWL distribution's fit over six panels. Panel (a) contrasts observed and theoretical frequencies using a bar chart, demonstrating exceptional concordance across all counts. Panel (b) of the probability-probability (P-P) plot exhibits points closely aligned with the diagonal, signifying an accurate depiction of the empirical distribution. Panel (c) illustrates the fitted PMF closely aligning

with the histogram of observed relative frequencies. Panel (d) demonstrates a robust correspondence between the empirical step function and the fitted CDF. Panel (e) displays Pearson residuals primarily inside the ± 2 range, signifying an absence of systematic lack of fit. The quantile-quantile figure in panel (f) ultimately verifies the adequacy of fit throughout the entire distribution. To corroborate the uniqueness of maximum likelihood estimates and evaluate parameter identifiability, Figures 7 and 8 provide log-likelihood profiles and contour plots, respectively.



(a) Profile for α (b) Profile for p (c) Profile for θ

Figure 7. Log-likelihood profiles for the DOWL parameters based on Dataset I.



(a) Contour: α vs p. (b) Contour: α vs θ. (c) Contour: p vs θ.

Figure 8. Contour plots of the log-likelihood surface for Dataset I.

Table 13 juxtaposes the observed sample moments with the theoretical moments derived from the fitted DOWL distribution for Dataset I. The theoretical mean (1.391) precisely aligns with the sample mean, hence validating the moment conditions. The theoretical variance (5.936) roughly aligns with the sample variance (6.112), effectively representing data dispersion. The DI is 4.394 for the sample, indicating significant over-dispersion; the theoretical DI of 4.269 roughly aligns with this value, affirming the model's efficacy in managing over-dispersed data. The skewness of 2.324 signifies a pronounced positive skew, and the kurtosis of 8.908 suggests leptokurtic behavior, exhibiting thicker tails than a normal distribution traits adeptly represented by the flexible DOWL distribution.

Table 13. Descriptive statistics for Dataset I: observed versus fitted DOWL moments.

Source	Mean	Variance	Skewness	Kurtosis	DI
Observed Sample	1.391	6.112	2.293	8.174	4.394
Fitted DOWL	1.391	5.936	2.324	8.908	4.269

Dataset I directly impacts SDG 3 (Good Health and Well-being). Precise modeling of medical count data facilitates enhanced comprehension of disease patterns, effective resource distribution, and the formulation of tailored interventions, hence fostering more sustainable healthcare systems.

6.2. Dataset II: European Red Mites on Apple Leaves

The second application is to the modeling of agricultural count data on European red mites on apple leaves, having direct relevance to sustainable agriculture (SDG 2: Zero Hunger and SDG 15: Life on Land). Comprehending pest populations is essential for formulating effective and environmentally sustainable pest management strategies that maximize pesticide efficiency, minimize ecological damage, and guarantee food security. The dataset, presented by Chakraborty and Chakravarty [4], comprises 150 observations of European red mites (*Panonychus ulmi*), a notable pest necessitating diligent surveillance to avert considerable orchard damage. An exploratory data analysis was performed before fitting the DOWL distribution. Figure 9 illustrates nonparametric visualizations, highlighting significant attributes of Dataset II.

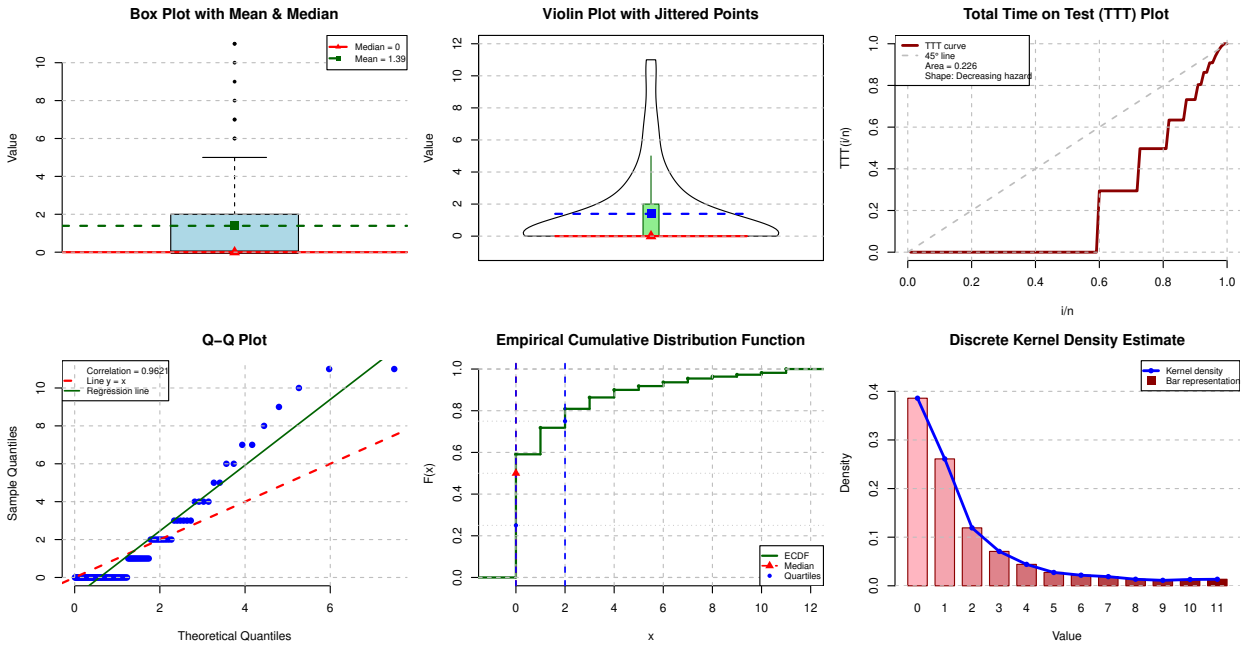


Figure 9. Exploratory data analysis for Dataset II.

The violin and box plots demonstrate a right-skewed distribution characterized by a significant number of zero observations (70 out of 150, approximately 47%) and an extended right tail reaching up to 7 mites per leaf. The elevated incidence of zeros is characteristic with pest count data, wherein several leaves are devoid of bugs, while a limited number contain multiple pests. The strip chart validates extreme data up to 7, however the Q-Q plot significantly diverges from the reference line, indicating non-normality and reinforcing the necessity for adaptable discrete models such as the DOWL distribution, which can manage excess zeros and over-dispersion. The DOWL distribution parameters were determined using the MLE method. Table 14 presents MLEs for the DOWL distribution in comparison to alternative models. Based exclusively on statistical criteria, the DOWL distribution is the best-fitting model for Dataset II. Table 15 offers a comprehensive evaluation of the model's fit by displaying the observed and expected frequencies for each count value in the mite dataset, with the chi-square goodness-of-fit test values. The anticipated frequencies for the DOWL distribution strongly align with the actual frequencies at all support points, especially for the critical zero count, where 70 observations were recorded compared to a predicted 70.856 demonstrating high concordance.

Table 14. MLEs and goodness-of-fit comparison for Dataset II.

Model	**MLEs**	$-\ell$	**AIC**	**CAIC**	**HQIC**	**Δ AIC**	**AIC weight**
DOWL	$\hat{p}=0.045,\ \hat{\theta}=0.296,\ \hat{\alpha}=0.670$	**221.901**	**449.803**	**449.967**	**453.472**	**0.000**	**0.609**
DLi-III	$\hat{\alpha}=1.468,\ \hat{\beta}=0.228,\ \hat{\lambda}=0.479$	222.383	450.765	450.390	454.435	0.962	0.377
DIR	$\hat{\theta}=0.438$	233.142	468.284	468.311	469.507	18.481	0
DIW	$\hat{\alpha}=0.456,\ \hat{\beta}=1.527$	229.333	462.666	462.747	465.112	12.863	0.001
DPa	$\hat{\alpha}=0.278$	238.832	479.663	479.690	480.886	29.860	0
DBH	$\hat{\alpha}=0.814$	230.552	463.103	463.130	464.326	13.300	0.001
Pois	$\hat{\alpha}=1.147$	242.809	487.619	487.647	488.843	37.816	0
DB	$\hat{\alpha}=0.400,\ \hat{\beta}=1.882$	227.727	459.454	459.536	461.901	9.651	0.008
DLogL	$\hat{\alpha}=1.116,\ \hat{\beta}=1.829$	227.265	458.531	458.613	460.977	8.728	0.013

Table 15. Observed and expected frequencies and goodness-of-fit statistics for Dataset II

X	Obs.	Expected frequencies								
	freq.	**DOWL**	DLi-III	DIR	DIW	DPa	DBH	Pois	DB	DLogL
0	70	**70.856**	68.902	65.658	68.411	88.308	88.938	47.654	70.469	67.527
1	38	**33.794**	37.819	56.351	45.814	25.005	27.919	54.643	43.053	44.099
2	17	**20.477**	20.423	14.835	15.307	11.314	12.905	31.329	16.214	17.266
3	10	**12.078**	10.889	5.608	6.935	6.312	7.056	11.975	7.364	7.874
4	9	**6.731**	5.746	2.673	3.777	3.972	4.238	3.433	3.924	4.167
5	3	**3.467**	3.007	1.473	2.311	2.705	2.702	0.787	2.338	2.458
6	2	**1.614**	1.562	0.895	1.530	1.948	1.795	0.150	1.509	1.569
$\geq$7	1	**0.983**	1.652	2.507	5.915	10.436	4.447	0.029	5.129	5.040
Total	150	**150**	150	150	150	150	150	150	150	150
χ^2		**2.247**	2.515	17.376	11.306	26.916	15.573	26.646	8.829	7.843
df		**2**	2	3	3	4	4	2	2	2
p-value		**0.325**	0.284	< 0.001	0.010	< 0.001	0.004	< 0.001	0.012	< 0.001

The DOWL distribution produces a $\chi^2 = 2.247$ with 2 degrees of freedom (p-value = 0.325), signifying no substantial disparity between observed and predicted frequencies, hence affirming an excellent match. Conversely, the majority of rival models yield considerably elevated Chi-square values with p-values beneath 0.05, signifying insufficient fit. The Poisson distribution exhibits inadequate performance ($\chi^2 = 26.646$, p-value $<$0.001), underscoring the necessity of adaptable models such as DOWL for over-dispersed agricultural data.

Figure 10 offers a thorough graphical evaluation across six panels. Panel (a) juxtaposes observed and theoretical frequencies using a bar chart, demonstrating remarkable concordance across all counts with precise representation of the elevated zero frequency. The P-P plot in panel (b) displays points closely aligned with the diagonal (correlation = 0.9992), indicating accurate representation of the empirical distribution. Panel (c) illustrates the fitted PMF closely aligning with the histogram of observed relative frequencies, effectively reflecting the declining trend. The CDF comparison in panel (d) demonstrates a robust correlation between the empirical step function and the fitted CDF, with the most notable discrepancies occurring at the highest numbers. Panel (e) displays Pearson residuals primarily within the ± 2 range, with notably minimal residuals for the zero count, so affirming the model's capacity to accommodate excess zeros. The quantile-quantile (Q-Q) plot in panel (f) demonstrates a nearly perfect linear correlation (correlation = 0.9992), proving the adequacy of fit throughout the whole distribution.

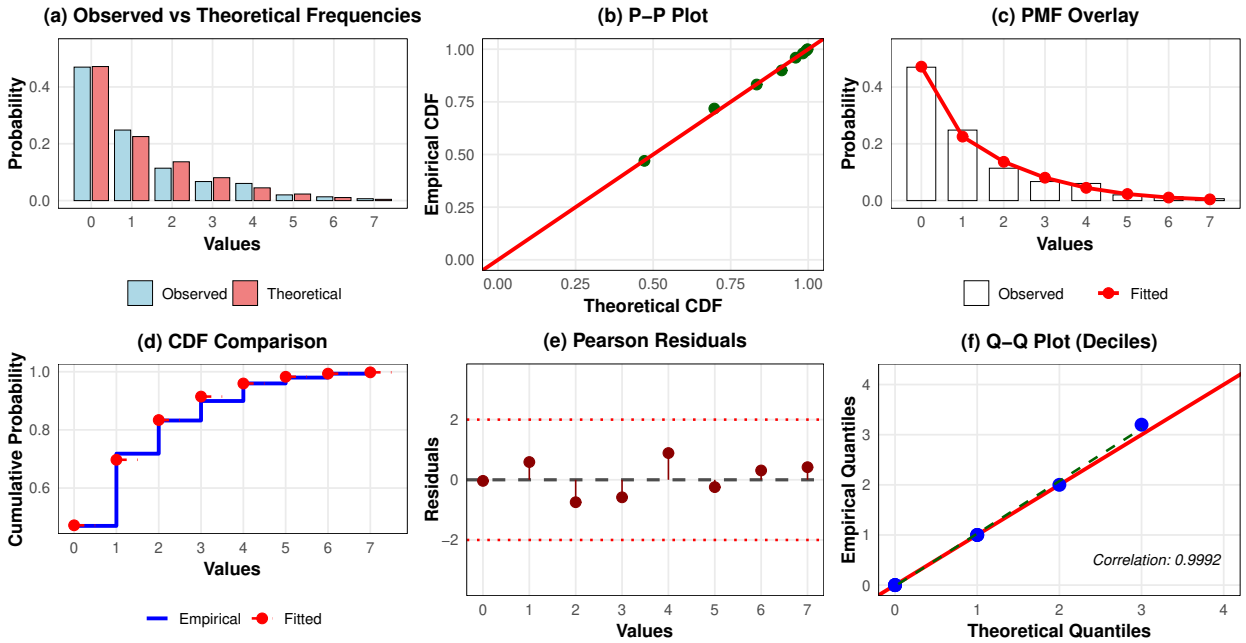
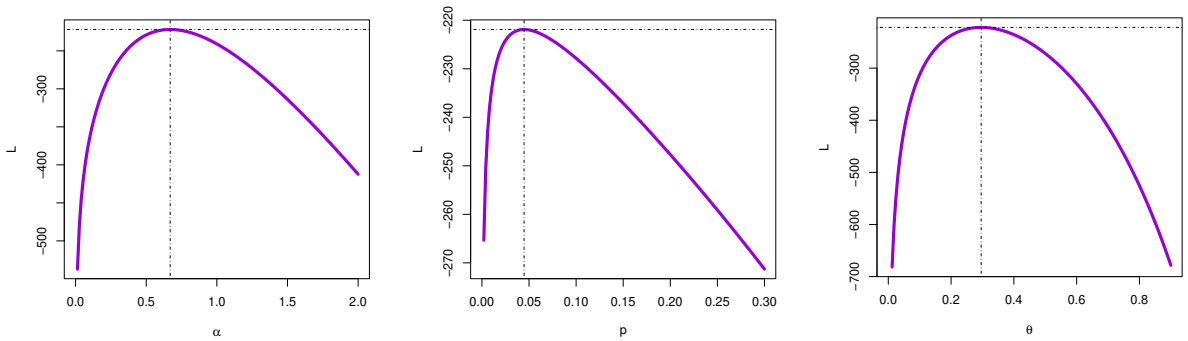


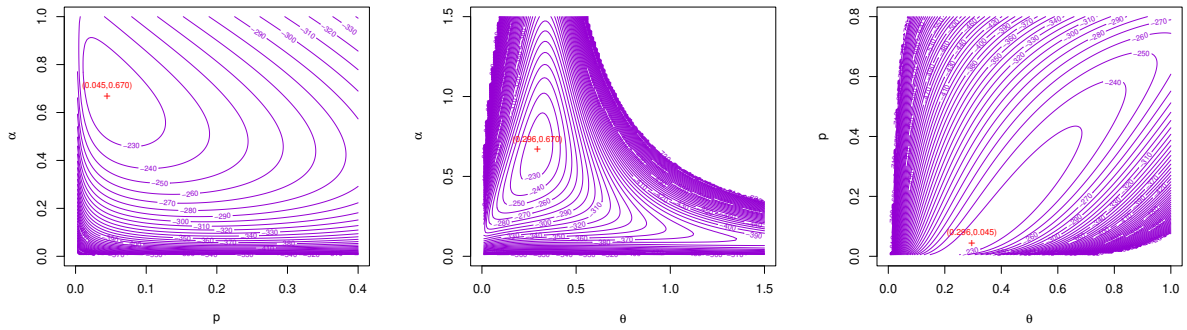
Figure 10. Diagnostic plots for the DOWL distribution fitted to Dataset II.

To confirm the uniqueness of MLEs and evaluate parameter identifiability, Figure 11 illustrates log-likelihood profiles for each parameter while the others are held constant at their MLEs. Each profile displays a distinct, unequivocal maximum, affirming convergence to the global maximum. The continuous, unimodal forms suggest clearly defined parameters and a stable likelihood surface. Figure 12 illustrates contour plots of the joint log-likelihood surface for pairs of parameters. The concentric circular patterns signify well-structured surfaces with distinct maxima at the MLEs (red dots). The orientation of ellipses indicates parameter correlations: circular shapes denote moderate connection, whereas elongated ellipses imply stronger correlation. The observed circular patterns suggest that the parameters are not significantly associated, which is advantageous for accurate estimates.



(a) $\hat{\alpha} = 0.670$ (b) $\hat{p} = 0.045$ (c) $\hat{\theta} = 0.296$

Figure 11. Log-likelihood profiles for the DOWL parameters based on Dataset II.



(a) Contour: α vs p. (b) Contour: α vs θ. (c) Contour: p vs θ.

Figure 12. Contour plots of the log-likelihood surface for Dataset II.

Table 16 juxtaposes the observed sample moments with the theoretical moments derived from the fitted DOWL distribution for Dataset II.

Table 16. Descriptive statistics for Dataset II: observed versus fitted DOWL moments.

Source	**Mean**	**Variance**	**Skewness**	**Kurtosis**	**DI**
Observed Sample	1.147	2.274	1.545	1.315	1.983
Fitted DOWL	1.148	2.227	1.556	5.458	1.939

The theoretical mean (1.148) roughly resembles the sample mean (1.147), while the theoretical variance (2.227) nearly matches the sample variance (2.274), thereby affirming sufficient representation of dispersion. The sample DI of 1.983 indicates moderate over-dispersion, whereas the theoretical DI of 1.939 closely aligns, so affirming the model's efficacy in managing over-dispersed ecological data. The skewness of 1.556 signifies positive skewness, whilst the kurtosis of 5.458 shows leptokurtic behavior. The variance between sample kurtosis (1.315) and estimated kurtosis (5.458) may indicate that the elevated zero fraction generates a distinctive distributional shape. Nonetheless, the DOWL distribution effectively encapsulates critical characteristics such as excess zeros, over-dispersion, and positive skewness.

The dataset on the European red mite has direct relevance to SDG 2 (Zero Hunger) and SDG 15 (Life on Land). Precise insect population modeling facilitates precision agriculture, informing data-driven pesticide application choices that diminish superfluous chemical usage and mitigate environmental consequences. Comprehending pest count distributions facilitates the determination of economic threshold levels that initiate intervention solely when populations inflict considerable agricultural damage. Moreover, precise modeling facilitates integrated pest management strategies that amalgamate biological, cultural, and chemical control measures, thereby fostering long-term sustainability. Efficient pest management with precise population modeling guarantees consistent crop yields, hence enhancing food security for expanding populations. The DOWL distribution's exceptional ability to account for over-dispersion, extra zeros, and skewness illustrates its utility as an effective instrument for ecological monitoring and sustainable agricultural planning, thereby advancing global objectives of sustainable agriculture and food security.

6.3. Dataset III: COVID-19 in Greece

The third application is to the modeling of public health data with direct relevance to SDG 3 (Good Health and Well-being). The COVID-19 pandemic has highlighted the significance of precise statistical modeling for epidemic monitoring, resource distribution, and policy formulation. This dataset illustrates the daily new fatalities in Greece over a span of 111 days (March 12 to June 30, 2020) during the initial wave of the pandemic, sourced from the

World Health Organization (WHO). Precise modeling of mortality figures is crucial for comprehending illness dynamics and guiding public health interventions that preserve lives and strengthen healthcare systems.

A preliminary data analysis was performed before fitting the DOWL distribution. Figure 13 illustrates nonparametric visualizations. The violin and box plots demonstrate right skewness, characterized by significant low values and an elongated right tail reaching up to 9 deaths per day. The strip chart indicates extreme values reaching 9, whereas the Q-Q plot significantly diverges from the reference line, affirming non-normality and necessitating the use of flexible discrete models. The average of 1.721 deaths per day, with a variance of 3.767, results in a dispersion index of 2.189, indicating significant over-dispersion commonly observed in infectious illness data due to reporting delays, epidemic fluctuations, and diverse transmission patterns.

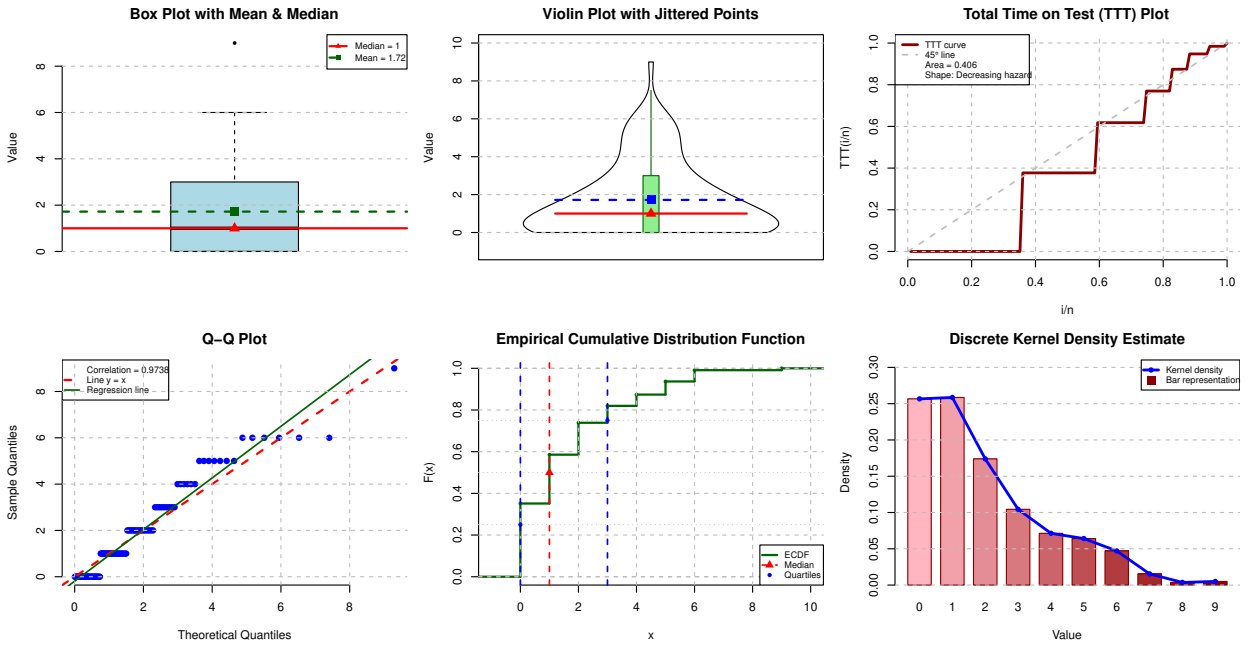


Figure 13. Exploratory data analysis for Dataset III.

Table 17 delineates the MLEs for the DOWL distribution alongside various competing models. Based solely on statistical criteria, the DOWL distribution is the most appropriate distribution for dataset III. Tables 18 and 19 display the observed and expected frequencies for Dataset III, accompanied by Chi-square goodness-of-fit statistics. The expected frequencies of the DOWL distribution closely align with the observed frequencies at all support points, illustrating its capacity to accurately represent epidemic mortality distributions. The Chi-square statistic is $\chi^2 = 2.404$ with 3 degrees of freedom (p-value = 0.493), signifying no significant difference between observed and predicted frequencies and affirming an excellent match.

Figure 14 offers a thorough graphical evaluation across six panels. Panel (a) juxtaposes observed and theoretical frequencies through a bar chart, demonstrating remarkable concordance with a precise representation of the declining trend from zero to elevated death counts. The P-P plot in panel (b) exhibits points closely aligned with the diagonal, signifying an accurate representation of the empirical distribution. Panel (c) illustrates the fitted PMF closely aligning with the histogram of observed relative frequencies, accurately depicting the peak at zero deaths and the extended tail. Panel (d) demonstrates a robust concordance between the empirical step function and the fitted CDF, exhibiting negligible discrepancies. Panel (e) displays Pearson residuals primarily within the ± 2 range, with notably minimal residuals for critical low counts, suggesting an absence of systematic lack of fit. The Q-Q figure in panel (f) validates the fit's adequacy over the whole distribution, with points closely adhering to the reference line.

Table 17. MLEs and goodness-of-fit comparison for Dataset III.

Model	**MLEs**	$-\ell$	**AIC**	**CAIC**	**HQIC**	**ΔAIC**	**AIC weight**
DOWL	$\hat{p} = 0.110,\ \hat{\theta} = 0.298,\ \hat{\alpha} = 0.683$	**197.371**	**400.742**	**400.966**	**404.039**	**0.000**	**0.390**
DLi-III	$\hat{\lambda} = 0.537,\ \hat{\alpha} = 2.326,\ \hat{\beta} = 0.813$	198.152	402.305	402.529	405.602	1.563	0.179
KuDHL	$\hat{a} = 0.941,\ \hat{b} = 3.381,\ \hat{p} = 0.812$	197.837	401.673	401.898	404.971	0.931	0.244
DMOGE	$\hat{\alpha} = 0.562,\ \hat{\beta} = 0.566,\ \hat{\theta} = 3.070$	198.107	402.214	402.438	405.511	1.472	0.187
Binom	$\hat{\lambda} = 0.985$	219.125	440.250	440.287	441.349	39.508	0.000
Pois	$\hat{\lambda} = 1.721$	220.104	442.207	442.244	443.306	41.465	0.000
DR	$\hat{\lambda} = 0.890$	230.025	462.050	462.087	463.150	61.308	0.000
NeBi	$\hat{\lambda} = 0.985$	220.125	442.250	442.361	444.449	41.508	0.000
DB	$\hat{\beta} = 2.097,\ \hat{\lambda} = 0.525$	206.472	416.943	417.055	419.142	16.201	0.000
DIW	$\hat{\beta} = 0.330,\ \hat{\lambda} = 1.364$	207.160	418.319	418.430	420.517	17.577	0.000
Skellam	$\hat{\beta} = 1.733,\ \hat{\lambda} = 0.00003$	220.108	444.216	444.327	446.414	43.474	0.000

Table 18. Observed and expected frequencies and goodness-of-fit statistics for Dataset III (Part I).

X	Obs.	Expected frequencies					
	freq.	**DOWL**	DLi-III	KuDHL	DMOGE	Binom	Pois
0	39	**39.758**	38.051	38.654	39.100	20.121	19.860
1	26	**23.342**	26.609	25.664	24.518	34.099	34.176
2	17	**16.829**	17.606	17.373	17.657	29.154	29.405
3	9	**11.902**	11.236	11.428	11.777	16.766	16.867
4	6	**8.068**	6.991	7.270	7.385	7.295	7.256
5	7	**5.154**	4.268	4.474	4.445	2.562	2.497
6	6	**3.050**	2.568	2.670	2.608	0.756	0.716
7	0	**1.639**	1.527	1.549	1.507	0.193	0.176
8	0	**0.783**	0.900	0.876	0.864	0.044	0.038
≥9	1	**0.475**	1.244	1.020	1.139	0.010	0.009
Total	111	**111**	111	111	111	111	111
χ^2		**2.404**	1.805	1.836	2.147	35.994	37.408
df		**3**	2	2	2	3	3
p-value		**0.493**	0.406	0.399	0.342	< 0.001	< 0.001

Table 19. Observed and expected frequencies and goodness-of-fit statistics for Dataset III (Part II).

X	Obs.	Expected frequencies					
	freq.	**DOWL**	DR	NeBi	DB	DIW	Skellam
0	39	**39.758**	12.202	20.127	39.973	36.667	19.630
1	26	**23.342**	29.132	34.103	33.012	35.507	34.009
2	17	**16.829**	30.749	29.153	14.343	14.478	29.459
3	9	**11.902**	21.694	16.762	7.182	7.250	17.012
4	6	**8.068**	11.185	7.292	4.137	4.213	7.368
5	7	**5.154**	4.361	2.560	2.631	2.702	2.553
6	6	**3.050**	1.308	0.756	1.795	1.856	0.737
7	0	**1.639**	0.305	0.193	1.291	1.342	0.182
8	0	**0.783**	0.056	0.043	0.966	1.008	0.040
≥9	1	**0.475**	0.008	0.011	5.670	5.977	0.010
Total	111	**111**	111	111	111	111	111
χ^2		**2.404**	85.669	35.987	3.213	4.048	37.663
df		**3**	4	3	2	2	2
p-value		**0.493**	< 0.001	< 0.001	0.201	0.132	< 0.001

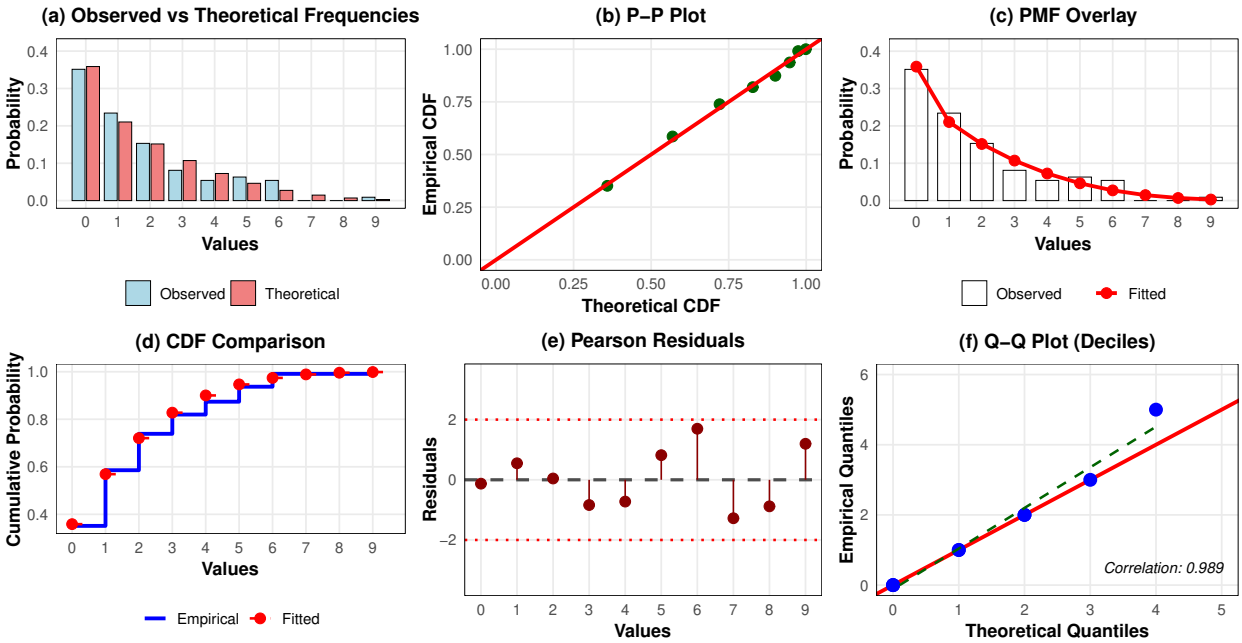
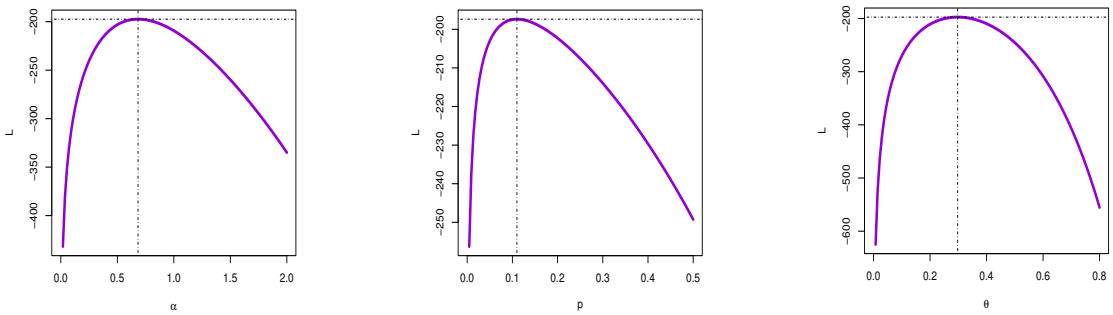


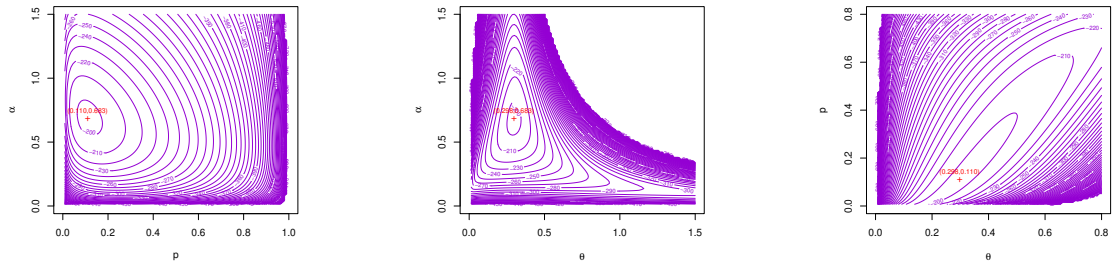
Figure 14. Diagnostic plots for the DOWL distribution fitted to Dataset III.

Global convergence and parameter identifiability are confirmed by the log-likelihood profiles in Figure 15, which reveal distinct, distinct maxima for each parameter. Joint log-likelihood contour plots with concentric elliptical patterns centered at MLEs (red dots) are shown in Figure 16, which shows well-behaved surfaces with mild parameter correlations typical of three-parameter models.



(a) $\hat{\alpha} = 0.683$ (b) $\hat{p} = 0.110$ (c) $\hat{\theta} = 0.298$

Figure 15. Log-likelihood profiles for the DOWL parameters based on Dataset III.



(a) Contour: α vs p. (b) Contour: α vs θ. (c) Contour: p vs θ.

Figure 16. Contour plots of the log-likelihood surface for Dataset III.

Table 20 juxtaposes the observed sample moments with the theoretical moments derived from the fitted DOWL distribution for the COVID-19 mortality dataset in Greece. The theoretical mean (1.723) substantially resembles

Table 20. Descriptive statistics for Dataset III: observed versus fitted DOWL moments.

Source	**Mean**	**Variance**	**Skewness**	**Kurtosis**	**DI**
Observed Sample	1.721	3.767	1.230	3.986	2.189
Fitted DOWL	1.723	3.683	1.256	4.299	2.137

the sample mean (1.721), and the theoretical variance (3.683) nearly approximates the sample variance (3.767), so validating the effective representation of dispersion in epidemic mortality data. The sample DI of 2.189 indicates significant over-dispersion, whereas the theoretical DI of 2.137 aligns closely, so affirming the model's efficacy in managing over-dispersed infectious disease data characterized by superspreading events, reporting delays, and epidemic waves that introduce additional Poisson variance. The skewness (1.256) signifies positive skewness with a right tail indicative of days with exceptionally high mortality, while the kurtosis (4.299) corroborates leptokurtic behavior with heavier tails associated with peak epidemic days vital for healthcare capacity planning.

The COVID-19 mortality dataset for Greece has obvious implications for Sustainable Development Goal 3 (Good Health and Well-being). Precise epidemic mortality modeling allows health officials to detect variations from anticipated trends, offering early alerts for potential additional waves or alterations in disease dynamics. Comprehending daily mortality distributions enables hospitals and ICUs to enhance preparedness for surges, guaranteeing the availability of essential resources (ventilators, ICU beds, medical personnel) when required. Precise baseline models enable policymakers to evaluate the efficacy of interventions (lockdowns, social distance, vaccination campaigns) by contrasting observed and anticipated death trends. Standardized modeling methodologies facilitate international comparisons, promoting the exchange of best practices and the identification of effective methods. Moreover, insights gained from modeling the initial COVID-19 wave enhance preparedness for future pandemics, fostering robust health systems capable of enduring emerging infectious diseases. The DOWL distribution's exceptional ability to address over-dispersion, positive skewness, and tail behavior underscores its utility as an instrumental resource for public health surveillance and pandemic response, facilitating informed decision-making that preserves lives and fortifies healthcare systems in alignment with SDG 3.

6.4. Policy and Decision-Relevant Outputs from DOWL Models

This subsection summarizes key actionable findings derived from the fitted DOWL models across the three empirical datasets:

- **Dataset I (Kidney cysts):** The fitted DOWL model estimates $P(X \geq 5) = 0.087$, indicating that approximately 9% of patients exhibit severe cyst burden. This information is clinically relevant for forecasting high-risk cases and guiding resource allocation in nephrology units.

- **Dataset II (Red mites):** The estimated hazard rate displays an increasing pattern after $x = 2$, suggesting accelerated population growth beyond the second generation. This implies that pesticide application should optimally occur before the second generation to prevent exponential pest expansion.
- **Dataset III (COVID-19 Greece):** The estimated 95th quantile of daily deaths is 7, providing a practical planning benchmark for ICU bed capacity. Additionally, the estimated DI of 2.14 confirms substantial overdispersion consistent with superspreading dynamics.

6.5. Connecting Statistical Outputs to SDG Targets and Policy Decisions

The DOWL distribution is not merely a theoretical contribution; its practical utility is demonstrated through three real-world datasets whose analysis directly informs evidence-based policy decisions aligned with the United Nations Sustainable Development Goals (SDGs). In what follows, we elaborate on how the statistical outputs derived from the DOWL model connect to specific SDG targets.

6.5.1. SDG 3 (Good Health and Well-Being) Two of the datasets analyzed in this study are directly related to public health, namely kidney cyst counts and COVID-19 case counts.

- **Kidney Cysts:** Accurate count modeling of kidney cyst frequencies enables clinicians and public health authorities to develop risk-stratified screening protocols. By capturing the distributional behavior of cyst counts more precisely than standard models such as the Poisson or Negative Binomial distributions, the DOWL model provides a reliable statistical basis for identifying high-risk patient subgroups. This directly supports **SDG Target 3.4**, which aims to reduce premature mortality from non-communicable diseases through prevention and treatment, as well as promote mental health and well-being.
- **COVID-19 Data:** The quantile function of the DOWL distribution provides precise estimates of case count thresholds at specified probability levels, which are essential for health system capacity planning. For instance, upper quantile estimates can inform hospital bed allocation, medical supply procurement, and surge response strategies. These outputs directly support **SDG Target 3.d**, which calls for strengthening the capacity of all countries to achieve health security and manage national and global health risks. The ability of the DOWL model to accommodate zero-inflated and overdispersed count data makes it particularly well-suited for modeling pandemic case distributions, where excess zero counts and heavy tails are commonly observed.

6.5.2. SDG 2 (Zero Hunger) and SDG 15 (Life on Land) The third dataset concerns red mite counts on apple leaves, which is an agronomic problem with direct implications for sustainable food production and ecosystem management.

- **Red Mite Infestation:** The shape of the hazard rate function estimated from the DOWL model carries important practical information for integrated pest management (IPM). Specifically, knowledge of whether the hazard rate is monotonically increasing, decreasing, or bathtub-shaped allows agricultural practitioners to determine the optimal timing for pesticide application or biological control interventions. A non-monotone hazard rate, for example, indicates periods of elevated infestation risk that may not be apparent from summary statistics alone. By targeting interventions at the appropriate stage of the infestation cycle, the DOWL model-based analysis supports a reduction in unnecessary pesticide overuse. This is directly aligned with **SDG Target 2.4**, which promotes sustainable food production systems and resilient agricultural practices that reduce ecological damage, and with **SDG 15**, which aims to protect, restore, and promote the sustainable use of terrestrial ecosystems and halt biodiversity loss caused by chemical overuse.

Table 21 summarizes the connections between the statistical outputs of the DOWL model, the datasets analyzed, and the relevant SDG targets.

Overall, these connections demonstrate that the DOWL distribution serves as a statistically rigorous and practically relevant tool for addressing real-world problems that are central to the global sustainability agenda. The flexibility of the model in capturing diverse count data patterns, including zero inflation, overdispersion, and non-monotone hazard rates, makes it a valuable resource for researchers and policymakers working toward the achievement of the SDGs.

Table 21. Summary of SDG linkages for the analyzed datasets.

Dataset	Statistical Output	Policy Implication	SDG Target
Kidney Cysts	PMF, hazard rate	Risk-stratified screening	SDG 3.4
COVID-19	Quantile function	Health system resilience	SDG 3.d
Red Mites	Hazard rate shape	Pest management timing	SDG 2.4, SDG 15

7. Concluding Remarks

A novel three-parameter discrete distribution, called the DOWL distribution, was created to model intricate count data in sustainable development research. The proposed model was developed by discretizing the continuous Lindley distribution using the discrete odd Weibull-G transformation framework, resulting in a versatile probability structure appropriate for various counting phenomena. A thorough examination of its mathematics and dependability characteristics was conducted. The parameter space was meticulously defined to guarantee model validity in sustainability-focused applications. The quantile function was formulated in closed form, enabling rapid random variate generation for simulation-based sustainability evaluations. The hazard rate function demonstrated various configurations, including increasing, decreasing, J-shaped, bathtub-shaped, and inverted bathtub-shaped failure rates, thereby encompassing a broad range of reliability aging categories pertinent to healthcare system resilience and environmental risk assessment. The inverse hazard rate function was also derived, enhancing the model's usefulness in retrospective survival analyses. Moments, variance, skewness, and kurtosis were calculated, demonstrating the model's ability to represent both symmetrical and asymmetrical data within mesokurtic, leptokurtic, and platykurtic frameworks. The index of dispersion was analyzed and shown to effectively handle over-dispersed, under-dispersed, and equi-dispersed data, even in the presence of extreme observations and outliers commonly found in epidemiological and ecological datasets. Information-theoretic metrics, such as Rényi and Shannon entropies, were developed to assess uncertainty in sustainability-related counting operations. Residual and historical lifetime functions, along with their corresponding aging classes, were developed, offering instruments for both prospective and retrospective evaluation of event durations in public health surveillance. The average previous lifetime function and equilibrium distribution were additionally derived to facilitate long-term sustainability planning analyses. Maximum likelihood estimators for the model parameters were formulated, and their efficacy was meticulously evaluated by comprehensive Monte Carlo simulation experiments including various sample sizes and parametric settings. The practical applicability of the DOWL distribution was illustrated through three empirical applications corresponding to the United Nations Sustainable Development Goals (SDGs): kidney cyst count data facilitated evidence-based clinical decision-making under SDG 3; European red mite infestation counts on apple leaves guided sustainable pest management strategies under SDG 2 and SDG 15; and daily COVID-19 mortality records from Greece during the initial pandemic wave offered insights into epidemic surveillance and healthcare capacity planning under SDG 3. Comparative analyses across various applications demonstrated that the DOWL model outperformed multiple competing discrete distributions in goodness-of-fit, establishing it as a robust and versatile framework for modeling over-dispersed, highly skewed, and zero-inflated count data in the context of sustainable development objectives.

Acknowledgements The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/35354)

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

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

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] Al-Babtain, A. A., Gemeay, A. M., & Afify, A. Z. (2021). Estimation methods for the discrete Poisson-Lindley and discrete Lindley distributions with actuarial measures and applications in medicine. *Journal of King Saud University–Science*, 33(2), 101224.

[3] Chan, S., Riley, P. R., Price, K. L., McElduff, F., & Winyard, P. J. (2009). 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, F346–F356.

[4] Chakraborty, S., & Chakravarty, D. (2012). Discrete gamma distributions: Properties and parameter estimations. *Communications in Statistics–Theory and Methods*, 41, 3301–3324.

[5] 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.

[6] El-Morshedy, M., Altun, E., & Eliwa, M. S. (2021). A new statistical approach to model the counts of novel coronavirus cases. *Mathematical Sciences*, 15, 1–14.

[7] El-Morshedy, M., Eliwa, M. S., & Tyagi, A. (2022). A discrete analogue of odd Weibull-G family of distributions: Properties, classical and Bayesian estimation with applications to count data. *Journal of Applied Statistics*, 49(11), 2928–2952.

[8] El-Morshedy, M. (2022). A discrete linear-exponential model: Synthesis and analysis with inference to model extreme count data. *Axioms*, 11(10), 531.

[9] El-Morshedy, M., Shahen, H. S., Almohaimeed, B., & Eliwa, M. S. (2022). Discrete single-factor extension of the exponential distribution: Features and modeling. *Axioms*, 11(12), 737.

[10] El-Morshedy, M., Eliwa, M. S., & Nagy, H. (2019). A new two-parameter exponentiated discrete Lindley distribution: Properties, estimation and applications. *Journal of Applied Statistics*, 47(2), 354–375.

[11] 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.

[12] 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.

[13] Eliwa, M. S., & El-Morshedy, M. (2022). A one-parameter discrete distribution for over-dispersed data: Statistical and reliability properties with applications. *Journal of Applied Statistics*, 49(10), 2467–2487.

[14] Das, D., Abouelenein, M. F., Das, B., Hazarika, P. J., El-Morshedy, M., Roushdy, N., & Eliwa, M. S. (2025). A discrete expansion of the Lindley distribution: Mathematical and statistical characterizations with estimation techniques, simulation, and goodness-of-fit analysis. *Computational Journal of Mathematical and Statistical Sciences*, 4(2), 622–645.

[15] Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. *TEST*, 19(2), 399–415.

[16] Gómez-Déniz, E., & Calderín-Ojeda, E. (2011). The discrete Lindley distribution: Properties and applications. *Journal of Statistical Computation and Simulation*, 81(11), 1405–1416.

[17] Hemeda, S. E., Elbarkawy, M. A., Al-Habib, K. H., Abdal-Hammed, M. K., Gamel, H. A., Almetwally, E. M., & Elgarhy, M. (2025). Development of a novel discrete distribution family from a continuous model: Applications to health, agricultural and crop-based fertilizer data. *Journal of Cultural Analysis and Social Change*, 2852–2861.

[18] Hussain, T., & Ahmad, M. (2014). Discrete inverse Rayleigh distribution. *Pakistan Journal of Statistics*, 30(2), 331–344.

[19] Hussain, T., Aslam, M., & Ahmad, M. (2016). A two-parameter discrete Lindley distribution. *Revista Colombiana de Estadística*, 39(1), 45–61.

[20] 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.

[21] Krishna, H., & Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, 6(2), 177–188.

[22] Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B (Methodological)*, 20(1), 102–107.

[23] Nagy, M., Almetwally, E. M., Gemeay, A. M., Mohammed, H. S., Jawa, T. M., Sayed-Ahmed, N., & Muse, A. H. (2021). The new novel discrete distribution with application on COVID-19 mortality numbers in Kingdom of Saudi Arabia and Latvia. *Complexity*, 2021, 7192833.

[24] 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(4), 407–422.

[25] Poisson, S. D. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile: Précédées des règles générales du calcul des probabilités*. Bachelier.

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

[27] Shahen, H. S., Eliwa, M. S., & El-Morshedy, M. (2025). Exploring the potential of the Kumaraswamy discrete half-logistic distribution in data science scanning and decision-making. *Annals of Data Science*, 12(3), 1013–1040.

[28] Shi, Y. (2022). Advances in big data analytics. *Advances in Big Data Analytics*, 10, 978–981.

[29] Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society, Series A*, 109(3), 296–296.

[30] United Nations Office for Disaster Risk Reduction (2023). *Global Assessment Report on Disaster Risk Reduction 2023: Mapping Resilience for the Sustainable Development Goals*. Stylus Publishing, LLC.