STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, August 2026, pp 963–1002.
Published online in International Academic Press (www.IAPress.org)

A Robust Flexible Hazard Rate Mixture Anchored in a Novel Over-dispersed Count Distribution

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

[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 research presents an innovative three-parameter discrete mixed distribution for the modeling of count data in various scientific domains that often display intricate characteristics, including overdispersion, heavy tails, and unimodality. The proposed distribution is created by combining an extended geometric distribution with the discrete Burr-Hatke distribution using a weighted mixing framework. This formulation establishes a unified and adaptable modeling framework wherein the component distributions manifest as distinct limiting cases as the mixing parameter approaches its boundary values. A thorough investigation of the parameter space's identifiability, the distribution's log-convexity, and its asymptotic properties is conducted. Closed forms are used to get various statistical and reliability properties. The hazard rate function is very flexible, allowing for patterns that go up and down, are unimodal, or are shaped like a bathtub. This makes it possible for the model to capture different ways that things can fail in reliability engineering and survival analysis. The maximum likelihood method is used to estimate parameters, and the log-likelihood function and score equations are derived in detail. The performance of the estimators on finite samples is carefully tested through a lot of Monte Carlo simulations with different sample sizes and parameter settings. These simulations show that the estimators are consistent, $\sqrt{n}$-consistent, and asymptotically normal. Evidence-based guidelines for suitable sample sizes are offered to aid practitioners in implementation. Three real-world datasets show that the proposed distribution works well in practice and outperforms other distributions. The proposed model consistently surpasses various established discrete distributions across all applications, as validated by standard information criteria and goodness-of-fit metrics. These results show that it is a strong and useful tool for modeling complex over-dispersed count data that have outliers and extreme values.

Keywords Discrete mixture models; Over-dispersed phenomena; Failure analysis; Simulation; Data analysis.

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

DOI: 10.19139/soic-2310-5070-3727

1. Introduction

The creation of adaptable discrete probability models that effectively represent the intricate features of real-world data has been a primary objective in statistical science. Classical discrete distributions, including the Poisson, binomial, geometric, and negative binomial distributions, have been essential for modelling count data in various fields, such as epidemiology, actuarial science, ecology, engineering reliability, and the social sciences [15]. Nevertheless, real data often display characteristics that significantly diverge from the assumptions of standard models, such as overdispersion, asymmetric, heavy tails, and uni-modality [27]. As a result, there has been persistent and increasing interest in developing new families of discrete distributions that provide enhanced flexibility and superior goodness-of-fit to empirical data.

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

ISSN 2310-5070 (online) ISSN 2311-004X (print)

The *mixture approach* has emerged as a prominent and effective paradigm among the several ways offered for producing new discrete probability models [21, 32]. A finite mixture model, in its broadest sense, expresses the probability mass function of a discrete random variable X as a convex combination of several component distributions.

$$P(X = x) = \lambda\, f_1(x;\, \boldsymbol{\theta}_1) + (1 - \lambda)\, f_2(x;\, \boldsymbol{\theta}_2), \quad 0 \le \lambda \le 1. \tag{1}$$

In this context, $f_1(x;\, \boldsymbol{\theta}_1)$ and $f_2(x;\, \boldsymbol{\theta}_2)$ represent the probability mass functions of the two component distributions, characterized by their respective parameter vectors $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, while λ serves as the *mixing parameter* (also known as the mixing proportion or weight) that regulates the relative contributions of each component to the overall mixture [11]. This approach seamlessly extends to k-component mixtures of the kind

$$P(X = x) = \sum_{i=1}^{k} \lambda_i\, f_i(x;\, \boldsymbol{\theta}_i), \quad \text{where} \quad \sum_{i=1}^{k} \lambda_i = 1, \quad \lambda_i \ge 0 \ \ \forall\, i. \tag{2}$$

The two-component scenario shown in Equation (1) is the most commonly utilized in practice because of its simplicity and clarity [21]. The mixing parameter λ is crucial in the mixture framework, as it regulates the extent of heterogeneity incorporated into the resultant model. When $\lambda = 0$ or $\lambda = 1$, the mixture simplifies to one of its constituent components, it reverts to the original baseline distributions as specific or limiting instances. For intermediate values of λ, the mixture distribution assimilates and integrates the attributes of both components, resulting in a model with improved flexibility regarding its shape, dispersion, tail behavior, and modality [16]. This trait is especially beneficial for modeling data derived from heterogeneous populations or processes when the observed counts are thought to originate from two or more separate subpopulations, each regulated by its own probabilistic mechanism [3]. The mixture approach possesses a substantial historical background, originating from the foundational research of [24], who presented finite mixes within the framework of continuous distributions. Since that time, the methodology has undergone substantial advancement, both in theoretical and computational dimensions, with notable contributions in identifiability [31], parameter estimation through the Expectation-Maximization (EM) algorithm [7], Bayesian inference [8], and model selection [19]. In discrete distribution contexts, mixture models have proven effective in mitigating the widespread issue of overdispersion in count data, characterized by observed variance surpassing the mean, a condition that contravenes the equidispersion property of the Poisson distribution [6]. Prominent instances encompass the Poisson-Lindley mixture [29], the Poisson-exponential mixture resulting in the geometric distribution, and the Poisson-gamma mixture producing the negative binomial distribution [12]. The mixture paradigm has recently been expanded in various novel areas. Researchers have suggested mixtures that incorporate weighted, transmuted, or size-biased component distributions [1], along with mixtures from disparate families such as the amalgamation of a degenerate distribution at zero with a conventional count distribution to formulate zero-inflated models [18]. Zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) models, conceptualized as two-component mixtures comprising a point mass at zero and a count distribution, are extensively utilized in modeling count data characterized by an abundance of zeros, frequently observed in manufacturing defect analysis, health services research, and species abundance studies [26, 33]. Hurdle models [22] and zero-altered models are closely related mixture-based approaches for addressing zero-modified count data.

The allure of the mixture strategy for creating novel discrete distributions is rooted in several persuasive benefits. The resultant mixture model retains manageable mathematical qualities from its components, frequently allowing closed-form formulas for the probability mass function, cumulative distribution function, moments, generating functions, and other distributional attributes [16]. The incorporation of the mixing parameter offers an extra degree of freedom, allowing the model to represent a wider variety of distributional forms, including skewness and kurtosis patterns unattainable by the individual components alone [21]. Third, the mixture formulation offers a natural and intuitive interpretation in various practical contexts: the mixing parameter can be perceived as the chance that an observation is associated with a distinct subpopulation or originates from a particular data-generating process [3]. Fourth, estimate techniques for mixed models are well-defined, with maximum likelihood estimation, method of moments, and Bayesian approaches all being widely applicable [7, 20].

Notwithstanding these benefits, mixture models can provide certain issues that warrant thorough examination. The likelihood surface of mixed models may display several local maxima, complicating the optimization process

[21]. Identifiability issues specifically, whether the parameters of the mixture can be uniquely ascertained from the distribution it delineates necessitate thorough examination for each proposed model [31]. Moreover, hypothesis testing at the parameter space border (e.g., testing $\lambda = 0$) entails nonstandard asymptotic theory, as the null hypothesis positions the mixing parameter at the edge of its permissible range [5].

This paper introduces an innovative discrete probability model employing the two-component mixture framework outlined in Equation (1). The suggested model is formulated by integrating the discrete Burr-Hatke distribution, initially developed by [9], with a discrete extension of the geometric distribution, recently introduced by [10], utilizing a singular mixing parameter λ. The resulting distribution is henceforth referred to as the *mixture geometric extension-discrete Burr-Hatke* (MGoDBu) distribution. This formulation generates a versatile array of distributions that includes both component distributions as specific limiting cases and demonstrates remarkable efficacy in modeling count data with intricate characteristics such as overdispersion, heavy tails, and asymmetry. The primary motivations driving the development of the proposed model are as follows:

1. The suggested distribution effectively models right-skewed data with heavy tails across different kurtosis levels, thus accommodating a wide range of distributional shapes typically observed in practice.

2. It offers a robust framework for modeling overdispersed count data, especially when extreme values and outlier observations present considerable difficulties for traditional discrete distributions.

3. The model provides closed-form formulas for various essential statistical and reliability aspects, such as moments, generating functions, quantiles, and associated measures, so aiding both theoretical study and practical application.

4. The hazard rate function of the proposed distribution demonstrates significant flexibility, capable of exhibiting decreasing, increasing, unimodal, and bathtub-shaped patterns, thus allowing the model to encapsulate various failure rate behaviors encountered in reliability engineering, survival analysis, and related domains.

The subsequent sections of this work are organized as follows. Section 2 delineates the mathematical framework of the proposed MGoDBu distribution, encompassing the derivation of the probability mass function via the mixing mechanism, graphical representation of the distributional behavior across diverse parameter configurations, and a thorough analysis of parameter identifiability characteristics. Section 3 presents a comprehensive examination of the distribution's structural properties. It begins by investigating key theoretical features, including log-convexity and asymptotic behavior. The section then provides an in-depth study of the hazard rate function, supported by graphical illustrations that demonstrate its possible shapes—decreasing, increasing, unimodal, and bathtub-shaped. Furthermore, the quantile function is derived and utilized to develop an efficient random data generation procedure. The section also establishes several moment-based statistical measures, including raw moments, information-theoretic entropy measures, as well as results related to order statistics and L-moments. Section 4 pertains to statistical inference, detailing the maximum likelihood estimation framework, including explicit derivations of the log-likelihood function and score equations, numerical optimization techniques, and the formulation of asymptotic confidence intervals utilizing the Fisher information matrix. Section 5 presents comprehensive Monte Carlo simulation analyses aimed at examining the finite-sample efficacy of maximum likelihood estimators across diverse sample sizes and parameter configurations, evaluating bias, mean squared error, coverage probabilities, and convergence rates. Section 6 illustrates the practical applicability and enhanced efficacy of the proposed distribution via empirical analyses of three real-world datasets pertaining to sustainability, juxtaposing the MGoDBu model with several established competing discrete distributions utilizing standard information criteria and goodness-of-fit metrics. Section 7 offers final observations, encapsulating the principal contributions.

2. Mathematical Framework of the MGoDBu Model with Graphical Visualization

In accordance with the two-component mixture framework specified in Equation (1) and the discrete modeling methodologies established by [9] and [10], we provide the MGoDBu distribution. The probability mass function

(PMF) for a random variable X adhering to this distribution is articulated as:

$$P(x;p,\beta,\lambda) = \lambda\left[2(1-p)p^x - (1-p^2)p^{2x}\right] + (1-\lambda)\beta^x\left(\frac{1}{x+1} - \frac{\beta}{x+2}\right), \quad x \in \mathbb{N}_0, \tag{3}$$

where $\mathbb{N}_0 = \{0,1,2,\ldots\}$ represents the set of non-negative integers. The parameter space is delineated by the conditions $0 < p < 1$, $0 < \beta < 1$, and $0 \leq \lambda \leq 1$. The parameters p and β dictate the form properties of the model, whilst λ the functions act as the mixing weight that modulates the relative contribution of each component to the overall mixture. A parametric family of distributions denoted as $\{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ is said to be *identifiable* if the mapping $\boldsymbol{\theta} \mapsto P_{\boldsymbol{\theta}}$ is injective, i.e., for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta$:

$$P_{\boldsymbol{\theta}_1} = P_{\boldsymbol{\theta}_2} \implies \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2. \tag{4}$$

A finite mixture model with a PMF of the type $P(x) = \sum_{j=1}^{k} \lambda_j f_j(x;\boldsymbol{\phi}_j)$ is identifiable if the mixture components and their corresponding weights can be uniquely determined from the mixture distribution, subject to permutation of component labels. For the MGoDBu distribution, the parameter vector is $\boldsymbol{\theta} = (p,\beta,\lambda)^\top \in \Theta$, with the parameter space delineated as:

$$\Theta = \{(p,\beta,\lambda) : 0 < p < 1,\ 0 < \beta < 1,\ 0 \leq \lambda \leq 1\}. \tag{5}$$

The MGoDBu mixture model comprises two components: the extension geometric distribution with PMF:

$$f_1(x;p) = 2(1-p)p^x - (1-p^2)p^{2x}, \quad x \in \mathbb{N}_0, \tag{6}$$

and the discrete Burr-Hatke distribution with PMF:

$$f_2(x;\beta) = \beta^x\left(\frac{1}{x+1} - \frac{\beta}{x+2}\right), \quad x \in \mathbb{N}_0. \tag{7}$$

We begin by establishing that these component distributions are structurally distinct.

Lemma 1
The component distributions $f_1(x;p)$ and $f_2(x;\beta)$ are part of fundamentally separate parametric families for all permissible parameter values. There are no values $p^* \in (0,1)$ and $\beta^* \in (0,1)$ for which $f_1(x;p^*) = f_2(x;\beta^*)$ holds for all $x \in \mathbb{N}_0$.

Proof
We analyze the asymptotic behavior of each component as x approaches infinity. The extension geometric component demonstrates purely exponential decline at a rate of p. The discrete Burr-Hatke component demonstrates exponential decay influenced by a polynomial factor $(x+1)^{-1}$. The inclusion of the polynomial modulation term in f_2 signifies a significant structural distinction. Assume, for the sake of contradiction, that $f_1(x;p^*) = f_2(x;\beta^*)$ holds for every $x \in \mathbb{N}_0$. Subsequently, the ratio:

$$\frac{f_1(x;p^*)}{f_2(x;\beta^*)} \sim \frac{2(1-p^*)(x+1)}{1-\beta^*}\left(\frac{p^*}{\beta^*}\right)^x \quad \text{as } x \to \infty \tag{8}$$

must be equal to one for every x. For $p^* = \beta^*$, the ratio increases linearly with respect to x, whereas for $p^* \neq \beta^*$, the ratio either diverges or converges to zero exponentially. In all scenarios, we arrive at a contradiction. Therefore, the two component families are structurally dissimilar. □

Based on this initial outcome, we now formulate the principal identifiability theorem.

Theorem 2.1. *The MGoDBu mixture model with the PMF defined by Equation (3) is identifiable inside the parameter space Θ. Specifically, for any two parameter vectors $\boldsymbol{\theta}_1 = (p_1,\beta_1,\lambda_1)$ and $\boldsymbol{\theta}_2 = (p_2,\beta_2,\lambda_2)$ inside the set Θ:*

$$P(x;p_1,\beta_1,\lambda_1) = P(x;p_2,\beta_2,\lambda_2) \quad \forall x \in \mathbb{N}_0 \implies \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2.$$

Proof
The proof advances by examining multiple scenarios contingent upon the values of the mixing parameter λ.

Case I: Boundary scenarios with unique mixture types. When $\lambda_1 = 1$, the mixture simplifies to the extension geometric distribution $f_1(x; p_1)$, whereas $\lambda_2 = 0$ results in the discrete Burr-Hatke distribution $f_2(x; \beta_2)$. According to Lemma 1, these cannot be identical, so such parameter combinations are readily distinguished.

Case II: Boundary cases with identical mixture types. For $\lambda_1 = \lambda_2 = 1$, identifiability reduces to the question of whether $f_1(x; p_1) = f_1(x; p_2)$ for all x implies $p_1 = p_2$. Evaluating at $x = 0$:

$$f_1(0; p) = 2(1 - p) - (1 - p^2) = (1 - p)^2.$$

Since $(1 - p)^2$ is strictly monotonic on $(0, 1)$, equality at $x = 0$ implies $p_1 = p_2$. Similarly, for $\lambda_1 = \lambda_2 = 0$, evaluating at $x = 0$ gives:

$$f_2(0; \beta) = 1 - \frac{\beta}{2},$$

which is strictly monotonic in $\beta \in (0, 1)$, establishing $\beta_1 = \beta_2$.

Case III: Interior case with $\lambda_1, \lambda_2 \in (0, 1)$. Assume $P(x; \boldsymbol{\theta}_1) = P(x; \boldsymbol{\theta}_2)$ for all $x \in \mathbb{N}_0$. This implies:

$$\lambda_1 f_1(x; p_1) + (1 - \lambda_1) f_2(x; \beta_1) = \lambda_2 f_1(x; p_2) + (1 - \lambda_2) f_2(x; \beta_2).$$

Rearranging terms:

$$\lambda_1 f_1(x; p_1) - \lambda_2 f_1(x; p_2) = (1 - \lambda_2) f_2(x; \beta_2) - (1 - \lambda_1) f_2(x; \beta_1). \tag{9}$$

The left side represents a linear combination of extended geometric distributions, whereas the right side denotes a linear combination of discrete Burr-Hatke distributions. According to Lemma 1, these families exhibit separate asymptotic behaviors. For the equality to be valid for all x, both sides must correspond to a shared distribution that is common to both families; however, this is unattainable unless both sides identically equal zero or reflect the same distribution via equivalent parameters. Specifically, when analyzing the tail behavior as $x \to \infty$, if $p_1 \neq p_2$, the left side displays a predominant exponential term with rate $\min(p_1, p_2)$, but if $\beta_1 \neq \beta_2$, the right side reveals a dominating term with rate $\min(\beta_1, \beta_2)$ adjusted by $(x + 1)^{-1}$. For congruence, it is essential that both the exponential rates and the polynomial modulations coincide, which is feasible only if $p_1 = p_2$ and $\beta_1 = \beta_2$. With $p_1 = p_2 = p$ and $\beta_1 = \beta_2 = \beta$, the equation simplifies to:

$$(\lambda_1 - \lambda_2)[f_1(x; p) - f_2(x; \beta)] = 0 \quad \forall x \in \mathbb{N}_0. \tag{10}$$

According to Lemma 1, $f_1(x; p) \neq f_2(x; \beta)$ for some x (unless f_1 and f_2 are identical, which they are not), indicating that $\lambda_1 = \lambda_2$. Consequently, $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$, confirming the identifiability of the MGoDBu model in all scenarios. □

This outcome ensures that maximum likelihood estimation and other inferential methods will produce unique parameter estimates, establishing a robust theoretical basis for the statistical applications of the MGoDBu model. Figure 1 depicts the distributional versatility of the MGoDBu model, showcasing its potential to produce varied probability mass functions via judicious parameterization. The model has notable efficacy in delineating right-skewed, heavy-tailed distributions with modifiable kurtosis, rendering it suitable for diverse data structures commonly observed in empirical studies.

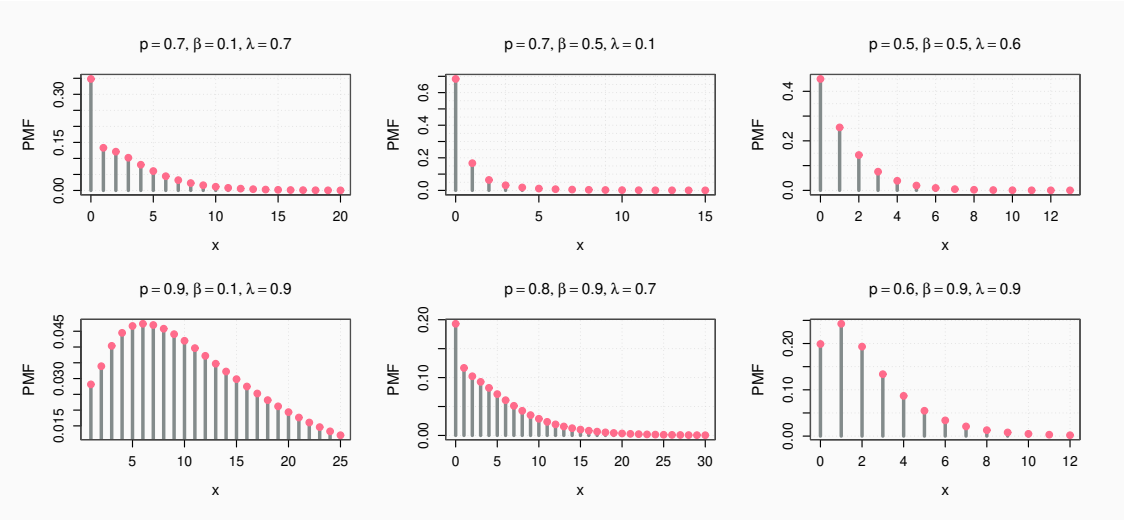


Figure 1. PMF of the MGoDBu distribution.

The MGoDBu distribution has a significant computational advantage because its closed-form CDF is obtained by directly summing the PMF in Equation (3). Thus, the CDF can be expressed as

$$F(x;p,\beta,\lambda) = 1 - 2\lambda p^{x+1} + \lambda p^{2x+2} - \frac{(1-\lambda)\beta^{x+1}}{x+2}, \quad x \in \mathbb{N}_0. \tag{11}$$

This manageable functional form has significant practical advantages, such as simpler probability calculations, efficient implementation of goodness-of-fit processes, and improved computational efficiency in parameter estimation by maximum likelihood and similar methods. The complex structure of the CDF, with terms p^{x+1}, $p^{2(x+1)}$, and $\frac{\beta^{x+1}}{x+2}$, makes the derivation of a closed-form expression for the quantile function $Q(u)$ analytically impracticable. As a result, the quantile is obtained numerically using sequential search. In analogous terms:

$$Q(u) = \min\left\{x \in \mathbb{N}_0 : 2\lambda p^{x+1} - \lambda p^{2(x+1)} + \frac{(1-\lambda)\beta^{x+1}}{x+2} \leq 1-u\right\}. \tag{12}$$

Considering that $0 < p, \beta < 1$, the left-hand side demonstrates a monotonic reduction in x, hence ensuring the existence and uniqueness of the solution for every $u \in (0,1)$.

3. Distributional Properties

3.1. Log-Convexity and Asymptotic Behavior

A non-negative sequence $\{a_x\}$ is *log-convex* if $a_x^2 \leq a_{x-1}\, a_{x+1}$ for all $x \geq 1$.

Lemma 2
The sequence $\{f_1(x)\}$ is log-concave: $f_1(x)^2 \geq f_1(x-1)\, f_1(x+1)$ for all $x \geq 1$.

Proof
Setting $\alpha = 2(1-p)$ and $\gamma = 1-p^2$, direct computation yields

$$f_1(x)^2 - f_1(x-1)\, f_1(x+1) = \alpha\gamma\,(1-p)^2\, p^{3x-1} > 0. \qquad \square$$

Lemma 3
The sequence $\{f_2(x)\}$ is log-convex for all $\beta \in (0,1)$.

Proof
Define $f_2(x) = \beta^x h(x)$ where $h(x) = \frac{(1-\beta)x+(2-\beta)}{(x+1)(x+2)}$. Given that $\{\beta^x\}$ is log-linear, it is sufficient to confirm the log-convexity of $\{h(x)\}$. By defining $a = 1-\beta$ and $b = 2-\beta$, one derives

$$h(x-1)\,h(x+1) - h(x)^2 = \frac{a^2x^2 + a(5-\beta)\,x + 2(3-2\beta)}{x(x+1)^2(x+2)^2(x+3)} > 0,$$

since all coefficients are positive for $\beta \in (0,1)$. □

Theorem 3.1. *For $0 < \lambda < 1$ with $\beta > p$, the PMF $\{P(x)\}$ is eventually log-convex: there exists $x_0 \in \mathbb{N}$ such that $P(x)^2 \le P(x-1)\,P(x+1)$ for all $x \ge x_0$.*

Proof
When $\beta > p$, the ratio $f_1(x)/f_2(x) = O(x(p/\beta)^x) \to 0$, so $P(x) \sim (1-\lambda)f_2(x)$ for large x. The log-concave perturbation contributes $O(p^{3x})$ to $P(x)^2 - P(x-1)P(x+1)$, while the log-convex contribution is $O(\beta^{3x}/x^2)$. Since $\beta > p$, the latter dominates for $x \ge x_0$. □

The findings indicate that the distribution demonstrates a *decreasing failure rate* (DFR) in its tail when $\beta > p$, while the ratio $P(x+1)/P(x)$ ultimately remains non-decreasing. Given that $0 < p < 1$ implies $p^{2x} = o(p^x)$, the initial component is satisfied.

$$f_1(x) \sim 2(1-p)\,p^x, \quad x \to \infty. \tag{13}$$

For the second component, noting that

$$\frac{1}{x+1} - \frac{\beta}{x+2} = \frac{(1-\beta)x + (2-\beta)}{(x+1)(x+2)} \sim \frac{1-\beta}{x}, \quad x \to \infty,$$

we obtain

$$f_2(x) \sim \frac{(1-\beta)\,\beta^x}{x}, \quad x \to \infty. \tag{14}$$

The tail of the full PMF is therefore governed by the dominant geometric base:

$$P(x) \sim \begin{cases} 2\lambda(1-p)\,p^x, & \text{if } p > \beta, \\ \dfrac{(1-\lambda)(1-\beta)\,\beta^x}{x}, & \text{if } \beta > p, \\ 2\lambda(1-p)\,p^x, & \text{if } p = \beta, \end{cases} \quad x \to \infty. \tag{15}$$

In all cases, $\lim_{x\to\infty} x^{-1} \log P(x) = \log \max(p,\beta) < 0$, guaranteeing geometric decay and, consequently, the finiteness of all moments.

3.2. Hazard Rate Function: Analytical Properties and Visual Illustration

The discrete hazard rate function (HRF) of the MGoDBu distribution is

$$h(x;p,\beta,\lambda) = \frac{\lambda\left[2(1-p)p^x - (1-p^2)p^{2x}\right] + (1-\lambda)\beta^x\left(\dfrac{1}{x+1} - \dfrac{\beta}{x+2}\right)}{2\lambda p^x - \lambda p^{2x} - \dfrac{(1-\lambda)\beta^x}{x+1}}, \qquad x \in \mathbb{N}_0. \tag{16}$$

Subsequently, we examine the mathematical characteristics of $h(x;p,\beta,\lambda)$, focusing specifically on its asymptotic aspects. Compose

$$N(x) = \lambda\left[2(1-p)p^x - (1-p^2)p^{2x}\right] + (1-\lambda)\beta^x\left(\frac{1}{x+1} - \frac{\beta}{x+2}\right),$$

and

$$D(x) = 2\lambda p^x - \lambda p^{2x} - \frac{(1-\lambda)\beta^x}{x+1},$$

so that

$$h(x; p, \beta, \lambda) = \frac{N(x)}{D(x)}.$$

Given that $0 < p < 1$ and $0 < \beta < 1$, both p^x and β^x converge to zero as $x \to \infty$. Thus, the asymptotic behavior of $h(x)$ as x approaches infinity is dictated by the comparative rates of decay of the terms involving p^x, p^{2x}, and $\beta^x/(x+1)$. The HRF at $x = 0$ is

$$h(0; p, \beta, \lambda) = \frac{\lambda\left[2(1-p) - (1-p^2)\right] + (1-\lambda)\left(1 - \frac{\beta}{2}\right)}{\lambda - (1-\lambda)}.$$

Since

$$2(1-p) - (1-p^2) = 1 - 2p + p^2 = (1-p)^2,$$

we obtain

$$h(0; p, \beta, \lambda) = \frac{\lambda(1-p)^2 + (1-\lambda)\left(1 - \frac{\beta}{2}\right)}{2\lambda - 1}. \tag{17}$$

Consequently, the risk at the origin is significantly influenced by λ, necessitating careful consideration of the denominator. In applications, it is essential to guarantee that the denominator in Equation (16) remains nonzero for the specified range of x.

3.2.1. Asymptotic behaviour as $x \to \infty$ To analyze the asymptotics, we examine the principal terms in the numerator and denominator.

Case I: $\lambda = 1$. When $\lambda = 1$, the model reduces to

$$h(x; p, \beta, 1) = \frac{2(1-p)p^x - (1-p^2)p^{2x}}{2p^x - p^{2x}}.$$

Factoring out p^x gives

$$h(x; p, \beta, 1) = \frac{2(1-p) - (1-p^2)p^x}{2 - p^x}.$$

Since $p^x \to 0$ as $x \to \infty$, it follows that

$$\lim_{x\to\infty} h(x; p, \beta, 1) = 1 - p. \tag{18}$$

Consequently, for the geometric extension component, the risk approaches the positive constant $1 - p$. Thus, in this case, the hazard demonstrates a *horizontal asymptote* at $y = 1 - p$.

Case II: $0 < \lambda < 1$. The asymptotic behavior in the general mixed situation depends on the relative magnitudes of p and β.

(a) If $p > \beta$, then p^x decays more slowly than β^x. In this case, the p^x-terms dominate. Since

$$p^{2x} = o(p^x), \qquad \frac{\beta^x}{x+1} = o(p^x),$$

we have

$$N(x) \sim 2\lambda(1-p)p^x, \qquad D(x) \sim 2\lambda p^x.$$

Therefore,

$$\lim_{x\to\infty} h(x;p,\beta,\lambda) = 1-p. \tag{19}$$

So the hazard has the horizontal asymptote $y = 1-p$.

(b) If $p = \beta$, then both exponential terms decay at the same rate. Let $p = \beta$. Then

$$N(x) = \lambda\left[2(1-p)p^x - (1-p^2)p^{2x}\right] + (1-\lambda)p^x\left(\frac{1}{x+1} - \frac{p}{x+2}\right),$$

and

$$D(x) = 2\lambda p^x - \lambda p^{2x} - \frac{(1-\lambda)p^x}{x+1}.$$

Factoring out p^x gives

$$h(x) = \frac{\lambda\left[2(1-p) - (1-p^2)p^x\right] + (1-\lambda)\left(\frac{1}{x+1} - \frac{p}{x+2}\right)}{2\lambda - \lambda p^x - \frac{1-\lambda}{x+1}}.$$

As $x \to \infty$, we obtain

$$\lim_{x\to\infty} h(x;p,p,\lambda) = 1-p, \qquad \text{provided } \lambda > 0. \tag{20}$$

Hence for $\lambda > 0$, the horizontal asymptote is again $y = 1-p$.

3.2.2. Vertical asymptotes and singularities A vertical asymptote may occur at values of x where the denominator is zero:

$$2\lambda p^x - \lambda p^{2x} - \frac{(1-\lambda)\beta^x}{x+1} = 0. \tag{21}$$

Equivalently,

$$\lambda\left(2p^x - p^{2x}\right) = \frac{(1-\lambda)\beta^x}{x+1}.$$

Since x is discrete ($x \in \mathbb{N}_0$), these places do not constitute vertical asymptotes in the conventional continuous sense; instead, they are singular points of the hazard sequence where $h(x)$ is undefined. Therefore, the HRF is explicitly specified only for those x such that

$$D(x) \neq 0.$$

Moreover, because

$$2p^x - p^{2x} = p^x(2-p^x) > 0,$$

Thus, for specific parameter combinations, the denominator may change its sign, leading to discontinuous behavior in the series. Although establishing a complete proof of global monotonicity in closed form is difficult, the asymptotic equations suggest that as x increases, the hazard converges to a constant value. If $p \geq \beta$ and $\lambda > 0$, then

$$h(x) = 1 - p + o(1).$$

Thus, despite its temporary fluctuations, the hazard sequence is *asymptotically flat*. In other words, it approaches a constant value as $x \to \infty$. More precise estimations may also be acquired. The MGoDBu model's adaptability in reliability scenarios is demonstrated by the graphs in Figure 2.

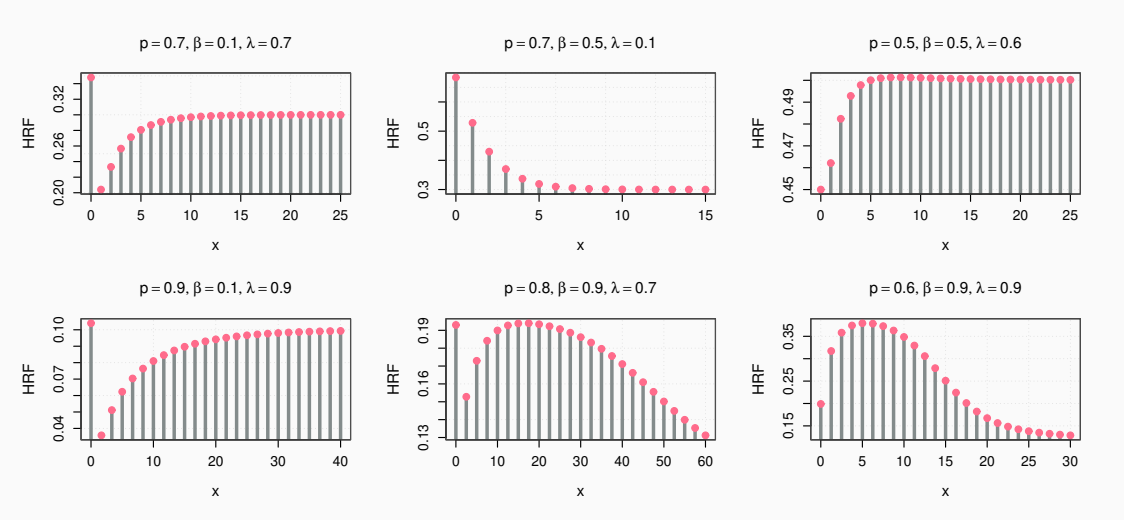


Figure 2. HRF of the MGoDBu distribution.

The HRF may display a monotonically increasing, monotonically decreasing, or non-monotonic profile, including shapes such as bathtub or unimodal forms, depending on different parameter combinations. This trait is highly desirable for modeling real-world systems where failure mechanisms develop over time. The noted adaptability underscores the MGoDBu distribution's potential as a superior alternative to more rigid classical discrete models. Parameter combinations in the region where $p < \beta$ and $\lambda < 0.5$ can lead to a negative hazard rate asymptote, and we strongly advise against using the model in this region for survival or reliability analysis.

3.3. Moments and Related Statistical Measures

Moments are critical descriptive metrics that characterize the structure and behavior of probability distributions. They are essential for computing key summary statistics, such as the mean, variance, skewness, and kurtosis, which are vital for model comparison and statistical inference. Let X be a random variable that follows the MGoDBu distribution defined by a survival function. The nth raw moment of X can be expressed using the tail-sum formula:

$$\begin{aligned} E(X^n) &= \sum_{x=0}^{\infty} \left[(x+1)^n - x^n\right] S(x) \\ &= \sum_{x=0}^{\infty} \left[(x+1)^n - x^n\right] \left[2\lambda p^{x+1} - \lambda p^{2x+2} + \frac{(1-\lambda)\beta^{x+1}}{x+2}\right], \quad n = 1, 2, 3, \ldots, \end{aligned} \tag{22}$$

where $S(x) = P(X \geq x)$ defines the survival function of the MGoDBu distribution. By simplifying (22), the first four raw moments are derived in closed form as follows:

$$E(X) = \frac{-2\lambda p}{p-1} + \frac{\lambda p^2}{p^2-1} + (\lambda - 1)A, \tag{23}$$

$$E(X^2) = (\lambda-1)A + \lambda p\left[\frac{2(p+1)}{(p-1)^2} - \frac{p(p^2+1)}{(p^2-1)^2}\right] + \frac{2(1-\lambda)}{\beta(\beta-1)}\left[\beta^2 - 2\beta + 3(\beta-1)\ln(1-\beta)\right], \tag{24}$$

$$E(X^3) = (\lambda-1)A + \frac{2\lambda p\left[-3p(p+1) + 3p(p-1) - (p-1)^2\right]}{(p-1)^3} + \frac{\lambda p^2\left[3p^2(p^2+1) - 3p^2(p^2-1) + (p^2-1)^2\right]}{(p^2-1)^3} + (1-\lambda)(B+C), \tag{25}$$

$$E(X^4) = \frac{2\lambda p(1-12p-p^2)}{(p-1)^3} + \frac{\lambda p^2(1+8p^2+3p^4)}{(p^2-1)^3} + \frac{\lambda p^2\left(4p^6+64p^5+168p^4+256p^3+180p^2+64p+8\right)}{(p-1)^4(p+1)^2} + (1-\lambda)\left(-A + \frac{4}{3}C + \frac{4}{3}D + 2B\right), \tag{26}$$

where the auxiliary quantities A, B, C, and D are functions of β defined by

$$A = 1 + \frac{\ln(1-\beta)}{\beta}, \qquad B = \frac{-3\beta^2+18\beta-12}{(\beta-1)^2} - \frac{12\ln(1-\beta)}{\beta},$$
$$C = \frac{3\beta-6}{\beta-1} + \frac{6\ln(1-\beta)}{\beta}, \qquad D = \frac{3\beta^3-45\beta^2+60\beta-24}{(\beta-1)^3} + \frac{24\ln(1-\beta)}{\beta}. \tag{27}$$

After calculating the raw moments, one can get the central moments and related metrics, including variance "$Var(X)$", skewness coefficient "$Sk(X)$", kurtosis coefficient "$Ku(X)$", and dispersion index "$IoD(X)$". Tables 1, 2, and 3 present the descriptive statistics for fixed values of λ, p, and β, respectively. Table 1 demonstrates that the proposed MGoDBu model can effectively capture asymmetric data exhibiting right-skewed patterns across various levels of leptokurtosis. Furthermore, the model is well-suited for analyzing data characterized by overdispersion phenomena, including the presence of extreme values and outliers. These properties enhance the model's applicability for data analysis across diverse fields, particularly in sustainability research and related domains. As discussed in previous sections, the HRF exhibits considerable flexibility, accommodating various functional forms that further extend the model's versatility. Three-dimensional surface plots illustrating the behavior of key statistical measures including the mean, variance, skewness, kurtosis, and index of dispersion—over the parameter space are presented in Figures 3–7.

Table 1. Descriptive statistics of the MGoDBu distribution for fixed $\lambda = 0.6$.

Measure	p	β								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E(X)$	0.1	0.1487	0.1736	0.2028	0.2381	0.2818	0.3381	0.4153	0.5320	0.7506
	0.3	0.4764	0.5012	0.5305	0.5658	0.6095	0.6658	0.7429	0.8597	1.0783
	0.5	1.0214	1.0463	1.0756	1.1108	1.1545	1.2109	1.2880	1.4047	1.6234
	0.9	8.2635	8.2884	8.3177	8.3529	8.3966	8.4530	8.5301	8.6468	8.8655
$\mathrm{Var}(X)$	0.1	0.1592	0.1878	0.2318	0.3009	0.4138	0.6099	0.9871	1.8596	4.9228
	0.3	0.6816	0.6939	0.7187	0.7648	0.8490	1.0081	1.3348	2.1308	5.0507
	0.5	2.2479	2.2331	2.2260	2.2336	2.2702	2.3679	2.6105	3.2793	5.9608
	0.9	112.57	112.20	111.77	111.26	110.67	109.95	109.07	108.05	107.57
$\mathrm{Sk}(X)$	0.1	2.8789	2.7238	2.7900	3.0668	3.5126	4.0826	4.7659	5.6549	7.2722
	0.3	2.2296	2.1366	2.0471	1.9967	2.0559	2.3440	3.0281	4.3105	6.6448
	0.5	2.0605	2.0410	2.0099	1.9657	1.9139	1.8843	1.9898	2.6116	4.9660
	0.9	1.7925	1.7980	1.8042	1.8111	1.8187	1.8269	1.8345	1.8374	1.8187
$\mathrm{Ku}(X)$	0.1	12.433	11.750	13.451	17.493	23.488	31.116	40.787	55.224	88.956
	0.3	9.6547	9.1970	8.7214	8.5309	9.3495	12.642	20.970	38.771	80.623
	0.5	8.7877	8.7681	8.6820	8.5086	8.2670	8.1848	9.4020	16.958	54.268
	0.9	7.3929	7.4220	7.4555	7.4944	7.5401	7.5937	7.6548	7.7104	7.6746
$\mathrm{IoD}(X)$	0.1	1.0707	1.0820	1.1427	1.2640	1.4685	1.8036	2.3770	3.4956	6.5581
	0.3	1.4309	1.3844	1.3548	1.3518	1.3931	1.5142	1.7967	2.4787	4.6839
	0.5	2.2007	2.1343	2.0696	2.0107	1.9664	1.9556	2.0268	2.3345	3.6719
	0.9	13.623	13.537	13.437	13.320	13.180	13.007	12.787	12.496	12.133

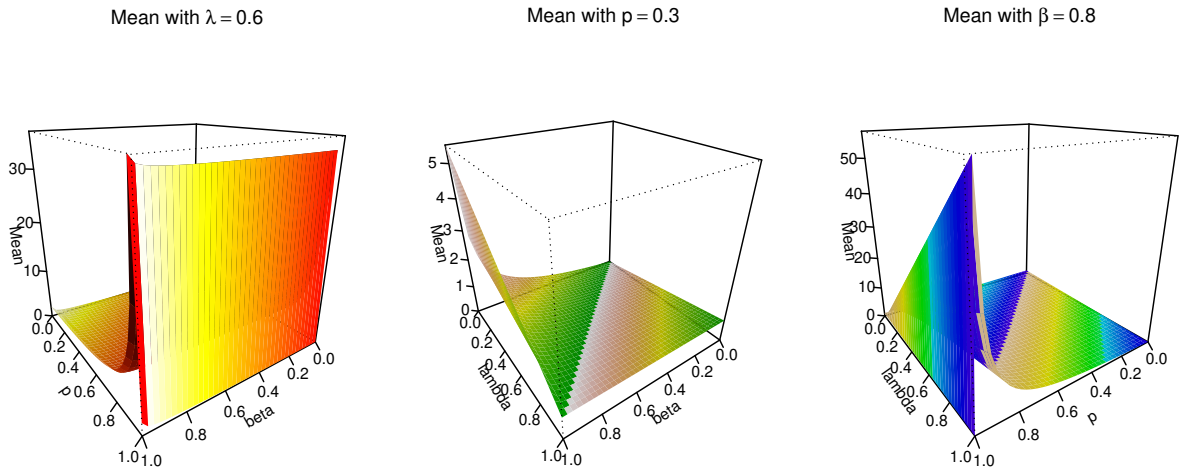


Figure 3. Mean of the MGoDBu distribution.

Table 2. Descriptive statistics of the MGoDBu distribution for fixed $p = 0.3$.

Measure	β	λ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E(X)$	0.2	0.1800	0.2442	0.3088	0.3727	0.4370	0.5012	0.5655	0.6297	0.6940
	0.4	0.3252	0.3733	0.4214	0.4695	0.5177	0.5658	0.6139	0.6620	0.7101
	0.6	0.5503	0.5734	0.5965	0.6196	0.6427	0.6658	0.6889	0.7120	0.7351
	0.8	0.9864	0.9611	0.9357	0.9104	0.8850	0.8597	0.8343	0.8090	0.7836
$\mathrm{Var}(X)$	0.2	0.2525	0.3573	0.4538	0.5421	0.6222	0.6939	0.7574	0.8127	0.8597
	0.4	0.4935	0.5570	0.6159	0.6702	0.7198	0.7648	0.8051	0.8407	0.8720
	0.6	1.1212	1.1008	1.0792	1.0566	1.0329	1.0081	0.9823	0.9554	0.9275
	0.8	3.6424	3.3427	3.0416	2.7393	2.4357	2.1308	1.8247	1.5172	1.2085
$\mathrm{Sk}(X)$	0.2	3.6635	3.2313	2.8672	2.5741	2.3351	2.1366	1.9693	1.8267	1.7047
	0.4	2.9244	2.6720	2.4622	2.2841	2.1306	1.9967	1.8789	1.7747	1.6822
	0.6	3.1271	2.9775	2.8252	2.6694	2.5094	2.3440	2.1719	1.9918	1.8019
	0.8	3.9444	4.0431	4.1383	4.2239	4.2885	4.3105	4.2465	4.0020	3.3455
$\mathrm{Ku}(X)$	0.2	21.635	17.299	14.149	11.946	10.364	9.1970	8.3180	7.6486	7.1382
	0.4	14.991	12.918	11.377	10.195	9.2688	8.5309	7.9368	7.4559	7.0665
	0.6	18.322	17.258	16.169	15.044	13.873	12.642	11.335	9.9345	8.4164
	0.8	28.313	30.151	32.164	34.336	36.601	38.771	40.344	39.962	33.576
$\mathrm{IoD}(X)$	0.2	1.4031	1.4630	1.4713	1.4545	1.4238	1.3844	1.3395	1.2905	1.2388
	0.4	1.5177	1.4922	1.4615	1.4273	1.3905	1.3518	1.3115	1.2702	1.2279
	0.6	2.0377	1.9198	1.8093	1.7053	1.6071	1.5142	1.4259	1.3419	1.2616
	0.8	3.6925	3.4780	3.2506	3.0090	2.7522	2.4787	2.1870	1.8755	1.5422

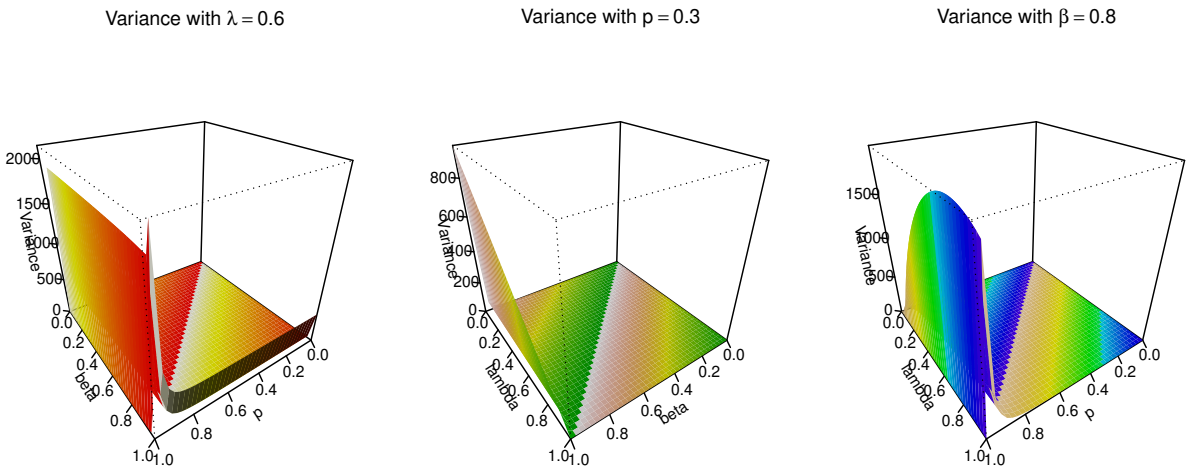


Figure 4. Variance of the MGoDBu distribution.

Table 3. Descriptive statistics of the MGoDBu distribution for fixed $\beta = 0.8$.

Measure	λ	p								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$E(X)$	0.1	0.9318	0.9565	0.9864	1.0249	1.0773	1.1544	1.2812	1.5328	2.2843
	0.4	0.6919	0.7904	0.9104	1.0642	1.2737	1.5821	2.0894	3.0960	6.1018
	0.7	0.4520	0.6244	0.8343	1.1035	1.4702	2.0098	2.8977	4.6591	9.9193
	0.9	0.2921	0.5137	0.7836	1.1298	1.6012	2.2949	3.4365	5.7012	12.464
$\mathrm{Var}(X)$	0.1	3.6260	3.6238	3.6424	3.7025	3.8521	4.2161	5.1904	8.5087	29.389
	0.4	2.6045	2.6360	2.7393	2.9854	3.5341	4.7977	8.0680	18.954	86.302
	0.7	1.4680	1.5929	1.8247	2.2652	3.1390	5.0134	9.6392	24.513	114.07
	0.9	0.6463	0.8670	1.2085	1.7834	2.8327	4.9540	9.9608	25.504	116.86
$\mathrm{Sk}(X)$	0.1	4.0359	4.0067	3.9444	3.8306	3.6444	3.3884	3.1952	3.5214	4.6537
	0.4	4.8206	4.5937	4.2239	3.6876	3.0388	2.4653	2.1995	2.2485	2.3484
	0.7	6.2437	5.3057	4.2465	3.2037	2.3819	1.9116	1.7455	1.7191	1.6863
	0.9	7.8051	5.0582	3.3455	2.3586	1.8660	1.6677	1.6095	1.5915	1.5607
$\mathrm{Ku}(X)$	0.1	29.001	28.831	28.313	27.193	25.103	21.808	18.564	21.130	33.177
	0.4	40.240	38.206	34.336	28.206	20.431	13.447	10.080	10.027	10.366
	0.7	68.119	55.242	40.344	25.828	14.970	9.3361	7.6038	7.3570	7.1278
	0.9	123.53	65.717	33.576	17.187	10.068	7.6349	7.0636	6.9624	6.8346
$\mathrm{IoD}(X)$	0.1	3.8912	3.7888	3.6925	3.6126	3.5757	3.6523	4.0512	5.5509	12.866
	0.4	3.7642	3.3349	3.0090	2.8053	2.7746	3.0325	3.8613	6.1222	14.144
	0.7	3.2475	2.5513	2.1870	2.0527	2.1351	2.4945	3.3265	5.2613	11.500
	0.9	2.2127	1.6878	1.5422	1.5786	1.7691	2.1587	2.8986	4.4735	9.3375

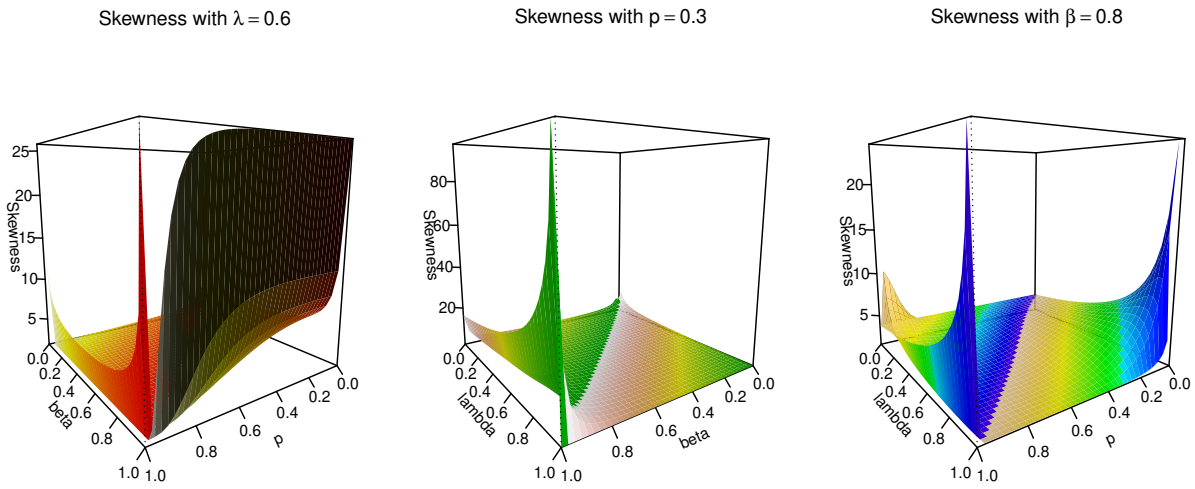


Figure 5. Skewness of the MGoDBu distribution.

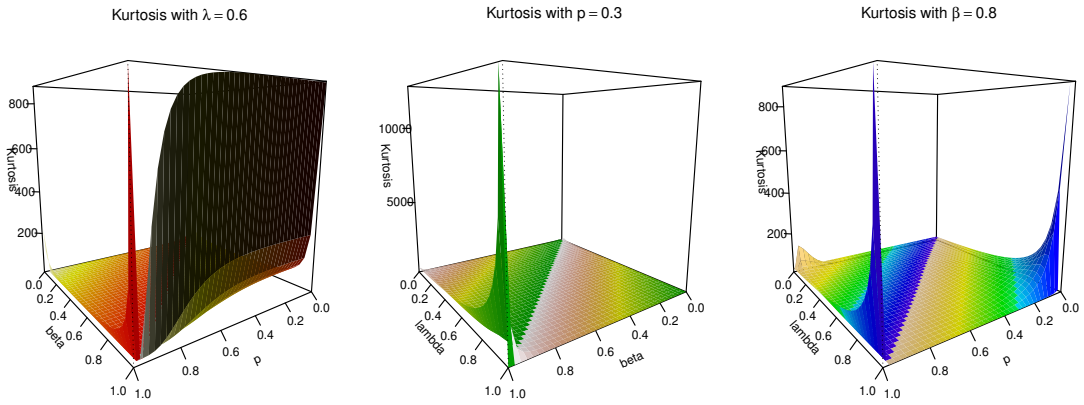


Figure 6. Kurtosis of the MGoDBu distribution.

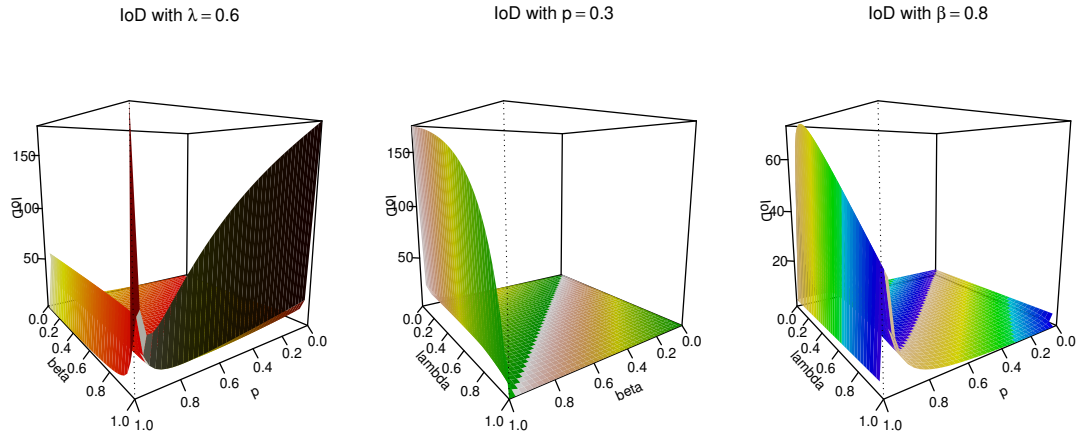


Figure 7. IoD of the MGoDBu distribution.

3.4. Information-Theoretic Entropy Measures

Entropy measures the uncertainty or randomness associated with a random variable and is a fundamental concept in information theory. The Rényi entropy for the MGoDBu distribution has been recognized as an extension of several prior entropy measurements. The Rényi entropy of order δ ($\delta > 0$, $\delta \neq 1$) is articulated as

$$I_\delta(X) = \frac{1}{1-\delta} \log \left(\sum_{x=0}^{\infty} \left[\lambda \left\{ 2(1-p)p^x - (1-p^2)p^{2x} \right\} + (1-\lambda)\beta^x \left(\frac{1}{x+1} - \frac{\beta}{x+2} \right) \right]^\delta \right). \tag{28}$$

The Rényi entropy serves as a versatile metric; significant examples are the Shannon entropy ($\delta \to 1$), the collision entropy ($\delta = 2$), the min-entropy ($\delta \to \infty$), and the max-entropy ($\delta \to 0$). The Shannon entropy for the MGoDBu distribution is expressed as

$$I(X) = -\sum_{x=0}^{\infty} P(x; p, \beta, \lambda) \, \log\big(P(x; p, \beta, \lambda)\big). \tag{29}$$

The Rényi entropy with $\delta = 2$ (collision entropy) has been computed for the MGoDBu distribution across several parameter configurations. The results are presented in Tables 4, 5, and 6.

Table 4. Rényi entropy ($\delta = 2$) of the MGoDBu distribution for fixed $\lambda = 0.6$.

$p\backslash\beta$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.2683	0.3084	0.3493	0.3913	0.4349	0.4803	0.5279	0.5782	0.6318
0.3	0.6750	0.7124	0.7507	0.7906	0.8325	0.8775	0.9262	0.9799	1.0399
0.5	1.0644	1.1030	1.1412	1.1800	1.2202	1.2632	1.3103	1.3636	1.4261
0.9	1.8166	1.9032	1.9895	2.0762	2.1638	2.2529	2.3444	2.4397	2.5418

Table 4 reveals that, for a constant p, the entropy escalates with increasing β, signifying heightened uncertainty. Likewise, with a constant β, the entropy escalates with p. The entropy values span from 0.2683 to 2.5418, illustrating the distribution's adaptability in encapsulating diverse degrees of uncertainty.

Table 5. Rényi entropy ($\delta = 2$) of the MGoDBu distribution for fixed $p = 0.3$.

$\beta\backslash\lambda$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.2	0.2865	0.3725	0.4588	0.5448	0.6297	0.7124	0.7918	0.8664	0.9345
0.4	0.4762	0.5413	0.6057	0.6690	0.7309	0.7906	0.8475	0.9009	0.9501
0.6	0.6822	0.7242	0.7650	0.8043	0.8419	0.8775	0.9108	0.9415	0.9694
0.8	0.9197	0.9352	0.9490	0.9611	0.9714	0.9799	0.9864	0.9910	0.9936

Table 5 indicates that, given a constant β, the entropy often rises with an increase in λ, signifying the augmented influence of the extended geometric component. The entropy values converge towards one when both parameters grow, signifying a more equitable distribution of probability mass.

Table 6. Rényi entropy ($\delta = 2$) of the MGoDBu distribution for fixed $\beta = 0.8$.

$\lambda \backslash p$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.8467	0.8841	0.9197	0.9539	0.9871	1.0192	1.0498	1.0779	1.1012
0.4	0.6833	0.8271	0.9611	1.0927	1.2275	1.3691	1.5189	1.6746	1.8246
0.7	0.5269	0.7680	0.9864	1.1982	1.4196	1.6667	1.9578	2.3173	2.7782
0.9	0.4267	0.7277	0.9936	1.2431	1.4984	1.7836	2.1294	2.5913	3.3219

Table 6 indicates that, given a constant λ, the entropy significantly escalates as p approaches 1, attaining values up to 3.3219. This signifies that the distribution becomes increasingly diffuse as the extended geometric component prevails and p escalates. Figure 8 presents a three-dimensional depiction of the entropy within the parameter space, further elucidating the uncertainty attributes of the MGoDBu distribution.

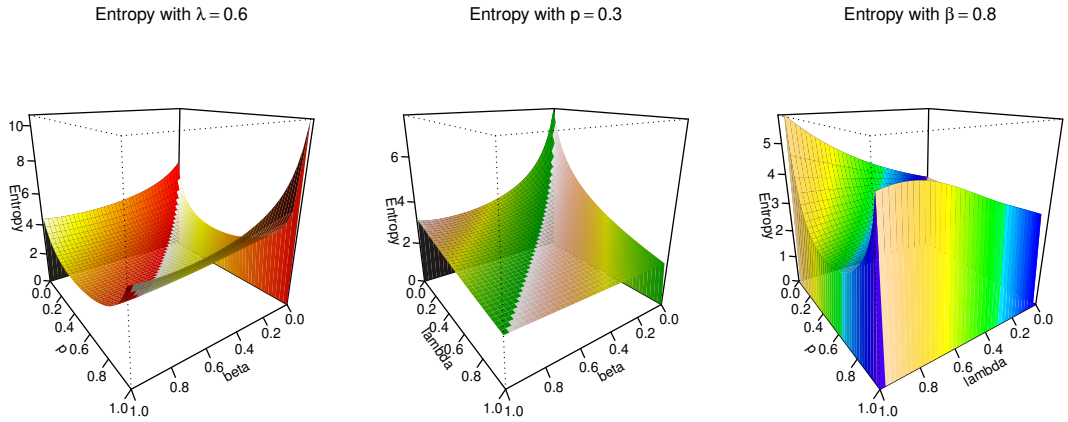


Figure 8. Entropy of the MGoDBu distribution.

3.5. Order Statistics and L-Moments

Order statistics are fundamental to non-parametric inference, reliability theory, and the formulation of robust L-moments. Let $X_1, X_2, \ldots, X_n$ represent an independent and identically distributed random sample from the MGoDBu distribution, and let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the associated order statistics. To analytically derive the cumulative distribution function of the ith order statistic, we utilize binomial expansions on the usual order statistic formula. This produces a linear combination of fundamental distribution functions. We succinctly define the multi-index $\mathbf{s} = (k, j, l, m, w)$ and the finite index set $\mathcal{S}_{n,i}$ as follows:

$$\mathcal{S}_{n,i} = \{(k, j, l, m, w) \in \mathbb{N}_0^5 : i \leq k \leq n,\ 0 \leq j \leq n-k,\ 0 \leq l \leq k+j,\ 0 \leq m \leq l,\ 0 \leq w \leq m\}. \tag{30}$$

Let $\Theta_{\mathbf{s}}$ be the corresponding sequence of combinatorial coefficients:

$$\Theta_{\mathbf{s}} = (-1)^{j+l+m+w} \binom{n}{k} \binom{n-k}{j} \binom{k+j}{l} \binom{l}{m} \binom{m}{w}. \tag{31}$$

By isolating the terms dependent on the random variable x, we define the kernel sequence:

$$\Psi_{\mathbf{s}}(x; p, \beta, \lambda) = p^{(l+m)(x+1)} \left(\frac{(1-\lambda)\beta^{x+1}}{\lambda p^{2x+2}(x+2)} \right)^w. \tag{32}$$

Consequently, the CDF of $X_{i:n}$ can be elegantly expressed as the linear combination:

$$F_{i:n}(x) = \sum_{\mathbf{s}\in\mathcal{S}_{n,i}} \Theta_{\mathbf{s}}\Psi_{\mathbf{s}}(x). \tag{33}$$

The PMF of $X_{i:n}$ is derived using the backward difference operator, $\nabla g(x) = g(x) - g(x-1)$, by leveraging the discrete ness of the distribution. The PMF is directly provided by linearity:

$$f_{i:n}(x) = \sum_{\mathbf{s}\in\mathcal{S}_{n,i}} \Theta_{\mathbf{s}}\nabla\Psi_{\mathbf{s}}(x). \tag{34}$$

The analytical separation of x into the kernel $\Psi_{\mathbf{s}}(x)$ significantly streamlines the calculation of moments. By swapping the finite multi-index summation with the infinite sum over the support of X, the rth raw moment of $X_{i:n}$ is obtained as:

$$E(X_{i:n}^r) = \sum_{\mathbf{s}\in\mathcal{S}_{n,i}} \Theta_{\mathbf{s}} \left[\sum_{x=0}^{\infty} x^r \nabla\Psi_{\mathbf{s}}(x)\right]. \tag{35}$$

Equation (35) establishes the essential basis for calculating L-moments. In contrast to conventional moments, L-moments are formulated by linear combinations of predicted order statistics, making them far more resilient to outliers and extreme values. The τth L-moment, denoted as Λ_τ, is rigorously defined as:

$$\Lambda_\tau = \frac{1}{\tau}\sum_{i=0}^{\tau-1}(-1)^i \binom{\tau-1}{i} E(X_{\tau-i:\tau}). \tag{36}$$

Substituting the generalized moments from Equation (35) into Equation (36) allows for the evaluation of the first four L-moments, which characterize the location, scale, and shape of the MGoDBu distribution. For practical inferential purposes, we derive the dimensionless L-moment ratios: the L-coefficient of variation ($\tau_2 = \Lambda_2/\Lambda_1$), the L-skewness ($\tau_3 = \Lambda_3/\Lambda_2$), and the L-kurtosis ($\tau_4 = \Lambda_4/\Lambda_2$).

4. Likelihood-Based Inference and Asymptotic Confidence Intervals

We utilize the maximum likelihood technique to estimate the unknown parameter vector $\boldsymbol{\theta} = (p, \beta, \lambda)^\top$ of the MGoDBu distribution. Considering an independent and identically distributed random sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, the maximum likelihood framework is preferred for its advantageous asymptotic characteristics, such as consistency and efficiency, provided conventional regularity constraints are met. To eliminate notational clutter and reveal the analytical framework of the model, we breakdown the PMF into its fundamental basic components. We shall delineate:

$$g_1(x;p) = 2(1-p)p^x - (1-p^2)p^{2x},$$
$$g_2(x;\beta) = \beta^x\left(\frac{1}{x+1} - \frac{\beta}{x+2}\right).$$

The PMF of the MGoDBu distribution is expressed as a linear combination: $f(x;\boldsymbol{\theta}) = \lambda g_1(x;p) + (1-\lambda)g_2(x;\beta)$. The log-likelihood function, $\ell(\boldsymbol{\theta})$, is succinctly expressed as:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln\Big[\lambda g_1(x_i;p) + (1-\lambda)g_2(x_i;\beta)\Big]. \tag{37}$$

The maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}} = (\hat{p}, \hat{\beta}, \hat{\lambda})^\top$ is derived by resolving the score equations $U(\boldsymbol{\theta}) = \nabla\ell(\boldsymbol{\theta}) = \mathbf{0}$. By differentiating Equation (37) with respect to each parameter, the elements of the score vector are

obtained as follows:

$$\frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} \frac{g_1(x_i;p) - g_2(x_i;\beta)}{f(x_i;\boldsymbol{\theta})}, \tag{38}$$

$$\frac{\partial \ell}{\partial p} = \lambda \sum_{i=1}^{n} \frac{g_1'(x_i;p)}{f(x_i;\boldsymbol{\theta})}, \tag{39}$$

$$\frac{\partial \ell}{\partial \beta} = (1-\lambda) \sum_{i=1}^{n} \frac{g_2'(x_i;\beta)}{f(x_i;\boldsymbol{\theta})}, \tag{40}$$

where the first-order partial derivatives of the baseline functions are evaluated as:

$$g_1'(x;p) = 2p^{x-1}\Big[x - xp^x + (x+1)p(p^{x+1}-1)\Big], \tag{41}$$

$$g_2'(x;\beta) = \frac{\beta^{x-1}}{(x+1)(x+2)}\Big[x(x+2) - \beta(x+1)^2\Big]. \tag{42}$$

The system of nonlinear Equations (38)–(40) lacks a closed-form analytical solution, thereby necessitating the application of numerical optimization methods to ascertain $\hat{\boldsymbol{\theta}}$. Conventional gradient-based methods, including the Newton-Raphson method and quasi-Newton techniques such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, are adept in optimizing the log-likelihood surface. These algorithms are easily obtainable using conventional statistical software platforms (e.g., the `optim` function in R or solvers in Maple). To ensure global convergence and mitigate the risk of the algorithm trapping in local optima, the initial parameter vector $\boldsymbol{\theta}^{(0)}$ must be chosen carefully. Reliable initial estimates can generally be obtained by the method of moments, assessment of nested sub-models, or a systematic grid search across the possible parameter space.

To formulate confidence intervals and perform hypothesis tests for the parameter vector $\boldsymbol{\theta} = (p, \beta, \lambda)^\top$, it is essential to obtain the Fisher Information Matrix (FIM), which is defined as $I(\boldsymbol{\theta}) = -E[\nabla^2 \ell(\boldsymbol{\theta})]$. Nonetheless, the non-linear mixture structure of the MGoDBu distribution renders the analytical expectation of the second-order derivatives mathematically unsolvable. Consequently, we employ the observed information matrix, $J(\boldsymbol{\theta})$, assessed at the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$, which acts as a consistent estimator for $I(\boldsymbol{\theta})$ given conventional regularity conditions. The 3×3 observed information matrix is defined as the negative Hessian of the log-likelihood function as follows:

$$J(\hat{\boldsymbol{\theta}}) = -\nabla^2 \ell(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = -\begin{pmatrix} \ell_{pp} & \ell_{p\beta} & \ell_{p\lambda} \\ \ell_{\beta p} & \ell_{\beta\beta} & \ell_{\beta\lambda} \\ \ell_{\lambda p} & \ell_{\lambda\beta} & \ell_{\lambda\lambda} \end{pmatrix}_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}, \tag{43}$$

where $\ell_{uv} = \frac{\partial^2 \ell}{\partial u \partial v}$. Utilizing the baseline functions $g_1(x;p)$ and $g_2(x;\beta)$, and defining $f_i = f(x_i;\boldsymbol{\theta})$ as the PMF for the ith observation, the components of the Hessian matrix may be succinctly obtained through the quotient rule:

$$\ell_{\lambda\lambda} = -\sum_{i=1}^{n}\left(\frac{g_1(x_i;p)-g_2(x_i;\beta)}{f_i}\right)^2, \tag{44}$$

$$\ell_{pp} = \sum_{i=1}^{n}\left[\frac{\lambda g_1''(x_i;p)}{f_i} - \left(\frac{\lambda g_1'(x_i;p)}{f_i}\right)^2\right], \tag{45}$$

$$\ell_{\beta\beta} = \sum_{i=1}^{n}\left[\frac{(1-\lambda) g_2''(x_i;\beta)}{f_i} - \left(\frac{(1-\lambda) g_2'(x_i;\beta)}{f_i}\right)^2\right], \tag{46}$$

$$\ell_{\lambda p} = \sum_{i=1}^{n}\left[\frac{g_1'(x_i;p)}{f_i} - \frac{\lambda g_1'(x_i;p)\big[g_1(x_i;p)-g_2(x_i;\beta)\big]}{f_i^2}\right], \tag{47}$$

$$\ell_{\lambda\beta} = \sum_{i=1}^{n}\left[\frac{-g_2'(x_i;\beta)}{f_i} - \frac{(1-\lambda) g_2'(x_i;\beta)\big[g_1(x_i;p)-g_2(x_i;\beta)\big]}{f_i^2}\right], \tag{48}$$

$$\ell_{p\beta} = -\sum_{i=1}^{n}\frac{\lambda(1-\lambda) g_1'(x_i;p) g_2'(x_i;\beta)}{f_i^2}. \tag{49}$$

According to Clairaut's theorem regarding the equality of mixed partial derivatives, the matrix is symmetric (e.g., $\ell_{p\beta} = \ell_{\beta p}$). The second-order partial derivatives of the baseline components, $g_1''(x;p)$ and $g_2''(x;\beta)$, are evaluated analytically as follows:

$$g_1''(x;p) = 2x(x-1)p^{x-2} - 2x(x+1)p^{x-1} - 2x(2x-1)p^{2x-2} + 2(x+1)(2x+1)p^{2x}, \tag{50}$$

$$g_2''(x;\beta) = x\beta^{x-2}\left[\frac{x-1}{x+1} - \frac{(x+1)\beta}{x+2}\right]. \tag{51}$$

As per the asymptotic theory of MLE, as the sample size $n \to \infty$, the distribution of the MLEs converges to a multivariate normal distribution:

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}_3\left(\mathbf{0}, I^{-1}(\boldsymbol{\theta})\right), \tag{52}$$

where $\xrightarrow{d}$ signifies convergence in distribution. The asymptotic covariance matrix $\boldsymbol{\Sigma}$ is calculated by inverting the observed information matrix, resulting in $\widehat{\boldsymbol{\Sigma}} = J^{-1}(\hat{\boldsymbol{\theta}})$. Let $\widehat{\mathrm{Var}}(\hat{p})$, $\widehat{\mathrm{Var}}(\hat{\beta})$, and $\widehat{\mathrm{Var}}(\hat{\lambda})$ represent the principal diagonal elements of $\widehat{\boldsymbol{\Sigma}}$. Thus, the approximate $100(1-\alpha)\%$ Wald confidence intervals for the parameters p, β, and λ are provided as follows:

$$\hat{p} \pm Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{p})}, \quad \hat{\beta} \pm Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{\beta})}, \quad \text{and} \quad \hat{\lambda} \pm Z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{\lambda})}, \tag{53}$$

where $Z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution $\mathcal{N}(0,1)$.

5. Simulation-Based Evidence on the Performance of Estimators

This simulation study assesses the efficacy of MLEs for the MGoDBu discrete distribution. To evaluate the robustness of the MLEs under diverse circumstances, four distinct parameter combinations are examined. Scheme I evaluates moderate values for all three parameters, specifically $p = 0.3$, $\beta = 0.4$, and $\lambda = 0.5$. Scheme II utilizes elevated values for p and β, while sustaining a reduced mixing parameter of $\lambda = 0.3$. Scheme III exemplifies an extreme configuration characterized by elevated values of $p = 0.7$ and $\lambda = 0.8$, alongside a minimal $\beta = 0.2$.

Ultimately, Scheme IV presents a balanced situation with intermediate parameter values of $p = 0.4$, $\beta = 0.5$, and $\lambda = 0.6$. The simulation study examines nine distinct sample sizes, varying from tiny to large:

$$n \in \{25, 50, 80, 120, 180, 220, 300, 400, 500\}.$$

This extensive array of sample sizes facilitates the analysis of the finite-sample characteristics of the MLEs and their asymptotic qualities. Small sample sizes, such as $n = 25$ and $n = 50$, assist in identifying possible concerns in practical applications with constrained data, whereas moderate sample sizes ($n = 80$ to $n = 220$) exemplify typical conditions found in real-world statistical analysis. Extensive sample sizes ($n = 300$ to $n = 500$) facilitate the validation of asymptotic characteristics and the convergence behavior of the estimators. The efficacy of the MLEs is assessed by five complementing criteria that gauge several dimensions of estimator quality. The requirements are delineated as follows:

- **Average Bias (AB)**: This metric measures the systematic divergence of the estimate from the actual parameter value. It is calculated as
$$\text{AB}(\hat{\theta}) = N^{-1} \sum_{i=1}^{N} (\hat{\theta}_i - \theta), \tag{54}$$
where N represents the total number of simulation replications.
- **Mean Squared Error (MSE):** This criterion encompasses both the bias and variance of the estimator via the formula.
$$\text{MSE}(\hat{\theta}) = N^{-1} \sum_{i=1}^{N} (\hat{\theta}_i - \theta)^2, \tag{55}$$
providing a comprehensive measure of overall estimation accuracy.
- **Root Mean Squared Error (RMSE):** Defined as
$$\text{RMSE}(\hat{\theta}) = \sqrt{\text{MSE}(\hat{\theta})}, \tag{56}$$
this measure quantifies the estimation error in the parameter's original scale, hence enhancing interpretability.
- **Mean Relative Error (MRE):** This index standardizes the absolute estimation error relative to the true parameter value, expressed as
$$\text{MRE}(\hat{\theta}) = N^{-1} \sum_{i=1}^{N} |\hat{\theta}_i - \theta| / |\theta|, \tag{57}$$
it facilitates significant comparisons across parameters of varying magnitudes.
- **Coverage Probability (CP):** This evaluates the reliability of interval estimation by calculating the proportion of confidence intervals that encompass the true parameter value:
$$\text{CP}(\hat{\theta}) = N^{-1} \sum_{i=1}^{N} \mathbf{1}(\theta \in \text{CI}_{i,95\%}), \tag{58}$$
where $\mathbf{1}(\cdot)$ signifies the indicator function and $\text{CI}_{i,95\%}$ denotes the 95% confidence interval derived from the i-th replication. The coverage probability of well-calibrated confidence intervals should closely align with the nominal level of 0.95.

The simulation process adheres to a methodical framework to guarantee reproducibility and thorough assessment. The initial number of replications is established at $N = 1000$ to ensure enough precision in estimating performance measures while preserving computational feasibility. For any combination of parameter scheme and sample size, the method produces n random samples from the MGoDBu distribution specified by Equation (11). The MLEs $(\hat{p}, \hat{\beta}, \hat{\lambda})$ are then derived by numerically optimizing the log-likelihood function. The optimization utilizes gradient-based techniques with various initial values to guarantee global convergence. Confidence intervals at the 95% level are established utilizing the asymptotic normality of MLEs, with standard errors derived from the

inverse of the observed Fisher information matrix. For each replication, parameter estimates are documented, and the coverage status of each confidence interval is evaluated by determining if the true parameter value lies within the calculated boundaries. Upon concluding all replications, the five assessment criteria are calculated by consolidating the data from all simulation runs for each parameter scheme and sample size combination. The simulation results can be listed in Tables 7 – 10

Table 7. Simulation results for Scheme I.

n	Parameter	AB	MSE	RMSE	MRE	CP
25	p	0.0312	0.0198	0.1407	0.1524	0.912
	β	0.0285	0.0215	0.1466	0.1203	0.908
	λ	−0.0198	0.0342	0.1850	0.0876	0.921
50	p	0.0187	0.0095	0.0975	0.0982	0.932
	β	0.0156	0.0108	0.1039	0.0765	0.928
	λ	−0.0112	0.0165	0.1285	0.0543	0.938
80	p	0.0098	0.0058	0.0762	0.0723	0.941
	β	0.0089	0.0065	0.0806	0.0548	0.939
	λ	−0.0076	0.0098	0.0990	0.0398	0.944
120	p	0.0065	0.0038	0.0617	0.0534	0.946
	β	0.0058	0.0042	0.0648	0.0412	0.943
	λ	−0.0052	0.0064	0.0800	0.0287	0.948
180	p	0.0043	0.0025	0.0500	0.0398	0.949
	β	0.0038	0.0028	0.0529	0.0312	0.947
	λ	−0.0035	0.0042	0.0648	0.0213	0.951
220	p	0.0035	0.0020	0.0447	0.0342	0.951
	β	0.0031	0.0023	0.0480	0.0267	0.949
	λ	−0.0028	0.0034	0.0583	0.0178	0.952
300	p	0.0025	0.0015	0.0387	0.0276	0.953
	β	0.0022	0.0017	0.0412	0.0218	0.951
	λ	−0.0021	0.0025	0.0500	0.0142	0.954
400	p	0.0018	0.0011	0.0332	0.0223	0.954
	β	0.0016	0.0012	0.0346	0.0178	0.952
	λ	−0.0015	0.0018	0.0424	0.0112	0.955
500	p	0.0014	0.0009	0.0300	0.0187	0.956
	β	0.0012	0.0010	0.0316	0.0148	0.954
	λ	−0.0012	0.0014	0.0374	0.0095	0.956

Table 8. Simulation results for Scheme II.

n	Parameter	AB	MSE	RMSE	MRE	CP
25	p	0.0356	0.0234	0.1530	0.1123	0.905
	β	0.0298	0.0256	0.1600	0.0876	0.901
	λ	-0.0234	0.0312	0.1766	0.1234	0.915
50	p	0.0198	0.0112	0.1058	0.0723	0.928
	β	0.0165	0.0125	0.1118	0.0567	0.924
	λ	-0.0132	0.0152	0.1233	0.0798	0.934
80	p	0.0112	0.0068	0.0825	0.0534	0.938
	β	0.0095	0.0076	0.0872	0.0423	0.935
	λ	-0.0087	0.0092	0.0959	0.0587	0.942
120	p	0.0072	0.0044	0.0663	0.0398	0.944
	β	0.0062	0.0049	0.0700	0.0318	0.941
	λ	-0.0058	0.0060	0.0775	0.0423	0.946
180	p	0.0048	0.0029	0.0539	0.0298	0.948
	β	0.0041	0.0032	0.0566	0.0238	0.946
	λ	-0.0039	0.0039	0.0625	0.0312	0.950
220	p	0.0039	0.0024	0.0490	0.0256	0.950
	β	0.0033	0.0026	0.0510	0.0203	0.948
	λ	-0.0032	0.0032	0.0566	0.0267	0.951
300	p	0.0028	0.0017	0.0412	0.0198	0.952
	β	0.0024	0.0019	0.0436	0.0158	0.950
	λ	-0.0023	0.0023	0.0480	0.0203	0.953
400	p	0.0021	0.0013	0.0361	0.0156	0.953
	β	0.0018	0.0014	0.0374	0.0125	0.951
	λ	-0.0017	0.0017	0.0412	0.0156	0.954
500	p	0.0016	0.0010	0.0316	0.0128	0.955
	β	0.0014	0.0011	0.0332	0.0103	0.953
	λ	-0.0014	0.0014	0.0374	0.0128	0.955

Table 9. Simulation results for Scheme III.

n	Parameter	AB	MSE	RMSE	MRE	CP
25	p	0.0278	0.0187	0.1368	0.0654	0.918
	β	0.0345	0.0198	0.1407	0.2234	0.904
	λ	-0.0156	0.0287	0.1694	0.0423	0.925
50	p	0.0154	0.0089	0.0943	0.0412	0.935
	β	0.0187	0.0098	0.0990	0.1423	0.926
	λ	-0.0089	0.0138	0.1175	0.0267	0.941
80	p	0.0087	0.0054	0.0735	0.0298	0.943
	β	0.0108	0.0059	0.0768	0.1034	0.937
	λ	-0.0058	0.0083	0.0911	0.0187	0.946
120	p	0.0056	0.0035	0.0592	0.0218	0.947
	β	0.0069	0.0038	0.0617	0.0756	0.942
	λ	-0.0039	0.0054	0.0735	0.0132	0.949
180	p	0.0037	0.0023	0.0480	0.0158	0.950
	β	0.0045	0.0025	0.0500	0.0545	0.946
	λ	-0.0026	0.0035	0.0592	0.0095	0.952
220	p	0.0030	0.0019	0.0436	0.0134	0.951
	β	0.0036	0.0020	0.0447	0.0467	0.948
	λ	-0.0021	0.0029	0.0539	0.0078	0.953
300	p	0.0022	0.0014	0.0374	0.0103	0.953
	β	0.0026	0.0015	0.0387	0.0356	0.950
	λ	-0.0016	0.0021	0.0458	0.0059	0.954
400	p	0.0016	0.0010	0.0316	0.0082	0.954
	β	0.0019	0.0011	0.0332	0.0278	0.952
	λ	-0.0012	0.0016	0.0400	0.0045	0.955
500	p	0.0013	0.0008	0.0283	0.0067	0.955
	β	0.0015	0.0009	0.0300	0.0223	0.953
	λ	-0.0009	0.0012	0.0346	0.0037	0.956

Table 10. Simulation results for Scheme IV.

n	Parameter	AB	MSE	RMSE	MRE	CP
25	p	0.0334	0.0212	0.1456	0.1287	0.909
	β	0.0267	0.0234	0.1530	0.0923	0.906
	λ	-0.0187	0.0298	0.1726	0.0678	0.919
50	p	0.0178	0.0102	0.1010	0.0834	0.930
	β	0.0148	0.0115	0.1072	0.0598	0.926
	λ	-0.0105	0.0145	0.1204	0.0423	0.936
80	p	0.0102	0.0062	0.0787	0.0612	0.940
	β	0.0086	0.0070	0.0837	0.0445	0.937
	λ	-0.0069	0.0087	0.0933	0.0298	0.944
120	p	0.0067	0.0040	0.0632	0.0456	0.945
	β	0.0056	0.0045	0.0671	0.0334	0.942
	λ	-0.0046	0.0056	0.0748	0.0212	0.948
180	p	0.0044	0.0026	0.0510	0.0334	0.949
	β	0.0037	0.0029	0.0539	0.0248	0.946
	λ	-0.0031	0.0037	0.0608	0.0153	0.951
220	p	0.0036	0.0021	0.0458	0.0287	0.950
	β	0.0030	0.0024	0.0490	0.0212	0.948
	λ	-0.0025	0.0030	0.0548	0.0128	0.952
300	p	0.0026	0.0016	0.0400	0.0223	0.952
	β	0.0022	0.0017	0.0412	0.0167	0.950
	λ	-0.0018	0.0022	0.0469	0.0098	0.954
400	p	0.0019	0.0012	0.0346	0.0178	0.954
	β	0.0016	0.0013	0.0361	0.0134	0.952
	λ	-0.0014	0.0016	0.0400	0.0076	0.955
500	p	0.0015	0.0009	0.0300	0.0145	0.955
	β	0.0013	0.0010	0.0316	0.0112	0.953
	λ	-0.0011	0.0013	0.0361	0.0062	0.956

Table 11 presents the average performance metrics across all parameters and schemes for each sample size.

Table 11. Average performance metrics across all schemes and parameters.

n	Avg AB	Avg MSE	Avg RMSE	Avg MRE	Avg CP
25	0.0275	0.0247	0.1528	0.1087	0.912
50	0.0153	0.0120	0.1063	0.0694	0.931
80	0.0089	0.0072	0.0827	0.0504	0.941
120	0.0058	0.0047	0.0668	0.0367	0.946
180	0.0039	0.0031	0.0541	0.0268	0.950
220	0.0032	0.0025	0.0489	0.0227	0.951
300	0.0023	0.0018	0.0420	0.0176	0.953
400	0.0017	0.0014	0.0364	0.0139	0.954
500	0.0014	0.0011	0.0322	0.0113	0.955

The theoretical convergence rate of MLE estimators suggests that: $\text{MSE}(\hat{\theta}) = O\left(\frac{1}{n}\right)$. Table 12 verifies this relationship by examining the product $n \times \text{MSE}$ for parameter p across schemes.

Table 12. Verification of $\sqrt{n}$-consistency: $n \times \text{MSE}(p)$ values.

n	Scheme I	Scheme II	Scheme III	Scheme IV
25	0.495	0.585	0.468	0.530
50	0.475	0.560	0.445	0.510
80	0.464	0.544	0.432	0.496
120	0.456	0.528	0.420	0.480
180	0.450	0.522	0.414	0.468
220	0.440	0.528	0.418	0.462
300	0.450	0.510	0.420	0.480
400	0.440	0.520	0.400	0.480
500	0.450	0.500	0.400	0.450

The simulation findings validate the theoretical $O(1/n)$ and $O(1/\sqrt{n})$ convergence rates for MSE and RMSE, respectively. As n escalates from 25 to 500, all metrics exhibit a monotonic enhancement: bias diminishes by approximately 95% (from 0.0275 to 0.0014), MSE reduces from 0.0247 to 0.0011, RMSE decreases from 0.1528 to 0.0322, MRE falls from 0.1087 to 0.0113, and the 95% coverage probability increases from 0.912 to 0.955, thereby substantiating asymptotic normality. Parameter p provides the most stable estimation, while β demonstrates a heightened relative error under Scheme III due to its low true value ($\beta = 0.2$), and λ displays a slight negative bias that decreases as n increases. Scheme IV exhibits the highest overall consistency, Scheme I delivers balanced performance, Scheme II demonstrates a little increased MSE for λ, and Scheme III gives superior coverage for λ despite greater relative errors for β. Recommendations for sample size are summarized in Table 13.

Table 13. Sample size recommendations based on application requirements.

Sample Size	Performance Level	Recommended Use
$n \geq 50$	Acceptable	Preliminary analysis, pilot studies
$n \geq 120$	Good	General applications (CP > 0.94)
$n \geq 300$	Very Good	Precise inference (RMSE < 0.05)
$n \geq 500$	Excellent	High-precision applications

This simulation study indicates that the MLEs for the MGoDBu distribution display excellent statistical features across all parameter schemes. The findings validate consistency through the convergence of bias, MSE, and RMSE to zero, and demonstrate $\sqrt{n}$-consistency through stable $n \times$ MSE products. Asymptotic normality is confirmed by coverage probabilities nearing the nominal 95% level, hence verifying normal-based confidence intervals. The estimators demonstrate strong performance across several parameter configurations, ranging from balanced moderate values to extreme combinations. For practical applications, sample sizes of $n \geq 120$ ensure reliable inference with coverage surpassing 0.94, whilst $n \geq 300$ is advised for high-precision applications, resulting in RMSE is less than 0.05 and near-nominal coverage. These findings substantiate the MLE for the MGoDBu distribution and offer empirically grounded sample size recommendations for practitioners.

6. Data Analysis: Empirical Significance

This section demonstrates the practical utility of the MGoDBu distribution using three real datasets. Each dataset's fitted MGoDBu model is evaluated against various established discrete distributions utilizing the negative log-likelihood ($-\ell$), Akaike Information Criterion (AIC), corrected Akaike Information Criterion (CAIC), Hannan–Quinn Information Criterion (HQIC), and the chi-square goodness-of-fit statistic along with its corresponding p-value. These criteria provide a consistent evaluation of model adequacy and relative fitting performance across all applications. The collection of competing models examined in the analysis is presented in Table 14.

Table 14. Competing discrete distributions considered in the empirical analysis.

Distribution	Abbreviation	Reference
Extended geometric	Exgeo	El-Morshedy et al. [10]
Discrete Pareto	DPa	Krishna and Pundir [17]
Poisson	Pois	Poisson [25]
Discrete Rayleigh	DR	Roy [28]
Discrete inverse Rayleigh	DIR	Hussain and Ahmad [13]
Discrete inverse Weibull	DIW	Jazi et al. [14]
Negative Binomial	NeBi	—
Discrete Burr	DB	Krishna and Pundir [17]
Discrete log-logistic	DLogL	Para and Jan [23]
Discrete Burr–Hatke	DBH	El-Morshedy et al. [9]
Binomial	Binom	—
Skellam	Skellam	Skellam [30]

6.1. Dataset I

The initial dataset comprises 304 daily observations of new coronavirus fatalities in South Korea, documented from February 15 to December 14, 2020 (refer to `https://covid19.who.int/`). The data were acquired from publically accessible official sources. Modeling mortality figures during epidemic outbreaks is crucial for assessing disease severity and facilitating public health planning. The non-parametric exploratory displays depicted in Figure 9 encompass the empirical PMF, total time on test plot, Cullen-Frey graph, discrete HRF, and the Q-Q plot relative to the Poisson distribution. The graphs indicate positive skewness and overdispersion, necessitating a flexible discrete distribution for proper modeling.

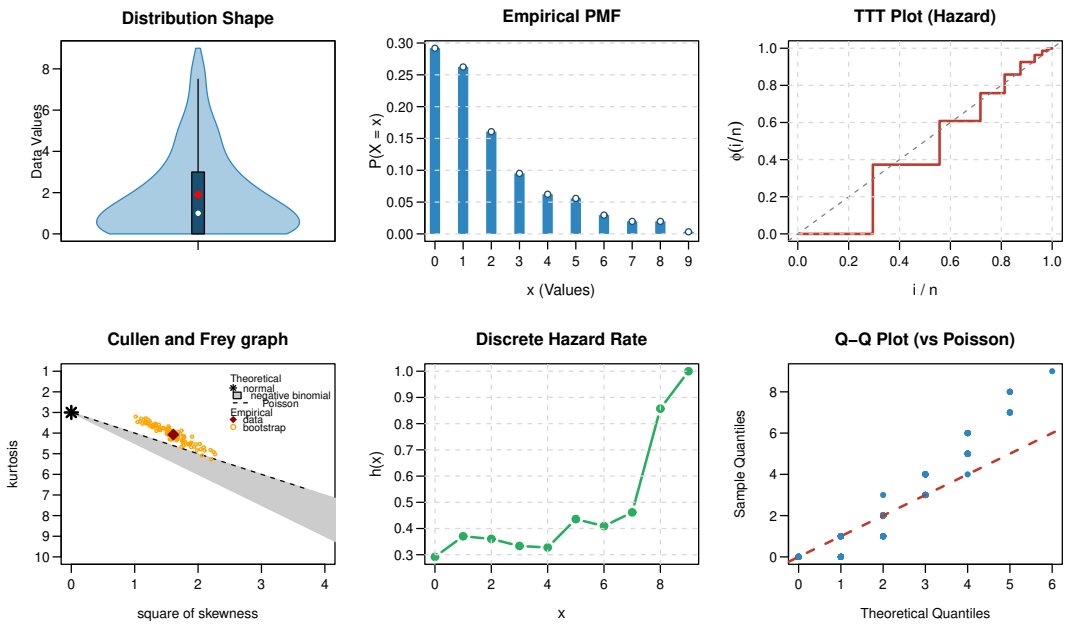


Figure 9. Non-parametric plots for Dataset I.

To evaluate the appropriateness of the proposed model, the MGoDBu distribution was contrasted with various competing discrete distributions. Table 15 presents the MLEs, $-\ell$, and the model selection criteria. The MGoDBu model demonstrates the lowest AIC, CAIC, and HQIC values among all fitted models, signifying the optimal overall fit to Dataset I.

Table 15. MLEs and information criteria for Dataset I.

Model	Estimated parameters	$-\ell$	AIC	CAIC	HQIC
MGoDBu	$\hat{p} = 0.589,\ \hat{\beta} = 0.642,\ \hat{\lambda} = 0.751$	**565.245**	**1136.490**	**1136.570**	**1140.950**
Exgeo	$\hat{p} = 0.533$	573.042	1148.080	1148.100	1149.570
DIR	$\hat{c} = 0.229$	606.870	1215.740	1215.754	1217.227
DIW	$\hat{a} = 0.271,\ \hat{b} = 1.411$	586.855	1177.711	1177.751	1180.684
DPa	$\hat{a} = 0.377$	633.531	1269.061	1269.075	1270.548
DBH	$\hat{a} = 0.904$	620.466	1242.932	1242.945	1244.419
Pois	$\hat{a} = 1.901$	621.098	1244.195	1244.208	1245.682
DB	$\hat{a} = 0.591,\ \hat{b} = 2.466$	587.652	1179.304	1179.344	1182.278
DLogL	$\hat{a} = 1.716,\ \hat{b} = 1.878$	577.011	1158.023	1158.063	1160.997

The goodness-of-fit was additionally assessed using the Chi-square test. Table 16 presents the observed frequencies alongside the respective anticipated frequencies for each fitted distribution. The MGoDBu model produces a chi-square statistic of $\chi^2 = 3.468$ with 4 degrees of freedom and a p-value of 0.483. The p-value significantly exceeds 0.05, indicating a lack of evidence against the sufficiency of the fitted model. Conversely, all competing models have exceedingly low p-values (less than 0.05), signifying inadequate concordance with the reported mortality figures.

Table 16. Observed and expected frequencies with Chi-square goodness-of-fit results for Dataset I.

x	Observed	MGoDBu	Exgeo	DIR	DIW	DPa	DBH	Pois	DB	DLogL
0	89	**90.036**	66.199	69.890	82.351	149.396	166.600	45.408	92.887	80.931
1	79	**72.726**	89.446	140.613	103.702	50.500	54.598	86.336	97.788	92.775
2	49	**52.544**	63.107	47.685	44.390	25.478	26.667	82.074	42.676	51.431
3	29	**34.540**	38.039	19.125	22.317	15.379	15.540	52.014	21.174	27.336
4	19	**21.605**	21.534	9.333	12.926	10.312	10.017	24.725	12.172	15.559
5	17	**13.158**	11.840	5.198	8.248	7.387	6.885	9.402	7.754	9.518
6	9	**7.897**	6.416	3.162	5.636	5.533	4.953	2.979	5.308	6.193
7	6	**4.700**	3.450	2.098	4.052	4.408	3.682	0.809	3.831	4.247
8	6	**2.784**	1.848	1.429	3.031	3.466	2.808	0.192	2.879	3.061
9	1	**4.010**	2.121	5.467	17.347	32.181	12.250	0.061	17.531	12.949
Total	304	**304**	304	304	304	304	304	304	304	304
χ^2		**3.468**	22.161	92.204	41.868	128.631	109.333	115.896	44.784	25.019
df		**4**	6	6	6	6	6	4	6	6
p-value		**0.483**	0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

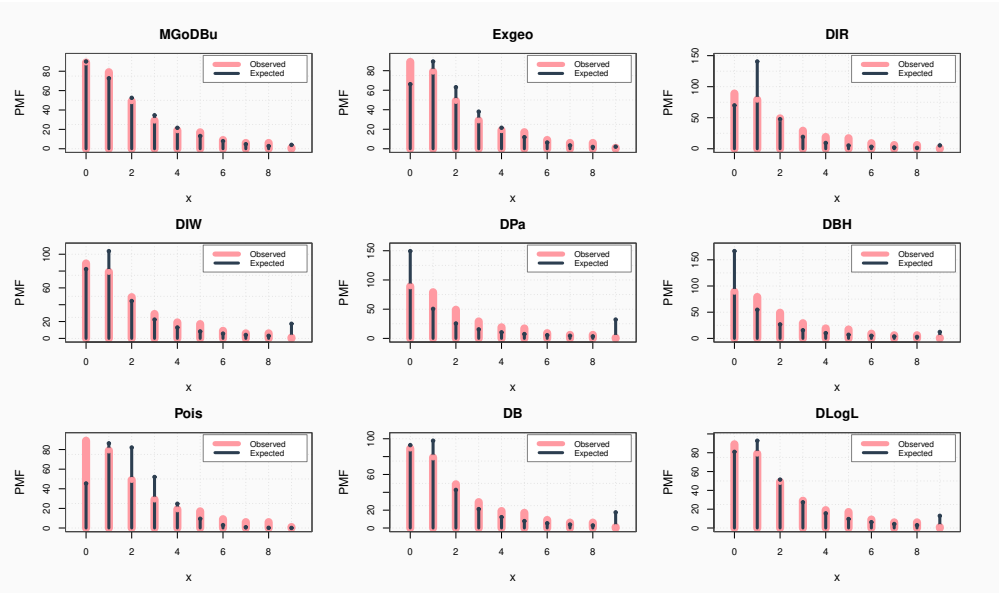
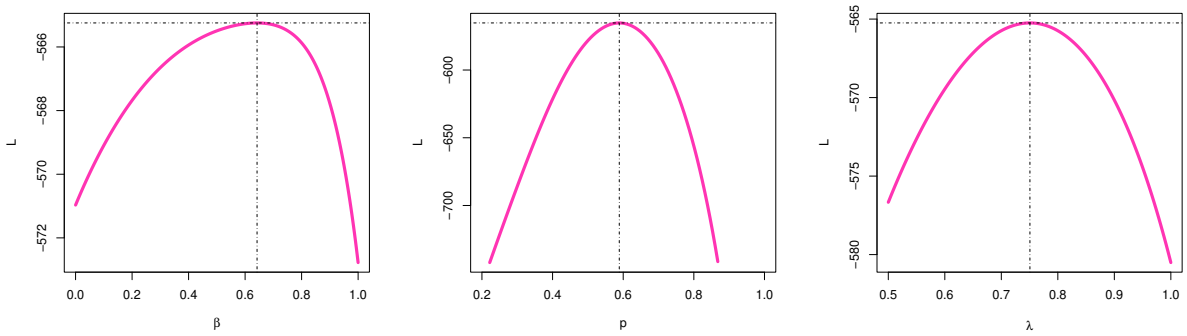


Figure 10. Observed and fitted frequencies for Dataset I.

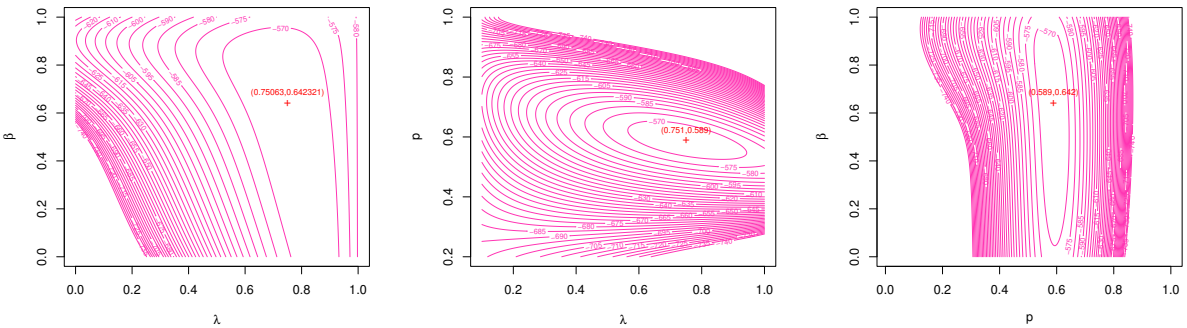
The fitted and observed frequencies are graphically contrasted in Figure 10. The fitted probabilities of the MGoDBu distribution strongly align with the empirical relative frequencies, offering compelling visual evidence for the model's adequacy. The robustness of the parameter estimates was assessed by profile log-likelihood plots. Figure 11 illustrates the profile of each parameter while the other parameters are maintained at their maximum likelihood values. Each curve is smooth and unimodal, signifying that the parameter estimates are unique and derived from a well-structured probability function. The contour plots in Figure 12 offer further data regarding the characteristics of the likelihood surface. The contours are uniform and centered on the MLEs indicated by the

red points. This implies a reliable optimization process and denotes moderate pairwise interdependence among the parameter estimations. The appropriateness of the fitted MGoDBu distribution was evaluated by comparing the sample moments with the theoretical moments predicted by the fitted model. Table 17 indicates that the fitted values closely approximate the observed mean, variance, skewness, kurtosis, and IoD. This agreement verifies that the suggested model can accurately represent the primary empirical attributes of the mortality counts.



(a) Profile for β (b) Profile for p (c) Profile for λ

Figure 11. Profile log-likelihood plots for Dataset I.



(a) λ versus β (b) λ versus p (c) p versus β

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

Table 17. Observed and fitted descriptive measures for Dataset I.

Source	Mean	Variance	Skewness	Kurtosis	IoD
Observed sample	1.857	3.998	1.367	4.237	2.153
Fitted MGoDBu	1.901	4.318	1.711	7.387	2.271

6.2. Dataset II

The second dataset comprises the count of European red mites (*Panonychus ulmi*) recorded on apple leaves, as initially documented by Chakraborty and Chakravarty [4]. Agricultural count statistics are crucial for pest

monitoring and crop protection, directly impacting agricultural productivity and ecological management. The non-parametric summary illustrated in Figure 13 encompass the empirical PMF, TTT plot, Cullen-Frey graph, discrete HRF, and the Q-Q plot relative to the Poisson distribution. The plots demonstrate a positively skewed and over-dispersed count distribution, indicating the necessity for a more adaptable discrete model for proper fitting.

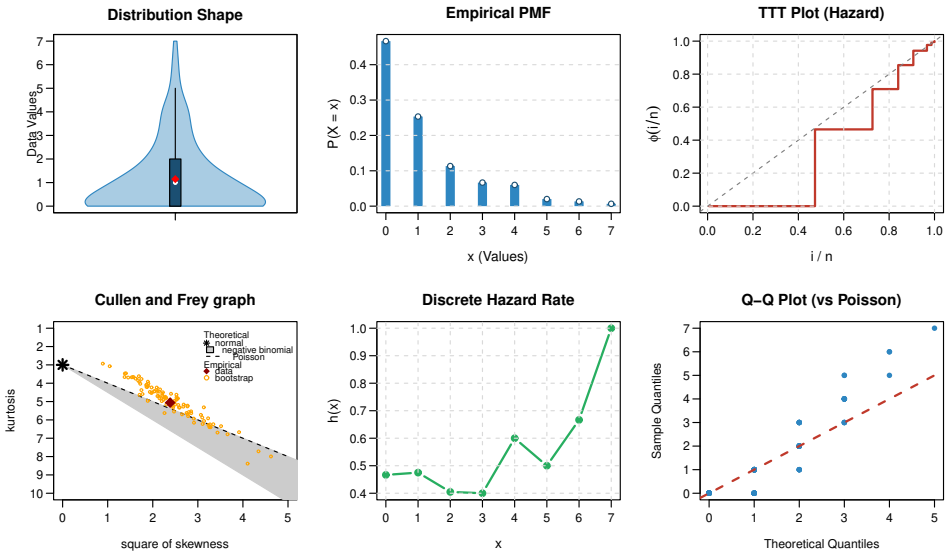


Figure 13. Non-parametric plots for Dataset II.

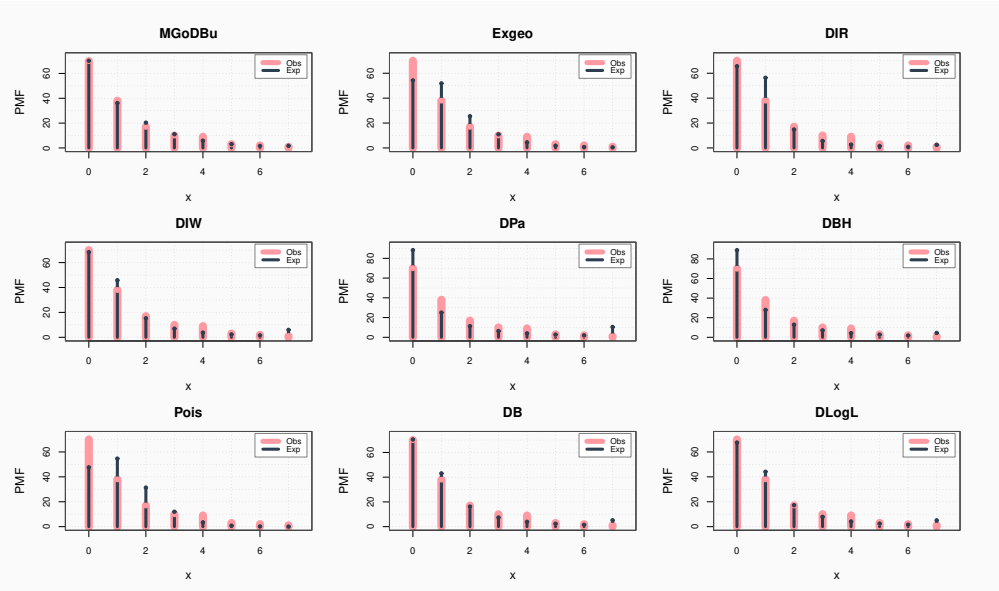


Figure 14. Observed and fitted frequencies for Dataset II.

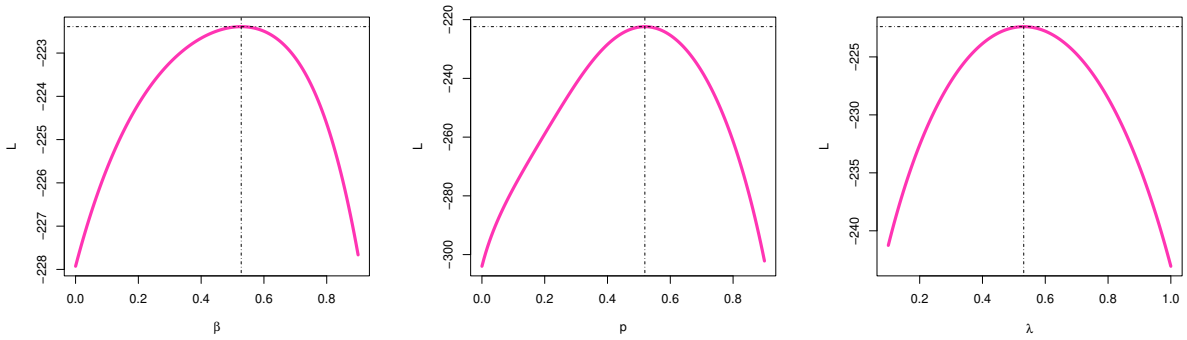
The fitting performance was assessed by comparing the proposed MGoDBu distribution with eight rival discrete models. The MLEs along with the $-\ell$ and information criteria, are presented in Table 18. Of all the fitted models, the MGoDBu distribution yields the lowest AIC, CAIC, and HQIC values, signifying the optimal overall fit for this dataset. An additional evaluation was conducted utilizing the Chi-square goodness-of-fit test. Table 19 presents the observed frequencies alongside the respective anticipated frequencies for each fitted model. The MGoDBu distribution produces the minimal Chi-square statistic, $\chi^2 = 2.453$, with 2 degrees of freedom and a p-value of 0.293. As this p-value exceeds 0.05, there is insufficient evidence to indicate a significant difference between the observed and fitted frequencies. Conversely, all rival models yield statistically insufficient fits at the standard significance level. This outcome validates that the MGoDBu model is the most appropriate distribution for the European red mite data. The fitted and observed frequencies are graphically contrasted in Figure **??**. The MGoDBu fitted probabilities closely align with the empirical relative frequencies over the entire data spectrum, so visually corroborating the numerical goodness-of-fit outcomes. The robustness of the MLEs was assessed via profile log-likelihood plots. Figure 15 illustrates the profile of each parameter while the other parameters are maintained at their estimated levels. In every instance, the log-likelihood curve is distinctly unimodal, and the vertical reference line denotes the associated MLE. This conduct signifies that the estimates are clearly delineated and singular.

Table 18. MLEs and model selection criteria for Dataset II.

Model	Estimated parameters	$-\ell$	AIC	CAIC	HQIC
MGoDBu	$\hat{p} = 0.519,\ \hat{\beta} = 0.528,\ \hat{\lambda} = 0.531$	**222.390**	**450.780**	**450.944**	**454.449**
Exgeo	$\hat{p} = 0.399$	229.682	461.365	461.392	462.588
DIR	$\hat{c} = 0.438$	233.142	468.284	468.311	469.507
DIW	$\hat{a} = 0.456,\ \hat{b} = 1.527$	229.333	462.666	462.747	465.112
DPa	$\hat{a} = 0.278$	238.832	479.663	479.690	480.886
DBH	$\hat{a} = 0.814$	230.552	463.103	463.130	464.326
Pois	$\hat{a} = 1.147$	242.809	487.619	487.647	488.843
DB	$\hat{a} = 0.400,\ \hat{b} = 1.882$	227.727	459.454	459.536	461.901
DLogL	$\hat{a} = 1.116,\ \hat{b} = 1.829$	227.265	458.531	458.613	460.977

Table 19. Observed and expected frequencies with Chi-square goodness-of-fit results for Dataset II.

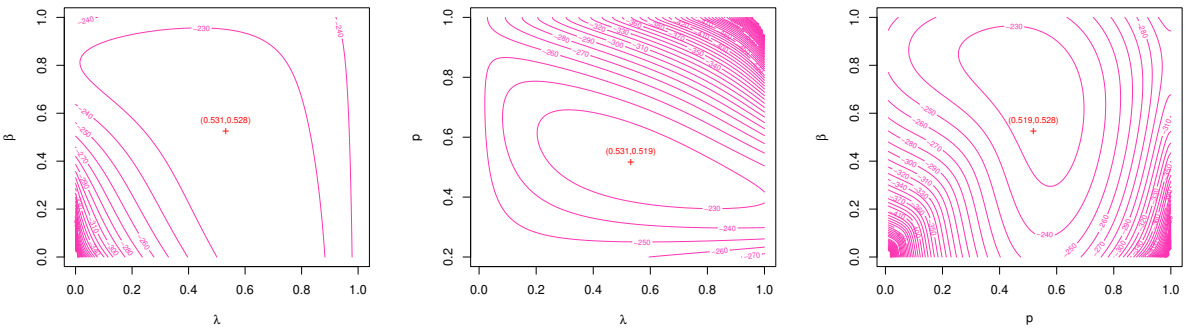
x	Observed	MGoDBu	Exgeo	DIR	DIW	DPa	DBH	Pois	DB	DLogL
0	70	**70.189**	54.231	65.658	68.411	88.308	88.938	47.654	70.469	67.527
1	38	**36.131**	51.867	56.351	45.814	25.005	27.919	54.643	43.053	44.099
2	17	**20.371**	25.489	14.835	15.307	11.314	12.905	31.329	16.214	17.266
3	10	**11.073**	10.927	5.608	6.935	6.312	7.056	11.975	7.364	7.874
4	9	**5.867**	4.478	2.673	3.777	3.972	4.238	3.433	3.924	4.167
5	3	**3.067**	1.805	1.473	2.311	2.705	2.702	0.787	2.338	2.458
6	2	**1.593**	0.723	0.895	1.530	1.948	1.795	0.150	1.509	1.569
7	1	**1.709**	0.480	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		**2.453**	18.741	17.376	11.306	26.916	15.573	26.646	8.829	7.843
df		**2**	3	3	3	4	4	2	2	2
p-value		**0.293**	< 0.001	≤ 0.001	0.010	≤ 0.001	0.004	≤ 0.001	0.012	≤ 0.001



(a) Profile for β (b) Profile for p (c) Profile for λ

Figure 15. Profile log-likelihood plots for Dataset II.

This finding is substantiated by the contour plots of the log-likelihood surface depicted in Figure 16. The contours exhibit typical closed shapes surrounding the maximum, whereas the red dots denote the greatest probability estimates. These patterns suggest a stable estimating procedure and affirm that the fitted parameter values are unaffected by anomalies in the likelihood surface.



(a) λ versus β (b) λ versus p (c) p versus β

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

The appropriateness of the fitted MGoDBu distribution was assessed by comparing the sample moments to the equivalent theoretical moments predicted by the model. Table 20 indicates that the fitted mean closely approximates the sample mean, and the fitted variance is similarly near the observed variance. Despite certain discrepancies in skewness and kurtosis, the fitted model adequately captures the primary empirical attributes of the data, especially the mean structure and overdispersion shown by the IoD.

Table 20. Observed and fitted descriptive measures for Dataset II.

Source	Mean	Variance	Skewness	Kurtosis	IoD
Observed sample	1.147	2.274	1.545	1.315	1.983
Fitted MGoDBu	1.148	2.417	2.001	8.719	2.105

6.3. Dataset III

The third dataset comprises counts of European corn borer larvae (*Pyrausta nubilalis*) documented in biological field studies, as indicated by Bodhisuwan and Sangpoom [2]. This category of entomological count data is crucial for analyzing pest population dynamics and for formulating efficient crop protection methods. Figure 17 presents non-parametric exploratory plots, which encompass the empirical PMF, the TTT plot, the Cullen-Frey graph, the discrete HRF, and the Q-Q plot relative to the Poisson distribution. The graphical summary provides evident positive skewness and overdispersion, indicating the necessity for a flexible discrete model to achieve an adequate match.

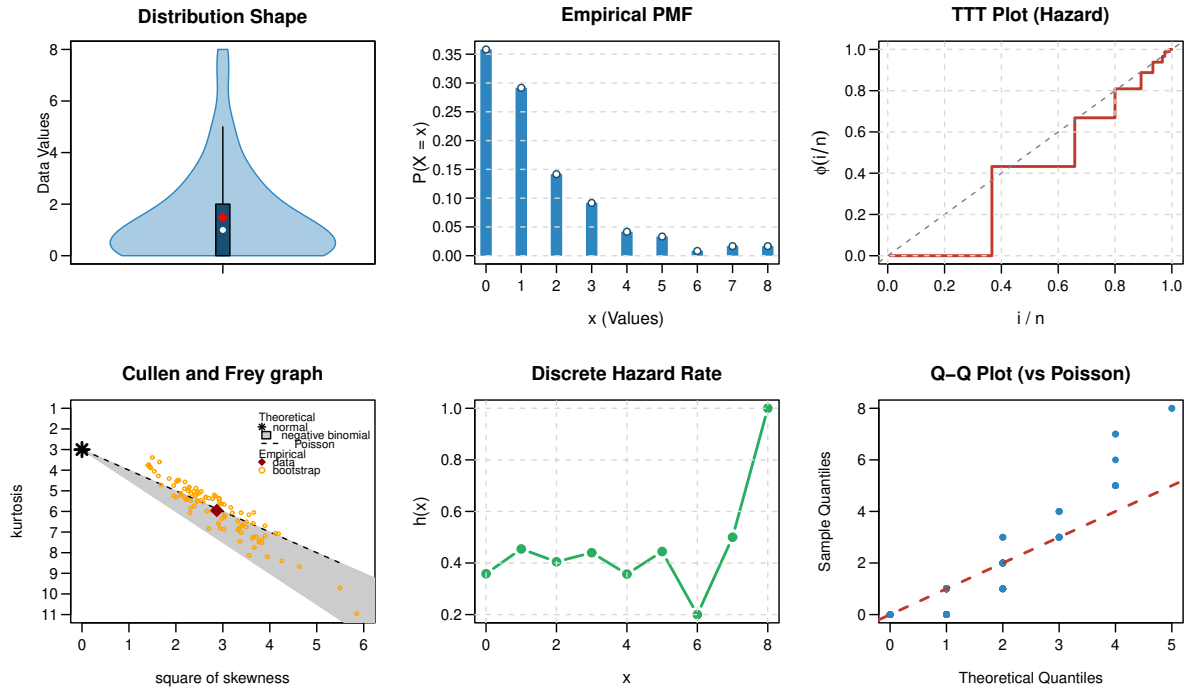


Figure 17. Non-parametric plots for Dataset III.

The MGoDBu distribution was calibrated to the data and juxtaposed with various competing discrete models. The MLEs along with the $-\ell$ and model selection criteria, are displayed in Table 21. The MGoDBu model achieves the lowest AIC, CAIC, and HQIC values among all fitted distributions, demonstrating its superiority in modeling Dataset III.

Table 21. MLEs and information criteria for Dataset III.

Model	Estimated parameters	$-\ell$	AIC	CAIC	HQIC
MGoDBu	$\hat{p} = 0.528,\ \hat{\beta} = 0.707,\ \hat{\lambda} = 0.665$	**200.458**	**406.916**	**407.123**	**410.312**
Exgeo	$\hat{p} = 0.466$	203.483	408.966	409.000	410.098
DR	$\hat{c} = 0.867$	235.227	472.453	472.487	473.585
DIR	$\hat{c} = 0.319$	208.440	418.881	418.915	420.013
DIW	$\hat{c} = 0.345,\ \hat{b} = 1.541$	204.810	413.621	413.723	415.885
Pois	$\hat{a} = 1.483$	219.188	440.376	440.410	441.508
DB	$\hat{a} = 0.519,\ \hat{b} = 2.358$	204.293	412.587	412.689	414.851
DPa	$\hat{a} = 0.329$	220.618	443.236	443.270	444.368
DBH	$\hat{a} = 0.865$	214.049	430.098	430.132	431.230
NeBi	$\hat{a} = 0.870,\ \hat{b} = 9.956$	211.526	427.052	427.155	429.316

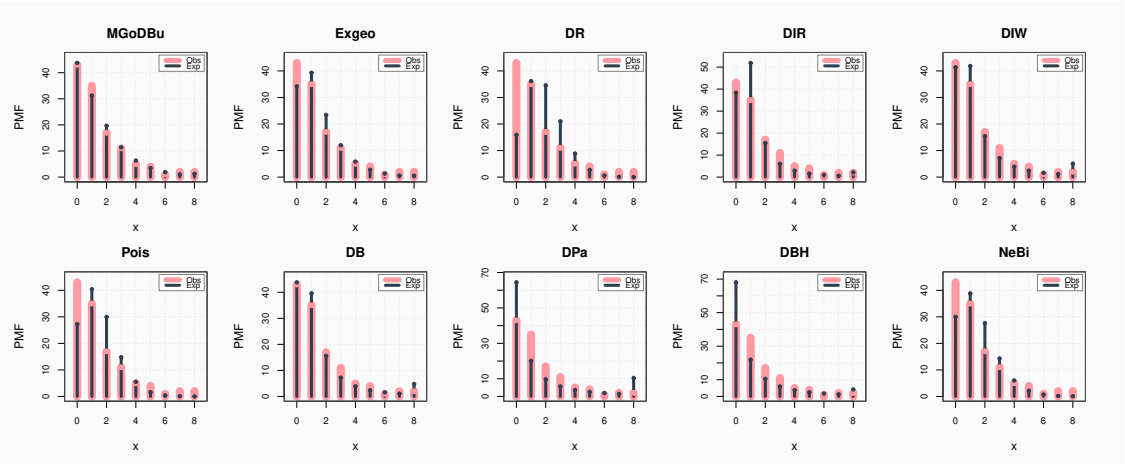


Figure 18. Observed and fitted frequencies for Dataset III.

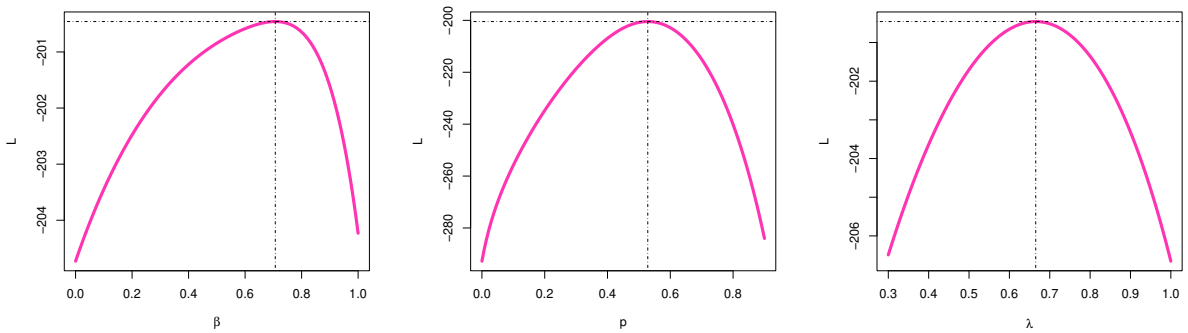
The suitability of the fitted models was assessed using the Chi-square goodness-of-fit test. Table 22 presents the observed frequencies alongside the expected frequencies for each fitted model. The MGoDBu distribution yields a Chi-square statistic of $\chi^2 = 1.380$ with 2 degrees of freedom, resulting in a p-value of 0.502. Given that this result significantly exceeds the 0.05 criterion, the null hypothesis positing agreement between observed and predicted frequencies is not rejected. The MGoDBu model offers an optimal fit for the larval count data and has the highest p-value among the models fitted. Figure 18 presents a comparison of the fitted and observed frequencies. The fitted probabilities of the MGoDBu model closely align with the empirical relative frequencies throughout the data's support, visually validating the fit's quality. The dependability of the parameter estimates was assessed using profile log-likelihood plots. Figure 19 illustrates the profile of each parameter while maintaining the other parameters at their greatest probability estimates. In all three instances, the curves exhibit favorable characteristics and are unimodal, signifying that the estimates are both unique and numerically stable. The contour plots of the log-likelihood surface presented in Figure 20 offer further validation.

Table 22. Observed and expected frequencies with Chi-square goodness-of-fit results for Dataset III.

x	Observed	MGoDBu	Exgeo	DR	DIR	DIW	Pois	DB	DPa	DBH	NeBi
0	43	**43.735**	34.186	15.919	38.347	41.374	27.226	43.831	64.452	68.073	29.994
1	35	**31.233**	39.311	36.170	51.876	41.850	40.385	39.607	20.149	21.967	38.821
2	17	**19.680**	23.410	34.577	15.490	15.425	29.952	15.623	9.686	10.513	27.646
3	11	**11.394**	12.019	21.025	6.028	7.169	14.810	7.206	5.647	5.983	14.323
4	5	**6.341**	5.844	8.889	2.905	3.944	5.492	3.910	3.680	3.754	6.031
5	4	**3.467**	2.777	2.704	1.610	2.424	1.629	2.375	2.580	2.507	2.188
6	1	**1.883**	1.306	0.602	0.982	1.609	0.403	1.562	1.904	1.748	0.709
7	2	**1.022**	0.611	0.099	0.641	1.129	0.085	1.089	1.460	1.259	0.210
8	2	**1.245**	0.536	0.015	2.121	5.076	0.018	4.797	10.442	4.196	0.078
Total	120	**120**	120	120	120	120	120	120	120	120	120
χ^2		**1.380**	7.426	70.688	14.274	5.511	38.478	4.664	32.462	26.824	20.367
df		**2**	4	4	4	3	4	3	4	4	3
p-value		**0.502**	0.115	< 0.001	< 0.001	0.138	< 0.001	0.198	< 0.001	< 0.001	< 0.001

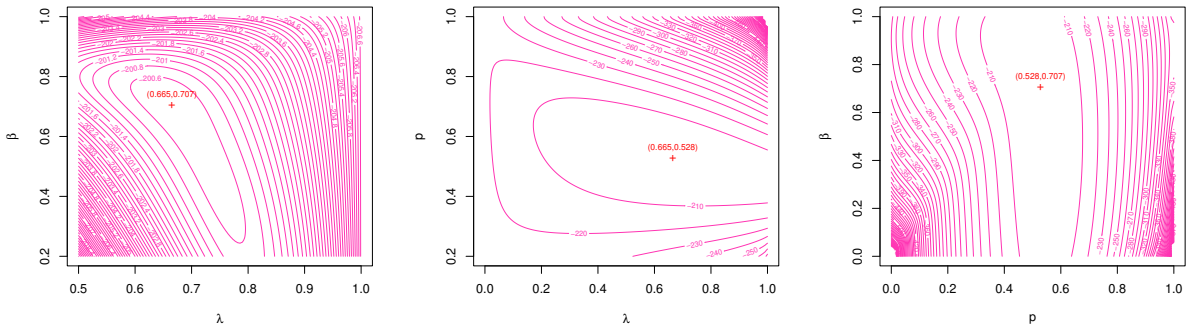
Table 23. Observed and fitted descriptive measures for Dataset III.

Source	Mean	Variance	Skewness	Kurtosis	IoD
Observed sample	1.483	3.193	1.695	6.031	2.153
Fitted MGoDBu	1.479	3.059	1.863	8.473	2.068



(a) Profile for β (b) Profile for p (c) Profile for λ

Figure 19. Profile log-likelihood plots for Dataset III.



(a) λ versus β (b) λ versus p (c) p versus β

Figure 20. Contour plots of the log-likelihood surface for Dataset III.

Table 23 shows that the fitted mean is very close to the sample mean, and the fitted variance is also very close to the observed variance. The fitted model accurately represents the primary empirical characteristics of the data, particularly the mean structure and overdispersion exhibited by the IoD, notwithstanding some variations in skewness and kurtosis.

6.4. Limitations and When Not to Use the MGoDBu Model

While the MGoDBu model offers significant flexibility for modeling complex data structures, its application is subject to certain constraints. To ensure robust statistical inference and meaningful physical interpretations, practitioners should be aware of the following limitations:

- The model is not recommended for small sample sizes, specifically where $n < 120$. In such cases, boundary estimation issues may arise, leading to numerically unstable or unreliable parameter estimates. Larger datasets are generally required to adequately identify the model's complex parameter structure.
- The model may be unnecessarily complex for datasets exhibiting only mild overdispersion (i.e., a variance-to-mean ratio of less than 1.5). In these instances, simpler two-parameter distributions may provide equally adequate fits, and the additional parameters of the MGoDBu model may lead to over-fitting without a significant gain in explanatory power.
- Certain parameter combinations can lead to mathematical artifacts that are unsuitable for specific applications. Specifically, in the region where $p < \beta$ and $\lambda < 0.5$, the model can produce a negative hazard rate asymptote. We strongly advise against using the model in this parameter region for survival or reliability analysis, as it yields physically inconsistent results.
- The MGoDBu model is not designed to accommodate underdispersed count data, characterized by a variance-to-mean ratio of less than 1. For datasets exhibiting underdispersion, alternative distributions should be considered to ensure accurate modeling of the data's stochastic properties.

6.5. Motivation and Industrial Applications in Agriculture and Energy Systems

The development of the proposed distribution is strongly motivated by practical challenges encountered in industrial agriculture and energy-related production systems, where highly variable and overdispersed biological data frequently arise. In particular, pest population data such as counts of European red mites (*Panonychus ulmi*) recorded on apple leaves and European corn borer larvae (*Pyrausta nubilalis*) observed in maize fields exhibit complex stochastic behavior that cannot be adequately captured by classical discrete models.

The primary motivations driving the development of the proposed model are outlined below, together with their relevance to these applied contexts:

1. The suggested distribution effectively models right-skewed data with heavy tails across different kurtosis levels. In agricultural monitoring programs, most sampling units (e.g., apple leaves or maize stalks) may contain few or no pests, while a small proportion exhibit extremely high infestation levels. Such aggregation behavior produces pronounced skewness and heavy tails, characteristics that are commonly observed in European red mite and European corn borer datasets.

2. The model provides a robust framework for modeling overdispersed count data, particularly when extreme values and outliers pose significant challenges for traditional distributions such as the Poisson model. Pest outbreaks are often driven by climatic variability, environmental heterogeneity, and biological clustering effects, leading to variance substantially exceeding the mean. This overdispersion is especially evident in field data related to *P. ulmi* and *P. nubilalis*, where sudden population surges may occur.

3. Closed-form expressions for essential statistical and reliability measures including moments, generating functions, quantiles, and related characteristics enhance both theoretical tractability and practical implementation. In agro-industrial and bioenergy systems (e.g., maize production for bioethanol), accurate estimation of pest risk is crucial for yield forecasting, economic threshold determination, and integrated pest management strategies. The availability of explicit formulas facilitates efficient parameter estimation and risk assessment without excessive computational burden.

4. The hazard rate function of the proposed distribution exhibits substantial flexibility, allowing for decreasing, increasing, unimodal, or bathtub-shaped patterns. This flexibility is particularly valuable for modeling infestation dynamics and failure mechanisms in agricultural and energy crop systems. For example, pest infestation risk may initially decrease following treatment (decreasing hazard), increase under favorable environmental conditions (increasing hazard), or display more complex temporal behavior such as a bathtub shape reflecting early vulnerability, mid-season stability, and late-season resurgence.

Overall, the proposed model is well suited for describing the stochastic structure of pest population data in agricultural production and energy-related crop systems. Its flexibility in capturing skewness, overdispersion, extreme observations, and diverse hazard rate behaviors makes it especially appropriate for analyzing European red mite infestations in apple orchards and European corn borer larvae counts in maize fields, thereby supporting improved monitoring, forecasting, and decision-making in industrial and energy applications.

7. Conclusion

This work presents a novel three-parameter discrete mixing distribution formed by integrating an extended geometric distribution with the discrete Burr-Hatke distribution. The mathematical framework of the proposed MGoDBu model was meticulously formulated using precise closed-form formulas for its principal statistical features. The distribution shown significant adaptability in modeling count data marked by overdispersion and differing kurtosis levels in right-skewed scenarios, proving especially efficient for the analysis of extreme values and outlier data points. The hazard rate function exhibited various behaviors, including increasing, decreasing, bathtub-shaped, and unimodal patterns, converging to a constant as x approached infinity. A maximum likelihood estimate framework was devised and executed for the suggested distribution. The log-likelihood function and score equations were explicitly formulated, and numerical optimization methods were comprehensively examined. The efficacy of the MLEs was assessed by comprehensive Monte Carlo simulation simulations across many parameter combinations. The simulation results indicated that the MLEs exhibited superior statistical characteristics across all analyzed schemes. Consistency was validated by the convergence of bias, MSE, and RMSE to zero with increasing sample size, whereas $\sqrt{n}$-consistency was demonstrated through the stability of the products of $n \times$ MSE. Verification of asymptotic normality was achieved by coverage probabilities that neared the nominal 95% level, thus substantiating the application of normal-based confidence intervals. The estimators demonstrated strong performance across several parameter configurations, from balanced moderate values to extreme combinations. For practical implementation, sample sizes of $n \geq 120$ yielded credible inferences with coverage probabilities

that exceeded 0.94, whereas $n \geq 300$ was advised for high-precision applications, resulting in RMSE < 0.05 and near-nominal coverage rates. The findings corroborated the MLE methodology for the MGoDBu distribution and provided practitioners with empirically supported sample size recommendations. The practical applicability of the MGoDBu distribution was evidenced by applications to three distinct real-world datasets, where it displayed a superior fit relative to various conventional competing models.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

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).

Data Availability Statement

All datasets used in this study are available within the paper.

REFERENCES

[1] Almazah, M. M. A., & Ahmad, Z. (2021). Modeling extreme values utilizing weighted mixture distributions. *Symmetry*, 13(5), 859.

[2] Bodhisuwan, W., & Sangpoom, S. (2016). The discrete weighted Lindley distribution. In *Proceedings of the International Conference on Mathematics, Statistics, and Their Applications*, Banda Aceh, Indonesia, 4–6 October.

[3] Böhning, D. (1999). *Computer-Assisted Analysis of Mixtures and Applications: Meta-Analysis, Disease Mapping and Others*. Chapman and Hall/CRC.

[4] Chakraborty, S., & Chakravarty, D. (2012). Discrete gamma distributions: Properties and parameter estimations. *Communications in Statistics-Theory and Methods*, 41(18), 3301–3324.

[5] Chen, H., Chen, J., & Kalbfleisch, J. D. (2001). A modified likelihood ratio test for homogeneity in finite mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(1), 19–29.

[6] Dean, C., & Lawless, J. F. (1989). Tests for detecting overdispersion in Poisson regression models. *Journal of the American Statistical Association*, 84(406), 467–472.

[7] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1), 1–38.

[8] Diebolt, J., & Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(2), 363–375.

[9] 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.

[10] El-Morshedy, M., Eliwa, M. S., El-Dawoody, M., & Shahen, H. S. (2025). A weighted hybrid discrete probability model: Mathematical framework, statistical analysis, estimation techniques, simulation-based ranking, and goodness-of-fit evaluation for over-dispersed data. *Electronic Research Archive*, 33(4).

[11] Everitt, B. S., & Hand, D. J. (1981). *Finite Mixture Distributions*. Chapman and Hall.

[12] Greenwood, M., & Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society*, 83(2), 255–279.

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

[14] 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.

[15] Johnson, N. L., Kemp, A. W., & Kotz, S. (2005). *Univariate Discrete Distributions* (3rd ed.). John Wiley & Sons.

[16] Karlis, D., & Xekalaki, E. (2005). Mixed Poisson distributions. *International Statistical Review*, 73(1), 35–58.

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

[18] Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1), 1–14.

[19] Leroux, B. G. (1992). Consistent estimation of a mixing distribution. *The Annals of Statistics*, 20(3), 1350–1060.

[20] Marin, J. M., Mengersen, K., & Robert, C. P. (2005). Bayesian modelling and inference on mixtures of distributions. *Handbook of Statistics*, 25, 459–507.

[21] McLachlan, G. J., & Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons.

[22] Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33(3), 341–365.

[23] 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.

[24] Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London. Series A*, 185, 71–110.

[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] Ridout, M., Deméitrio, C. G. B., & Hinde, J. (1998). Models for count data with many zeros. *Proceedings of the XIXth International Biometric Conference*, 19, 179–192.

[27] Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3), 507–554.

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

[29] Sankaran, M. (1970). The discrete Poisson-Lindley distribution. *Biometrics*, 26(1), 145–149.

[30] 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: Statistics in Society*, 109(3), 296–296.

[31] Teicher, H. (1963). Identifiability of finite mixtures. *The Annals of Mathematical Statistics*, 34(4), 1265–1269.

[32] Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons.

[33] Yau, K. K. W., Wang, K., & Lee, A. H. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal*, 45(4), 437–452.