AIMS Mathematics, 11(11): 36469–36516.
DOI: 10.3934/math.20261411
Received: 20 June 2026
Revised: 29 September 2026
Accepted: 19 October 2026
Published: 04 November 2026

https://www.aimspress.com/journal/Math

Research article

The discrete modified weighted exponential distribution: mathematical theory with applications

Mohamed S. Eliwa[1,2], **Mahmoud El-Morshedy**[3], **Afrah Al-Bossly**[3], **Mohamed El-Dawoody**[3,*] **and M. E. Sobh**[2]

[1] Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia

[2] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

[3] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

* **Correspondence:** Email: m.eldawoody@psau.edu.sa.

Abstract: This paper presents a new two-parameter discrete probability distribution called the discrete modified weighted exponential (DM-WE) distribution. A thorough examination of its mathematical and structural aspects is performed, resulting in closed-form formulas for essential statistical attributes. This encompasses the form, unimodality, and characteristics of both the hazard rate and reversed hazard rate functions. Additionally, essential foundational aspects, including parameter identifiability, asymptotic tail behavior, mode, quantile function, random data generation, and moments, are derived. The DM-WE distribution demonstrates a discontinuous growing failure rate and features a highly adaptable probability mass function that can represent various distributional forms. The examination of dispersion, zero-modification indices, and entropy metrics indicates the model is specifically tailored for over-dispersed, leptokurtic, and zero-inflated count data environments. Numerical investigations show that the coefficients of skewness and kurtosis exhibit a pronounced bathtub-shaped pattern within the explored parameter grid, indicating significant variability in tail thickness and asymmetry. Order statistics and their descriptive characteristics are also determined. The maximum likelihood estimation framework is established for parameter estimation, along with its asymptotic features. A thorough Monte Carlo simulation analysis is conducted to assess and compare the finite-sample performance and efficacy of the estimators. The practical applicability and superiority of the DM-WE distribution are proved using a comprehensive model comparison methodology and goodness-of-fit criteria applied to real-world datasets, establishing its effectiveness over existing rival discrete models.

Keywords: statistical model; ordered statistics; failure analysis; simulation; decision-making
Mathematics Subject Classification: 60F05, 62G09, 62G20, 62G30, 62G32

1. Introduction

In recent decades, data science has grown increasingly important in deriving reliable insights from complex data. A key phase in this process is the choice of a probability model that accurately represents the distributional characteristics of the observations and produces interpretable inferential metrics (e.g., tail probabilities, predictive distributions, and uncertainty measures). In reliability and survival contexts, although continuous-time models are mathematically advantageous, they frequently prove inadequate when the measuring process is inherently discrete, when the recording apparatus has insufficient resolution, or when the phenomenon is only observed at inspection intervals. Common instances include the number of rounds fired before the first malfunction, the number of cycles until the first failure of cyclic devices, the daily duration of patient stays in a hospital ward, discrete examination scores, the number of computer failures within designated intervals, the number of deaths in a specific locale during a defined timeframe, and the occurrence rate of natural phenomena such as wildfires or earthquakes. In these applications, the observed "lifetime" is a count, thereby representing a non-negative, integer-valued random variable, necessitating the use of discrete lifetime/count models. Despite their prevalence, traditional discrete distributions, particularly the Poisson and geometric, often prove insufficient for contemporary data. The Poisson distribution requires equidispersion (variance equal to the mean), yet actual data frequently display overdispersion or underdispersion due to unobserved heterogeneity, clustering, contagion effects, or reporting processes [1, 2]. The geometric distribution, as the discrete counterpart of the exponential distribution, implies a constant discrete hazard, which may be unrealistic in reliability and survival scenarios where risk generally varies with age or usage. Moreover, numerous empirical datasets exhibit an abundance of zeros that exceed the predictions of conventional models; this necessitates the use of zero-inflated and hurdle-type frameworks commonly employed in count regression [3, 4]. These challenges have driven ongoing interest in developing adaptable discrete distributions that may depict various dispersion regimes, skewness/kurtosis characteristics, heavy tails, and an array of hazard-rate forms.

1.1. Discrete lifetime distributions via survival discretization

The literature mainly concerns the discretization that leads to the development of discrete analogs of continuous lifetime distributions. Detailed studies of discretization techniques and their statistical implications can be found in standard texts on discrete distributions [5, 6]. Among these strategies is a particularly influential, dependability-based approach founded on the survival (reliability) function. Alternatively, discretization can be achieved through mixed-Poisson construction, where the discrete shape is inherited from a continuous mixing distribution on the Poisson rate, or through infinite-series method, where the probability mass function (PMF) is obtained by normalizing the continuous density directly over the integers:

$$P(Y = k) = f_X(k) / \sum_j f_X(j).$$

Table 1 summarizes how the three constructions differ along the properties most relevant to reliability modeling.

Table 1. Comparison of discretization strategies for a continuous lifetime Z.

Property	Survival	Mixed-Poisson	Infinite-series
Preserves $S_Z(x)$ at integers	Yes	No	No
Automatically normalized (no extra constant)	Yes	Yes	No
Inherits monotone (IFR/DFR) aging	Yes	No	No

It is important to note that survival discretization is not presented here as mathematically unique from an optimality viewpoint; rather, it is the natural construction when the primary modeling objective, as in the present work, is the retention of aging and reliability interpretations inherited from the continuous baseline. No discretization scheme is uniquely optimal in an unqualified sense, as each is designed to preserve a different property of the continuous baseline: Survival discretization preserves the survival function (SF) itself at integer points, the mixed-Poisson construction preserves a specific overdispersion mechanism, and the infinite-series method preserves the shape of the density but neither the SF nor, in general, a closed-form normalizing constant. Other discretization schemes may therefore be preferred for other priorities, such as moment matching or preservation of a particular dispersion structure, but those concerns are beyond the reliability-centered scope of this paper. Let Z be a non-negative continuous failure time with its SF $S_Z(z) = P(Z > z)$. If we take the integer-valued variable to be $X = \lfloor Z \rfloor$, then the PMF becomes

$$p(X = x) \quad = \quad S_Z(x) - S_Z(x+1), \qquad x = 0, 1, 2, \ldots. \tag{1.1}$$

An important advantage of Eq (1.1) is that it is interpretable, as the discrete model keeps the functional form of the continuous survival at integer values. Many reliability properties (e.g., monotonic tendencies of survival and related aging concepts) are thus preserved in an analytically tractable way.

Survival-based discretization is advantageous for modeling discrete-time failure mechanisms while preserving a direct connection to continuous-time physics. The reliability-focused discrete Weibull distribution proposed by Nakagawa and Osaki [7] provided a fundamental alternative to the geometric model by allowing for nonconstant hazard rates. In the following decades, the literature has broadened to encompass many discrete analogues aimed at representing intricate hazard structures and tail behaviors. Significant advancements include the discrete normal and Rayleigh models [8, 9], the discrete Maxwell and Burr families [10, 11], the discrete inverse Weibull [12], and the discrete Lindley distribution [13]. This progression continued with increasingly flexible models, including the discrete Burr type III [14], discrete generalized Rayleigh [15], discrete Burr–Hatke [16], discrete Gumbel [17], and exponentiated discrete Lindley [18]. Recent advances have significantly enriched this field, particularly with the Burr–Hatke class extensions [19], the discrete power quasi-Xgamma distribution [20], the discrete Hjorth distribution [21], a discrete analog of the exponentiated generalized Weibull-G family [22], a discrete analog of the generalized exponential distribution [23], and the discrete logarithmic-transformed exponential model [24].

1.2. Zero-inflated and overdispersed

At the same time, the extensive count-data literature has developed alternative mechanisms for flexible mixtures, compounding, generalized Poisson-type models, and regression-oriented

frameworks primarily driven by dispersion and zero-frequency concerns [1, 3]. Although these approaches are robust, they do not consistently yield reliability-centric metrics (survival and hazard) in a definitive, interpretable form, and they may not be designed to preserve the aging characteristics that are pivotal in lifetime analysis. Furthermore, discrete-time survival analysis has a distinct inferential heritage (e.g., discrete hazard modeling at inspection intervals), highlighting the significance of hazard-based interpretations [25]. Despite the considerable extant literature, several important limitations persist in practice. The primary issue is the limited joint treatment of reliability shape and dispersion; although several discrete analogs emphasize hazard-rate behavior, they frequently do not adequately address the overdispersion and excess zeros typical of contemporary count and lifetime datasets. We note that reliability evaluation methodology has also advanced substantially in adjacent engineering contexts, for instance, through adaptive surrogate-model-based approaches for structural reliability assessment [26], underscoring the broader and continuing relevance of flexible, computationally efficient reliability-analysis frameworks across disciplines. This inflexibility affects tail behavior, as numerous models are limited in their potential skewness and kurtosis, hindering their use for heavy-tailed or highly asymmetric data. Moreover, analytical tractability remains a significant difficulty, as several proposed distributions lack closed-form expressions for essential quantities such as moments, generating functions, and entropy measures. The field is ultimately hindered by insufficient inferential analyses. Even when likelihood inference is attainable, practitioners frequently lack thorough comparisons of diverse estimators, including maximum likelihood, least squares, and Cramér–von Mises-type approaches, under finite samples and differing parameter regimes.

These identified limitations call for discrete models that (i) maintain a reliability interpretation through survival discretization, (ii) offer adequate flexibility to accommodate overdispersion and possible zero-inflation, and (iii) remain analytically tractable for both theoretical and practical applications. In light of the above, we propose a novel two-parameter discrete model, designated as the discrete modified weighted exponential (DM-WE) distribution, derived through the survival discretization approach outlined in Eq (1.1). The DM-WE model provides closed-form expressions for various essential quantities, such as survival and hazard functions, generating functions, and moments. It is specifically designed to handle positively skewed data and overdispersed counts; depending on the parameter values, it may also improve fit in the presence of excess zeros. It should be noted from the outset that, like its geometric limiting case, the DM-WE distribution accommodates only constant or increasing hazard rates and is therefore unsuitable for phenomena governed by decreasing or bathtub-shaped failure rates, such as infant mortality or burn-in reliability data.

The remainder of this paper is organized as follows. Section 2 presents the DM-WE distribution, explaining its mathematical derivation and reliability attributes. Section 3 analyzes its statistical aspects, including shape behavior, unimodality, characteristics of the hazard-rate function (HRF), parameter identifiability, mode, quantile function, random variate generation, moments and associated statistical measures, order statistics, and entropy analysis. Section 4 addresses maximum likelihood estimation (MLE) and the corresponding asymptotic properties of the estimators. Section 5 presents a simulation study to assess the performance of the proposed methods. Section 6 describes the model comparison framework and the goodness-of-fit criteria used. Section 7 concludes the paper with a summary of the principal findings and observations.

2. The DM-WE distribution: mathematical foundations and reliability characterizations

To create a thorough mathematical framework, we initially examine a baseline continuous random variable X_0 that adheres to the continuous two-parameter modified weighted exponential (M-WE) distribution. The SF of this continuous baseline model is defined by:

$$S_0(x;\alpha,\lambda) = \frac{1}{\alpha+1}\left(\alpha+2-e^{-\alpha\lambda x}\right)e^{-\lambda x}, \quad x>0, \tag{2.1}$$

where the parameter spaces are strictly bounded such that $\alpha \geq 0$ and $\lambda > 0$. To adapt this framework for empirical scenarios governed by discrete time intervals, digital counting processes, or integers, a discrete analog is generated via the survival discretization methodology. By mapping the continuous survival profile onto the natural support domain $\mathbb{N} = \{0,1,2,\dots\}$ and applying the transformation parameter relation $\beta = e^{-\lambda}$ (where $0<\beta<1$), the SF of the DM-WE distribution is formalized as:

$$S(x;\alpha,\beta) = \frac{\beta^{(\alpha+1)(x+1)} - (\alpha+2)\beta^{x+1} - \beta^{(\alpha+1)(x+2)} + (\alpha+2)\beta^{x+\alpha+2}}{(\alpha+1)\left(\beta^{\alpha+1}-1\right)}, \quad x\in\mathbb{N}. \tag{2.2}$$

The analytical flexibility of the DM-WE distribution is demonstrated by evaluating its asymptotic boundaries across specific parameter configurations. The SF exhibits the following conditional limits:

$$S(x;\alpha,\beta) = \begin{cases} \frac{\beta(2+\alpha-\beta^{\alpha})}{\alpha+1}, & x=0,\\ \beta^{x+1}, & \alpha\to 0,\\ -\frac{1}{2}\beta^{x+1}\left(\beta^{x+1}-3\right), & \alpha\to 1,\\ 1, & \beta\to 1. \end{cases} \tag{2.3}$$

Notably, as $\alpha \to 0$, the SF converges cleanly to a standard geometric distribution survival profile, emphasizing its utility as a generalized framework. From the primary axiomatic definitions of discrete probability, the corresponding cumulative distribution function (CDF) is expressed as

$$F(x;\alpha,\beta) = 1 - \frac{\beta^{(\alpha+1)(x+1)} - (\alpha+2)\beta^{x+1} - \beta^{(\alpha+1)(x+2)} + (\alpha+2)\beta^{x+\alpha+2}}{(\alpha+1)\left(\beta^{\alpha+1}-1\right)}, \quad x\in\mathbb{N}. \tag{2.4}$$

Symmetrically, the PMF is derived by taking the first difference of the SF, $P(x;\alpha,\beta) = S(x-1;\alpha,\beta) - S(x;\alpha,\beta)$, which yields the following explicit algebraic form:

$$P(x;\alpha,\beta) = \frac{2+\alpha-\alpha\beta-2\beta-\beta^{\alpha x}+\beta^{\alpha x+\alpha+1}}{\alpha+1}\beta^{x}, \quad x\in\mathbb{N}. \tag{2.5}$$

Evaluating the structural dynamics of the PMF across boundary conditions clarifies the distributional behavior under diverse shape configurations:

$$P(x;\alpha,\beta) = \begin{cases} \frac{\beta^{\alpha+1}-\alpha\beta-2\beta+\alpha+1}{\alpha+1}, & x=0,\\ (1-\beta)\beta^{x}, & \alpha\to 0,\\ (1-\beta)\beta^{x}, & \alpha\to\infty,\\ \frac{1}{2}(\beta-1)\left(\beta^{x+1}+\beta^{x}-3\right)\beta^{x}, & \alpha\to 1,\\ 0, & \beta\to 1. \end{cases} \tag{2.6}$$

To provide a comprehensive reliability assessment, we define the operational functions describing systemic aging and failure rates. The HRF, which quantifies the conditional probability of immediate failure given survival up to step x, is expressed as

$$h(x;\alpha,\beta) = \frac{P(x;\alpha,\beta)}{S(x-1;\alpha,\beta)} = \frac{\left(\beta^{\alpha+1}-1\right)\left(\left(\beta^{\alpha+1}-1\right)\beta^{\alpha x}-(\alpha+2)(\beta-1)\right)}{\beta^{(\alpha+1)(x+1)}-(\alpha+2)\beta^{x+1}-\beta^{(\alpha+1)(x+2)}+(\alpha+2)\beta^{x+\alpha+2}}\beta^{x}, \quad x \in \mathbb{N}. \tag{2.7}$$

In parallel lifetime settings, such as left-censored data analysis, the reversed HRF (RHRF) provides a vital complementary measure, formalized as:

$$r(x;\alpha,\beta) = \frac{P(x;\alpha,\beta)}{F(x;\alpha,\beta)} = \frac{\left(\beta^{\alpha+1}-1\right)\left((\alpha+2)(\beta-1)-\left(\beta^{\alpha+1}-1\right)\beta^{\alpha x}\right)}{\psi(x;\alpha,\beta)}\beta^{x}, \quad x \in \mathbb{N}, \tag{2.8}$$

where the comprehensive denominator expansion kernel $\psi(x;\alpha,\beta)$ is given by:

$$\psi(x;\alpha,\beta) = \beta^{(\alpha+1)(x+1)}-\beta^{(\alpha+1)(x+2)}-\beta^{\alpha+1}-2\beta^{x+1}+2\beta^{x+\alpha+2}+\alpha\left(\beta^{\alpha+1}-1\right)\left(\beta^{x+1}-1\right)+1.$$

To visually assess the theoretical flexibility of these metrics, Figures 1 and 2 display the diverse shapes of the PMF, HRF, and RHRF curves across different configurations of the parameters α and β. These figures highlight how well the DM-WE distribution fits asymmetric unimodal data subject to increasing risk.

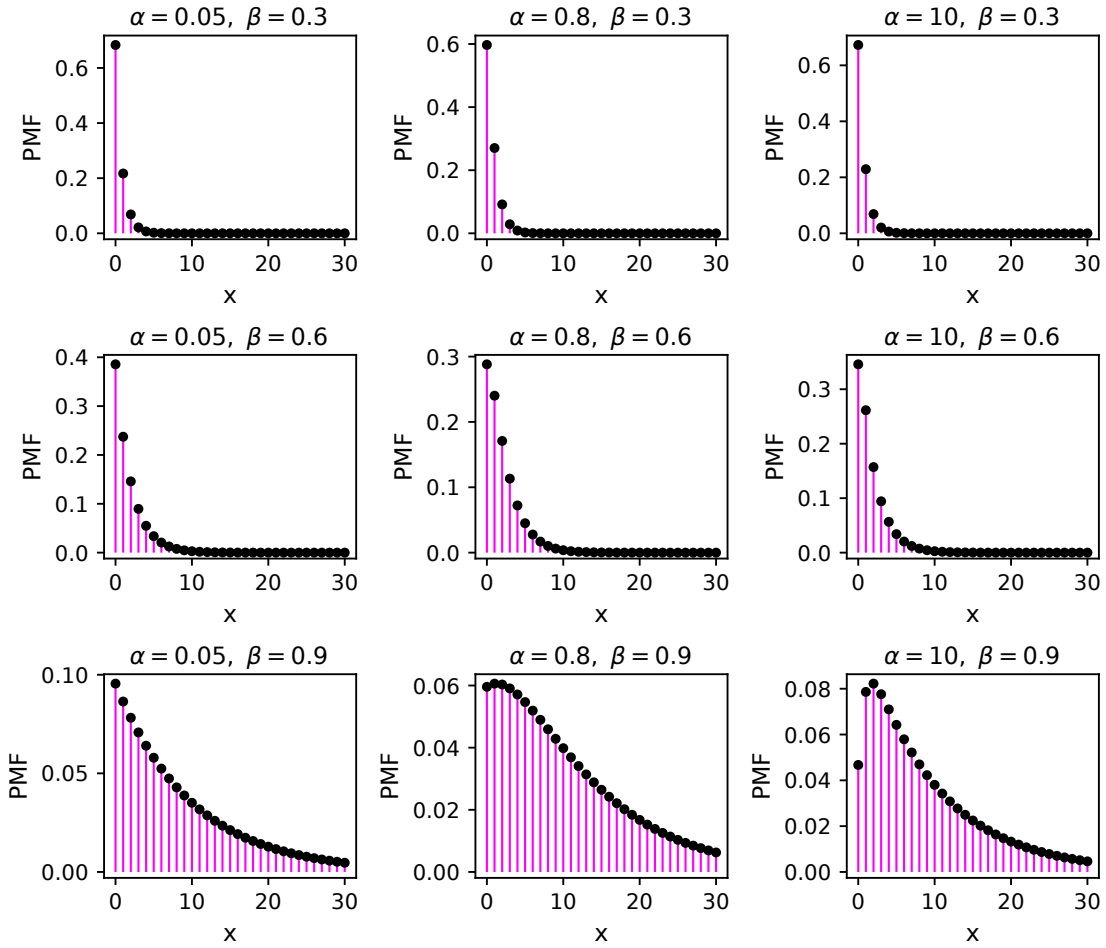


Figure 1. The PMF plots of the DM-WE distribution for various values of α and β.

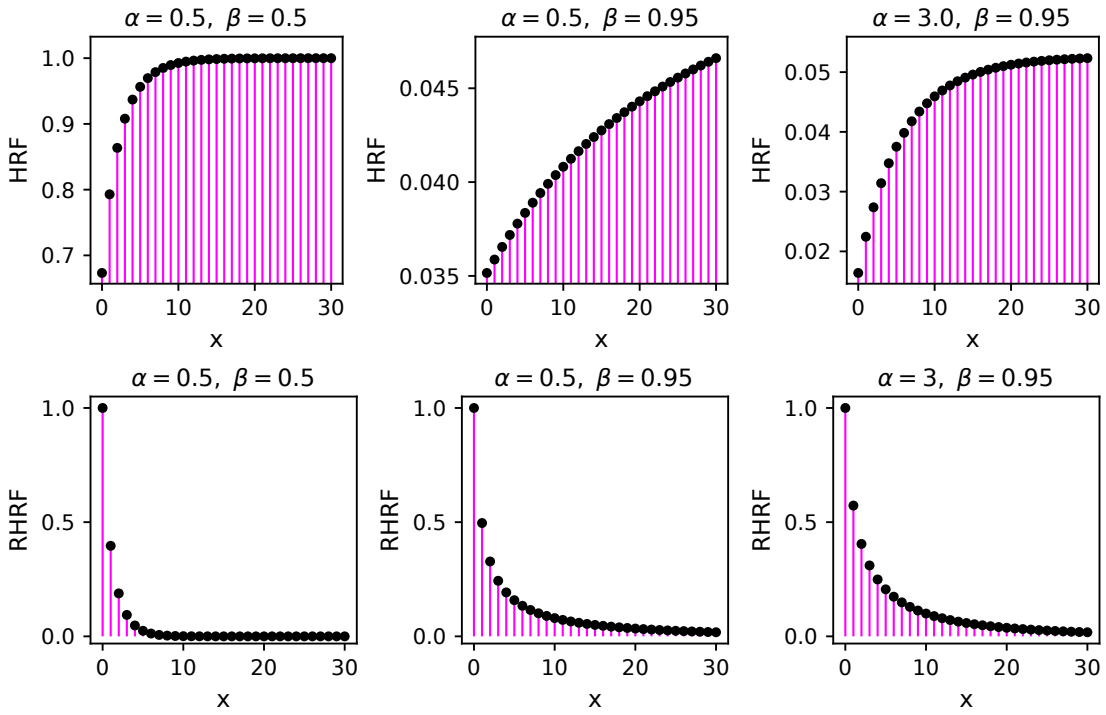


Figure 2. The HRF and RHRF plots of the DM-WE distribution for various values of α and β.

Furthermore, to look beyond standard hazard intervals, we map the structural aging dynamics via the second rate of failure $r^*(x;\alpha,\beta)$, which highlights the log-proportional changes within the survival profile:

$$\begin{aligned} r^*(x;\alpha,\beta) &= \log\left[\frac{S(x;\alpha,\beta)}{S(x+1;\alpha,\beta)}\right] \\ &= \log\left[\frac{(\alpha+2)\beta^{x+1}-\beta^{(\alpha+1)(x+1)}+\beta^{(\alpha+1)(x+2)}-(\alpha+2)\beta^{x+\alpha+2}}{(\beta^{\alpha+1}-1)(\beta^{\alpha(x+2)}-\alpha-2)}\beta^{-x-2}\right]. \end{aligned} \tag{2.9}$$

Complementing these aging characterizations, the Mills ratio (MI), which frequently arises in asymptotic calculations and lifetime risk modeling, is defined as the ratio of survival probability to immediate mass allocation:

$$\begin{aligned} MI(x;\alpha,\beta) &= \frac{S(x;\alpha,\beta)}{P(x;\alpha,\beta)} \\ &= \frac{\beta^{(\alpha+1)(x+1)}+(\alpha+2)\beta^{x+\alpha+2}-\beta^{(\alpha+1)(x+2)}-(\alpha+2)\beta^{x+1}}{(\beta^{\alpha+1}-1)(\beta^{\alpha x+\alpha+1}-\beta^{\alpha x}-\alpha\beta-2\beta+\alpha+2)}\beta^{-x}. \end{aligned} \tag{2.10}$$

3. Statistical properties

3.1. Shape, unimodality, and HRF characteristics

This section examines the fundamental structural frameworks of the PMF and the HRF that regulate the DM-WE distribution. Theorems 3.1 to 3.3 meticulously delineate the distribution's shape characteristics, log-concavity, hazard aging patterns, and unimodal attributes.

Theorem 3.1. *Let X be a discrete, random variable following the discrete DM-WE distribution characterized by parameters $\alpha > 0$ and $0 < \beta < 1$, with the corresponding PMF denoted as $P(x;\alpha,\beta)$. Then, the sequence $\{P(x;\alpha,\beta)\}_{x\in\mathbb{N}}$ is strictly log-concave, satisfying the inequality*

$$P^2(x;\alpha,\beta) \geq P(x-1;\alpha,\beta)\,P(x+1;\alpha,\beta), \qquad \forall\, x \in \mathbb{N}. \tag{3.1}$$

Proof. By directly referencing the explicit algebraic architecture of Eq (2.5), the structural difference between the squared localized mass and its adjacent bounds resolves into:

$$P^2(x;\alpha,\beta) - P(x-1;\alpha,\beta)P(x+1;\alpha,\beta) = \frac{(\alpha+2)(1-\beta)\left(1-\beta^{\alpha+1}\right)(\beta^{\alpha}-1)^2\,\beta^{(\alpha+2)x-\alpha}}{(\alpha+1)^2}. \tag{3.2}$$

Given the strict parameter space restrictions where $\alpha > 0$ and $0 < \beta < 1$, the terms $(1-\beta) > 0$ and $(1-\beta^{\alpha+1}) > 0$ hold true, forcing the numerator expression to remain strictly positive across all elements. Consequently, we establish that:

$$P^2(x;\alpha,\beta) \geq P(x-1;\alpha,\beta)P(x+1;\alpha,\beta), \qquad x \in \mathbb{N}, \tag{3.3}$$

which definitively proves the condition of log-concavity. □

Theorem 3.2. *Let X follow the discrete DM-WE distribution with parameters $\alpha > 0$ and $0 < \beta < 1$. Then, the discrete DM-WE distribution profile is strictly unimodal.*

Proof. By leveraging foundational discrete probability axioms established by Keilson and Gerber [27], it is known that any PMF exhibiting strict log-concavity is inherently strongly unimodal. Because the discrete DM-WE distribution's PMF was proven to be log-concave in Theorem 3.1, its properties of unimodality follow directly and immediately as a structural consequence ([27, Theorem 3]). □

Theorem 3.3. *Let X follow the discrete DM-WE distribution with parameter spaces constrained to $\alpha > 0$ and $0 < \beta < 1$. Then, the corresponding HRF is strictly increasing in x.*

Proof. To establish the monotonic trajectory of the HRF profile, we employ Glaser's analytical approach [28]. We define the auxiliary continuous-analog function $\eta(x)$ as the negative log-derivative ratio of the probability mass:

$$\eta(x) = -\frac{P'(x;\alpha,\beta)}{P(x;\alpha,\beta)} = \frac{\left((\alpha+2)(\beta-1) - (\alpha+1)\left(\beta^{\alpha+1}-1\right)\beta^{\alpha x}\right)\log(\beta)}{(\beta^{\alpha+1}-1)\beta^{\alpha x} - (\alpha+2)(\beta-1)}. \tag{3.4}$$

Differentiating $\eta(x)$ with respect to the spatial domain variable x yields:

$$\eta'(x) = \frac{\alpha^2(\alpha+2)(1-\beta)\left(1-\beta^{\alpha+1}\right)(\log(\beta))^2}{\left((\alpha+2)(\beta-1) - (\beta^{\alpha+1}-1)\beta^{\alpha x}\right)^2}\beta^{\alpha x}. \tag{3.5}$$

Under the valid domain conditions $\alpha > 0$ and $0 < \beta < 1$, all isolated multiplicative components in the numerator of (3.5) evaluate to strictly positive values, yielding $\eta'(x) > 0$. By Glaser's theorem, this criterion confirms that the discrete HRF exhibits an increasing hazard rate (IHR) trend relative to x. □

Remark 3.1. *The IHR property established in Theorem 3.2 holds for every finite $\alpha > 0$ and characterizes the DM-WE distribution as a model for deteriorating/wear-out aging mechanisms. The HRF shape is not unrestricted, however: As $\alpha \to 0$ and as $\alpha \to \infty$, the model reduces exactly to the geometric distribution with constant HRF, so the family ranges from constant hazard at both boundaries to strictly increasing hazard for every finite $\alpha > 0$ in between, but it is never decreasing or bathtub-shaped. Consequently, the DM-WE distribution is not suited to phenomena exhibiting decreasing or bathtub-shaped HRFs, such as infant mortality or burn-in reliability data. Extensions capable of accommodating nonmonotonic aging, such as via mixture, compounding, or additional shape parameters, are left for future research.*

3.2. Parameter identifiability

Prior to executing statistical inference, parameter estimation, or real data modeling, it is crucial to ascertain the structural identifiability of the model parameters. Parameter identifiability guarantees a one-to-one correspondence between the parameter space and the space of probability distributions. A distribution is identifiable if different parameter vectors produce different probability mass functions. Let $\boldsymbol{\theta} = (\alpha, \beta)^T$ and $\boldsymbol{\theta}^* = (\alpha^*, \beta^*)^T$ be two parameter vectors residing within the valid parameter space $\Theta = \{(\alpha, \beta) : \alpha \geq 0, 0 < \beta < 1\}$. The DM-WE distribution is identifiable if and only if:

$$P(x; \boldsymbol{\theta}) = P(x; \boldsymbol{\theta}^*) \quad \forall\, x \in \mathbb{N} \implies \boldsymbol{\theta} = \boldsymbol{\theta}^*. \tag{3.6}$$

Theorem 3.4. *The parameters α and β of the DM-WE distribution, characterized by the PMF in* (2.5)*, are strictly identifiable.*

Proof. Assume that for two parameter vectors $\boldsymbol{\theta} = (\alpha, \beta)^T$ and $\boldsymbol{\theta}^* = (\alpha^*, \beta^*)^T$, the identity $P(x; \boldsymbol{\theta}) = P(x; \boldsymbol{\theta}^*)$ holds across the entire support $x \in \mathbb{N}$. That is:

$$\frac{2 + \alpha - \alpha\beta - 2\beta - \beta^{\alpha x} + \beta^{\alpha x + \alpha + 1}}{\alpha + 1}\beta^x = \frac{2 + \alpha^* - \alpha^*\beta^* - 2\beta^* - \beta^{*\alpha^* x} + \beta^{*\alpha^* x + \alpha^* + 1}}{\alpha^* + 1}\beta^{*x}. \tag{3.7}$$

To prove that $\boldsymbol{\theta} = \boldsymbol{\theta}^*$, we systematically evaluate the behavior of the PMF at the boundary states of the support $\mathbb{N}$.

Case 1: Evaluation at $x = 0$. Substituting $x = 0$ into both sides of Eq (3.7) isolates the localized mass at the origin:

$$\frac{2 + \alpha - \alpha\beta - 2\beta - 1 + \beta^{\alpha+1}}{\alpha + 1} = \frac{2 + \alpha^* - \alpha^*\beta^* - 2\beta^* - 1 + \beta^{*\alpha^*+1}}{\alpha^* + 1}. \tag{3.8}$$

Simplifying the constant integer terms in the numerators yields

$$\frac{1 + \alpha - \alpha\beta - 2\beta + \beta^{\alpha+1}}{\alpha + 1} = \frac{1 + \alpha^* - \alpha^*\beta^* - 2\beta^* + \beta^{*\alpha^*+1}}{\alpha^* + 1}. \tag{3.9}$$

Case 2: Asymptotic tail behavior ($x \to \infty$). Consider the ratio of the PMF at two consecutive long-tail survival steps, $x + 1$ and x, as x approaches infinity. Because $0 < \beta < 1$ and $\alpha \geq 0$, the exponential terms involving $\beta^{\alpha x}$ decay asymptotically to zero; i.e.,

$$\lim_{x \to \infty} \beta^{\alpha x} = 0.$$

Taking the limit of the probability ratio for both parametric representations yields

$$\lim_{x\to\infty} \frac{P(x+1;\boldsymbol{\theta})}{P(x;\boldsymbol{\theta})} = \lim_{x\to\infty} \frac{P(x+1;\boldsymbol{\theta}^*)}{P(x;\boldsymbol{\theta}^*)}. \tag{3.10}$$

Substituting the PMF structures into the limit functional, the complex parameter brackets cancel out, isolating the geometric decay factor:

$$\lim_{x\to\infty} \left(\frac{\text{Bracket}(\alpha,\beta)\cdot\beta^{x+1}}{\text{Bracket}(\alpha,\beta)\cdot\beta^{x}} \right) = \lim_{x\to\infty} \left(\frac{\text{Bracket}(\alpha^*,\beta^*)\cdot\beta^{*x+1}}{\text{Bracket}(\alpha^*,\beta^*)\cdot\beta^{*x}} \right). \tag{3.11}$$

This directly simplifies to:

$$\beta = \beta^*. \tag{3.12}$$

Thus, the scale parameter β is uniquely determined by the asymptotic geometric decay rate of the distribution's upper tail.

Case 3: Uniqueness of α. Having established that $\beta = \beta^*$, we substitute this identity back into the initial origin state equation derived in (3.9). This reduces the system to a single-variable equation dependent strictly on α and α^*:

$$\frac{1+\alpha-\alpha\beta-2\beta+\beta^{\alpha+1}}{\alpha+1} = \frac{1+\alpha^*-\alpha^*\beta-2\beta+\beta^{\alpha^*+1}}{\alpha^*+1}. \tag{3.13}$$

Let us define a continuous auxiliary function $g(y)$ over the domain $y \geq 0$ for a fixed valid value of β:

$$g(y) = \frac{1+y-y\beta-2\beta+\beta^{y+1}}{y+1} = 1-\beta+\frac{\beta^{y+1}-2\beta}{y+1}. \tag{3.14}$$

Differentiating $g(y)$ with respect to y to verify its monotonicity yields:

$$g'(y) = \frac{\beta^{y+1}\ln(\beta)(y+1)-(\beta^{y+1}-2\beta)}{(y+1)^2}. \tag{3.15}$$

Since, $0 < \beta < 1$, $\ln(\beta) < 0$, and $\beta^{y+1} < 2\beta$ for all $y \geq 0$, the derivative $g'(y)$ remains strictly signed (injective) across its defined parameter space. Because $g(y)$ is a strictly monotonic, injective function, the equality $g(\alpha) = g(\alpha^*)$ fundamentally demands that:

$$\alpha = \alpha*. \tag{3.16}$$

Combining the results from Eqs (3.12) and (3.16), we obtain $(\alpha,\beta)^T = (\alpha^*,\beta^*)^T$, which completes the proof. The DM-WE distribution parameters are strictly identifiable. □

3.3. Mode

Theorem 3.5. *Let X be a discrete random variable following the discrete, DM-WE distribution. The mode of X, denoted by x_0, satisfies*

$$x_0 \in \left[\frac{1}{\alpha\log(\beta)}\log\left(\frac{(\alpha+2)(\beta-1)^2}{(\beta^{\alpha+1}-1)^2}\right), \frac{1}{\alpha\log(\beta)}\log\left(\frac{(\alpha+2)(\beta-1)^2\beta^{\alpha}}{(\beta^{\alpha+1}-1)^2}\right)\right].$$

Furthermore, $x_0 = 0$ if the upper bound is negative.

Proof. Because the discrete DM-WE distribution is unimodal, its PMF, $P(x;\alpha,\beta)$, satisfies the following conditions for the mode x_0:

$$P(x;\alpha,\beta) \geq P(x-1;\alpha,\beta), \quad \text{for all } x \leq x_0 \tag{3.17}$$

and

$$P(x;\alpha,\beta) \geq P(x+1;\alpha,\beta), \quad \text{for all } x \geq x_0. \tag{3.18}$$

Evaluating the inequality (3.17) leads to the upper bound for x:

$$x \leq \frac{1}{\alpha \log(\beta)} \log\left(\frac{(\alpha+2)(\beta-1)^2\beta^{\alpha}}{(\beta^{\alpha+1}-1)^2}\right). \tag{3.19}$$

Similarly, evaluating the inequality (3.18) leads to the lower bound:

$$x \geq \frac{1}{\alpha \log(\beta)} \log\left(\frac{(\alpha+2)(\beta-1)^2}{(\beta^{\alpha+1}-1)^2}\right). \tag{3.20}$$

Combining the inequalities (3.19) and (3.20) yields the required interval for x_0. In the event that the expression in (3.19) is negative, the PMF is nonincreasing for all $x \in \mathbb{N} \cup \{0\}$, which implies $x_0 = 0$. □

3.4. Quantile function and random data generation

The quantile function, sometimes referred to as the inverse CDF or percent-point function, is an essential instrument in statistical modeling. It fulfills an essential dual function: First, it delineates the theoretical structural limits of the distribution (including the median and quartiles); second, it offers a precise, nonparametric framework for simulating pseudo-random variables through the inversion approach. Let X be a discrete random variable that adheres to the DM-WE distribution. The DM-WE distribution is rigorously discrete, supported on the natural numbers $\mathbb{N}$, resulting in discontinuous cumulative steps. Therefore, we must delineate the quantile function utilizing the generalized inverse framework:

$$Q(u) = \inf\{x \in \mathbb{N} : F(x;\alpha,\beta) \geq u\}, \tag{3.21}$$

where u is a continuous, uniform, random variable drawn from the standard open interval $u \in (0,1)$. To construct an operational form for Eq (3.21), we isolate the algebraic terms depending on x within the inequality $F(x;\alpha,\beta) \geq u$. Substituting the explicit framework of Eq (2.4) into the boundary condition yields:

$$1 - \frac{\beta^{(\alpha+1)(x+1)} - (\alpha+2)\beta^{x+1} - \beta^{(\alpha+1)(x+2)} + (\alpha+2)\beta^{x+\alpha+2}}{(\alpha+1)\left(\beta^{\alpha+1}-1\right)} \geq u. \tag{3.22}$$

Rearranging the terms to isolate the fraction gives:

$$\frac{\beta^{(\alpha+1)(x+1)} - (\alpha+2)\beta^{x+1} - \beta^{(\alpha+1)(x+2)} + (\alpha+2)\beta^{x+\alpha+2}}{(\alpha+1)(\beta^{\alpha+1}-1)} \leq 1-u. \tag{3.23}$$

By factoring out the common base term β^{x+1} from the numerator expressions, we can cleanly separate the structural role of x from the distribution parameters:

$$\beta^{x+1}\left[\frac{\beta^{\alpha(x+1)} - (\alpha+2) - \beta^{\alpha(x+2)+1} + (\alpha+2)\beta^{\alpha+1}}{(\alpha+1)(\beta^{\alpha+1}-1)}\right] \leq 1-u. \tag{3.24}$$

Applying the natural logarithm (ln) to both sides of the inequality moves x out of the exponent on the leading factor, though a residual copy of x remains trapped inside the inner logarithmic term. Because the base parameter fulfills the condition $0 < \beta < 1$, the logarithmic transformation $\ln(\beta)$ yields a negative coefficient, so unequivocally reversing the direction of the inequality operator, we obtain

$$x + 1 \geq \frac{\ln(1-u) - \ln\left(\frac{\beta^{\alpha(x+1)} - (\alpha+2) - \beta^{\alpha(x+2)+1} + (\alpha+2)\beta^{\alpha+1}}{(\alpha+1)(\beta^{\alpha+1}-1)}\right)}{\ln(\beta)}. \tag{3.25}$$

Applying the ceiling function $\lceil \cdot \rceil$ to enforce the discrete integer constraint over the domain $\mathbb{N}$ yields the following operational, self-referential form of the quantile function:

$$Q(u) = \left\lceil \frac{\ln(1-u) - \ln\left(\frac{\beta^{\alpha(x+1)} - (\alpha+2) - \beta^{\alpha(x+2)+1} + (\alpha+2)\beta^{\alpha+1}}{(\alpha+1)(\beta^{\alpha+1}-1)}\right)}{\ln(\beta)} \right\rceil - 1. \tag{3.26}$$

It should be emphasized that Eq (3.26) is an implicit transcendental equation: The unknown quantile x appears both outside and inside the logarithmic term on the right-hand side, so no explicit closed-form solution $Q(u)$ exists in terms of elementary functions of u, α, and β. Consequently, the quantile function, and in particular the median (Eq (3.28)) and the quartiles $Q_1 = Q(0.25)$ and $Q_3 = Q(0.75)$, must be evaluated numerically for each fixed (α, β) by applying a standard root-finding procedure, such as the bisection method or the Newton–Raphson algorithm, to the defining equation $F(x; \alpha, \beta) \geq u$, followed by application of the ceiling operator to enforce the integer-valued support. In practice, such root-finding converges rapidly and accurately, as $F(x; \alpha, \beta)$ is monotonically nondecreasing and bounded on $[0, 1]$. The median, M, is a crucial measure of central tendency that divides the distribution's probability mass into two equal halves. It has significant resilience to the leverage effects of long tails and extreme outliers. In this context, the median is determined by assessing the quantile function at precisely $u = 0.5$:

$$M = Q(0.5) = \inf\{x \in \mathbb{N} : F(x; \alpha, \beta) \geq 0.5\}. \tag{3.27}$$

By utilizing the implicit relation in Eq (3.26), the theoretical median M is characterized as the solution of the following transcendental equation, again requiring numerical root-finding rather than an explicit closed form:

$$M = \left\lceil \frac{\ln(0.5) - \ln\left(\frac{\beta^{\alpha(M+1)} - (\alpha+2) - \beta^{\alpha(M+2)+1} + (\alpha+2)\beta^{\alpha+1}}{(\alpha+1)(\beta^{\alpha+1}-1)}\right)}{\ln(\beta)} \right\rceil - 1. \tag{3.28}$$

Other essential descriptive partition parameters, including the first quartile (Q_1) and the third quartile (Q_3), are derived symmetrically, and likewise numerically, by substituting $u = 0.25$ and $u = 0.75$ into Eq (3.26) and solving the resulting implicit relation by root-finding. To run Monte Carlo simulations, carry out parameter estimations, or perform empirical validation tests on the DM-WE model, it is essential to generate dependable pseudo-random samples. Utilizing the probability-integral-transform theorem, continuous uniform random variables can be directly mapped into the discrete sample space of the target distribution. The procedure for generating a random sample of size N from the DM-WE distribution is detailed below:

(1) Initialization: Define the fixed parameter values for α ($\alpha \geq 0$) and β ($0 < \beta < 1$), and specify the desired sample size N.

(2) Stochastic sampling: Generate a pseudo-random sequence of N independent and identically distributed (i.i.d.) continuous observations from a standard uniform distribution:

$$U_i \sim \text{Uniform}(0, 1), \quad \text{for } i = 1, 2, \ldots, N.$$

(3) Inversion mapping: For each uniform realization U_i, find the smallest non-negative integer $X_i \in \mathbb{N}$ that satisfies the cumulative boundary condition. This can be handled efficiently via a sequential lookup loop or by computing the roots of the expression:

$$X_i = \min \{x \in \mathbb{N} : F(x; \alpha, \beta) \geq U_i\}.$$

(4) Output generation: Return the synthesized integer vector $\mathbf{X} = [X_1, X_2, \ldots, X_N]^T$ as the final validated random sample representing the DM-WE distribution.

3.5. Moments and associated statistical measures

This subsection examines the analytical and structural characteristics of the DM-WE distribution through raw moments and specific structural indices. The probability-generating function, defined as

$$G_X(z) = E[z^X],$$

can be articulated in a closed form as follows:

$$G_X(z) = \sum_{x=0}^{\infty} z^x P(x; \alpha, \beta) = \frac{\alpha - \alpha\beta + (z\alpha\beta - z\alpha + z\beta - 2z + 1)\beta^{\alpha+1} + \beta(z-2) + 1}{(\alpha+1)(\beta z - 1)\left(z\beta^{\alpha+1} - 1\right)}. \tag{3.29}$$

By substituting $z = e^t$ into Eq (3.29), the corresponding moment-generating function, $M_X(z) = E[e^{zX}]$, maps directly to:

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} P(x; \alpha, \beta) = \frac{\alpha - \alpha\beta + (\alpha\beta e^t - \alpha e^t + \beta e^t - 2e^t + 1)\beta^{\alpha+1} + \beta\left(e^t - 2\right) + 1}{(\alpha+1)\left(\beta e^t - 1\right)\left(e^t\beta^{\alpha+1} - 1\right)}. \tag{3.30}$$

The first four noncentral, raw moments (μ'_r) of the DM-WE distribution are explicitly calculated by differentiating the moment generating function at the origin ($z = 0$). Thus,

$$\mu'_1 = \frac{(2 + \alpha - (\alpha\beta + \beta + 1)\beta^{\alpha})\beta}{(\alpha+1)(\beta-1)\left(\beta^{\alpha+1} - 1\right)}, \tag{3.31}$$

$$\mu'_2 = \frac{-\frac{3}{\beta^{\alpha+1}-1} - \frac{2}{\left(\beta^{\alpha+1}-1\right)^2} + \frac{\beta(\alpha\beta+\alpha+\beta+4)-1}{(\beta-1)^2}}{\alpha+1}, \tag{3.32}$$

$$\mu'_3 = \frac{\frac{7}{\beta^{\alpha+1}-1} + \frac{12}{\left(\beta^{\alpha+1}-1\right)^2} + \frac{6}{\left(\beta^{\alpha+1}-1\right)^3} - \frac{\beta(\alpha\beta(\beta+4)+\alpha+\beta(\beta+11)-1)+1}{(\beta-1)^3}}{\alpha+1}, \tag{3.33}$$

$$\mu'_4 = \frac{-\frac{15}{\beta^{\alpha+1}-1} - \frac{50}{\left(\beta^{\alpha+1}-1\right)^2} - \frac{60}{\left(\beta^{\alpha+1}-1\right)^3} - \frac{24}{\left(\beta^{\alpha+1}-1\right)^4} + \frac{\beta(\alpha(\beta+1)(\beta(\beta+10)+1)+\beta(\beta(\beta+26)+16)+6)-1}{(\beta-1)^4}}{\alpha+1}. \tag{3.34}$$

Utilizing these raw moments, the theoretical variance ($\sigma^2 = \mu_2' - (\mu_1')^2$) of the DM-WE distribution resolves to:

$$\sigma^2 = \frac{\beta\Big((\alpha+2)(\alpha\beta+\beta-1)\beta^{2\alpha+1} + (\alpha+2)(\alpha-\beta+1) - \Big(2\alpha^2\beta + \alpha(\beta+1)^2 + (\beta-1)^2\Big)\beta^{\alpha}\Big)}{(\alpha+1)^2(\beta-1)^2\left(\beta^{\alpha+1}-1\right)^2}. \quad (3.35)$$

By employing the specific mathematical frameworks delineated in Eq (3.31) to (3.34), conventional shape metrics, including the coefficient of skewness (CS) and the coefficient of kurtosis (CK), can be computed. Furthermore, the model's over- or under-dispersion characteristics can be evaluated using the dispersion index (DI), which is defined as the variance-to-mean ratio:

$$DI = \frac{\sigma^2}{\mu_1'} = \frac{\frac{2\alpha+3}{1-\beta^{\alpha+1}} + \frac{2(\alpha+2)(\alpha+1)^2\beta}{(\alpha\beta+\beta+1)((\alpha\beta+\beta+1)\beta^{\alpha}-\alpha-2)} - \frac{2(\alpha+1)}{\alpha\beta+\beta+1} - \frac{\alpha}{\beta-1}}{\alpha+1}. \quad (3.36)$$

A DI number below 1 indicates acceptability for under dispersed systems, whereas a DI value above 1 signifies appropriateness for over dispersed processes. The zero-modification index (ZI) is employed to address certain inflation or deflation anomalies in empirical count observations. This index measures the surplus frequency of zero observations relative to conventional classical distributions:

$$\begin{aligned} ZI &= 1 + \frac{\log[P(X=0;\alpha,\beta)]}{\mu_1'} \\ &= \frac{(\alpha+1)(\beta-1)\left(\beta^{\alpha+1}-1\right)\log\left(\frac{\beta(\beta^{\alpha}-2)-\alpha\beta+\alpha+1}{\alpha+1}\right)}{\beta\left(2+\alpha-(\alpha\beta+\beta+1)\beta^{\alpha}\right)} + 1. \end{aligned} \quad (3.37)$$

The ZI sign dictates the zero-modification structure: $ZI > 0$ signifies zero inflation, $ZI < 0$ denotes zero deflation, and $ZI = 0$ indicates no modification. Having established the general closed-form expressions for the DM-WE distribution, it is instructive to verify their consistency in the $\alpha \to 0$ limit, where the model reduces to the classical geometric distribution (Eq (2.3)). This serves two purposes: It confirms that the closed-form expressions derived in Sections 2–3 are internally consistent with this known special case, and it clarifies precisely which of the qualitative features of the DM-WE distribution overdispersion, zero-inflation, and so on are already present at this baseline, as opposed to those introduced by the shape parameter α.

Corollary 3.1. *(Geometric limit) At $\alpha = 0$ (exactly, by direct substitution) and as $\alpha \to \infty$ (asymptotically), the DM-WE distribution reduces to the geometric distribution with parameter β, denoted*

$$G(\beta) = Geometric\,(p = 1-\beta).$$

This reduction is exact for every quantity derived in Sections 2 and 3 at $\alpha = 0$ and is approached in the limit as $\alpha \to \infty$; the two cases are distinguished formally in Proposition 3.1 and Remark 3.2 below. In Table 2, these closed-form formulas for the degenerate case are presented.

Table 2. Exact closed-form expressions for the DM-WE distribution at the boundary $\alpha = 0$ (geometric distribution limit).

Quantity	Expression	Notes
PMF	$P(X = x; 0, \beta) = (1 - \beta)\beta^x$	$x = 0, 1, 2, \ldots$
CDF	$F(X = x; 0, \beta) = 1 - \beta^{x+1}$	$x = 0, 1, 2, \ldots$
SF	$S(X = x; 0, \beta) = \beta^{x+1}$	$x = 0, 1, 2, \ldots$
HRF	$h(X = x; 0, \beta) = 1 - \beta$	Constant hazard
RHRF	$r(X = x; 0, \beta) = \dfrac{(1 - \beta)\beta^x}{1 - \beta^{x+1}}$	Decreasing in x
Probability-generating function	$G_X(z; 0, \beta) = \dfrac{1 - \beta}{1 - \beta z}$	$\lvert z \rvert < 1/\beta$
Mills ratio	$MI(X = x; 0, \beta) = \dfrac{\beta}{1 - \beta}$	Surviving-to-mass ratio
Mean	$E[X] = \dfrac{\beta}{1 - \beta}$	First moment
Variance	$\mathrm{Var}(X) = \dfrac{\beta}{(1 - \beta)^2}$	
Coefficient of skewness	$CS = \dfrac{1 + \beta}{\sqrt{\beta}}$	
Coefficient of kurtosis	$CK = 9 + \dfrac{(1 - \beta)^2}{\beta}$	Kurtosis excess
Dispersion index	$DI = \dfrac{1}{1 - \beta}$	Variance-to-mean ratio
Zero-inflation index	$ZI = 1 + \dfrac{(1 - \beta)\log(1 - \beta)}{\beta}$	

Proposition 3.1 below derives each entry of this table directly from Eqs (2.2)–(3.35) by algebraic substitution at $\alpha = 0$, and Remark 3.2 clarifies the resulting nested relationship between the two models.

Proposition 3.1. *Let $\alpha = 0$ be substituted directly, not as a limit, into the closed-form expressions derived in Sections 2 and 3. Then, the PMF, SF, HRF, RHRF, probability–generating function, mean, and variance of the DM-WE distribution reduce exactly, term by term, to the corresponding closed-form expressions of the geometric distribution $G(\beta) = \mathrm{Geometric}(p = 1 - \beta)$ reported in Table 2.*

Proof. **PMF:** Substituting $\alpha = 0$ into Eq (2.5) gives

$$P(x; 0, \beta) = \frac{2 - 2\beta - \beta^0 + \beta^1}{1}\beta^x = \frac{2 - 2\beta - 1 + \beta}{1}\beta^x = (1 - \beta)\beta^x,$$

matching Table 2, with no indeterminate form arising.

SF: Substituting $\alpha = 0$ into the numerator of Eq (2.2) gives

$$\beta^{x+1} - 2\beta^{x+1} - \beta^{x+2} + 2\beta^{x+2} = \beta^{x+2} - \beta^{x+1} = \beta^{x+1}(\beta - 1),$$

and the denominator becomes $(0 + 1)(\beta^1 - 1) = \beta - 1$. Both carry a common, nonzero factor $(\beta - 1)$ (recall $0 < \beta < 1$, so $\beta \neq 1$), which cancels algebraically:

$$S(x; 0, \beta) = \frac{\beta^{x+1}(\beta - 1)}{\beta - 1} = \beta^{x+1}.$$

The CDF follows immediately as

$$F(x;0,\beta) = 1-\beta^{x+1},$$

both matching Table 2.

HRF: Using the reductions above,

$$h(x;0,\beta) = \frac{P(x;0,\beta)}{S(x-1;0,\beta)} = \frac{(1-\beta)\beta^x}{\beta^x} = 1-\beta,$$

a constant hazard, exactly reproducing the memoryless property of the geometric distribution and consistent with the constant-HRF boundary already noted in Remark 3.1.

RHRF: Directly,

$$r(x;0,\beta) = \frac{P(x;0,\beta)}{F(x;0,\beta)} = \frac{(1-\beta)\beta^x}{1-\beta^{x+1}},$$

strictly decreasing in x, matching Table 2.

Probability generating function. Substituting $\alpha = 0$ into Eq (3.29), the numerator becomes

$$\beta(z\beta - 2z + 1) + \beta(z-2) + 1 = z\beta^2 - z\beta - \beta + 1 = (\beta-1)(z\beta-1),$$

and the denominator becomes $(1)(\beta z - 1)(z\beta - 1) = (\beta z - 1)^2$. The common factor $(z\beta - 1)$ cancels, giving

$$G_X(z;0,\beta) = \frac{(\beta-1)(z\beta-1)}{(\beta z-1)^2} = \frac{1-\beta}{1-\beta z}, \qquad |z| < 1/\beta.$$

Mean. Substituting $\alpha = 0$ into Eq (3.31): Because $(\alpha\beta+\beta+1)\beta^\alpha = \beta+1$ at $\alpha = 0$, the numerator is $(2-(\beta+1))\beta = (1-\beta)\beta$, and the denominator is $(1)(\beta-1)(\beta-1) = (1-\beta)^2$. Hence,

$$\mu_1' = \frac{(1-\beta)\beta}{(1-\beta)^2} = \frac{\beta}{1-\beta}.$$

No indeterminate form arises here either, as $\beta^{\alpha+1} - 1 = \beta - 1 \neq 0$ throughout the admissible range $0 < \beta < 1$.

Variance. Substituting $\alpha = 0$ into Eq (3.35), the bracketed numerator term reduces to

$$2\beta(\beta-1) + 2(1-\beta) - (\beta-1)^2 = (\beta-1)^2,$$

so the full numerator is $\beta(\beta-1)^2$, and the denominator is $(1)^2(\beta-1)^2(\beta-1)^2 = (\beta-1)^4$. Hence,

$$\sigma^2 = \frac{\beta(\beta-1)^2}{(\beta-1)^4} = \frac{\beta}{(1-\beta)^2}.$$

The same substitution mechanism extends termwise to μ_2', μ_3', μ_4', and hence to DI, ZI, CS, and CK, all of which reduce to the closed forms already reported in Table 2, completing the proof. □

Remark 3.2. *(Nested relationship between the DM-WE and geometric distributions) Proposition 3.1 shows that the reduction to the geometric distribution at $\alpha = 0$ is not merely asymptotic: Because the parameter space is $\Theta = \{(\alpha,\beta) : \alpha \geq 0,\ 0 < \beta < 1\}$, the point $\alpha = 0$ lies* inside Θ*, and every closed-form expression in Sections 2 and 3 is finite and well-defined there, with no* $0/0$ *form requiring*

a limiting argument. The one-parameter geometric family is therefore a genuine nested submodel *of the two-parameter DM-WE family:*

$$\{\text{Geometric}(1-\beta) : 0 < \beta < 1\} = \{\text{DM-WE}(\alpha,\beta) : \alpha = 0,\ 0 < \beta < 1\}$$
$$\subset \{\text{DM-WE}(\alpha,\beta) : \alpha \geq 0,\ 0 < \beta < 1\}.$$

Consequently, testing $\alpha = 0$ against $\alpha > 0$ within the DM-WE family is, in principle, a well-posed nested-model comparison rather than a boundary case excluded from the parameter space a structural consequence of $\alpha = 0$ lying in the interior of Θ, distinguishing the exact nesting established in Proposition 3.1 from the asymptotic reduction discussed below. A formal likelihood-ratio implementation is left for future work.

By contrast, the second boundary $\alpha \to \infty$ referenced above is a genuine asymptotic *limit rather than an attained submodel, as $\alpha = \infty \notin \Theta$: For every finite α, however large, $P(x;\alpha,\beta) \neq (1-\beta)\beta^x$ in general, and the geometric form is only approached, not reached, as $\alpha \to \infty$. The DM-WE family thus has the geometric distribution as an* exact, attained *boundary submodel at $\alpha = 0$ (Proposition 3.1) and as an* approached, unattained *asymptotic limit as $\alpha \to \infty$ (Table 2) two distinct senses of "reduction to the geometric distribution" that are worth distinguishing explicitly when the result is used for hypothesis testing or asymptotic interpretation.*

The limiting behavior confirms the following

(1) All derivations in Sections 2 and 3 maintain internal consistency when evaluated at $\alpha = 0$.

(2) The DM-WE family positions the geometric distribution as a strict boundary between constant hazard ($\alpha = 0$) and strictly increasing hazard ($\alpha > 0$).

(3) The pattern of overdispersion ($DI > 1$) and zero-inflation ($ZI > 0$) persists in the geometric limit for all $\beta \in (0, 1)$.

Remark 3.3. *(Monotonicity of ZI and practical zero-inflation bands) Table 5 and a dense grid over $\alpha \in [10^{-6}, 10^{6}]$, $\beta \in (0, 1)$ (Figure 3, panel d) together make the joint dependence of ZI on (α,β) explicit:*

(1) For every fixed α, ZI increases strictly and monotonically with β, with no exception found across the tested grid.

(2) For fixed β, ZI is nonmonotonic in α: It decreases from the $\alpha \to 0$ geometric baseline (Corollary 3.1) to an interior minimum, then rises back toward the same limiting value as $\alpha \to \infty$, consistent with the bathtub-type behavior documented for CS and CK. The location of this minimum shifts toward larger α as β increases: Numerically, $\alpha \approx 1.0$ at $\beta = 0.1$, moving beyond the range tabulated in Table 5 for $\beta \geq 0.7$ (specifically, $\alpha \approx 4$ at $\beta = 0.7$ and $\alpha \approx 7$ at $\beta = 0.9$).

(3) Across the range tabulated in Table 5 ($\alpha \in [0.5, 3.0]$), ZI varies by only 0.008–0.087 *at fixed β, but by* 0.66–0.72 *at fixed α across $\beta \in [0.1, 0.9]$. Hence, β is the dominant driver of the zero-inflation level, with α acting as a secondary refinement; this asymmetry supports using β as a direct, practical indicator of the zero-inflation degree.*

(4) Table 3 reports the corresponding zero-inflation degree at the five tabulated values of β. Because ZI increases continuously and monotonically with β rather than in discrete steps, intermediate values of β produce intermediate ZI values along the same curve, with the exact value at any β given directly by Eq (3.37).

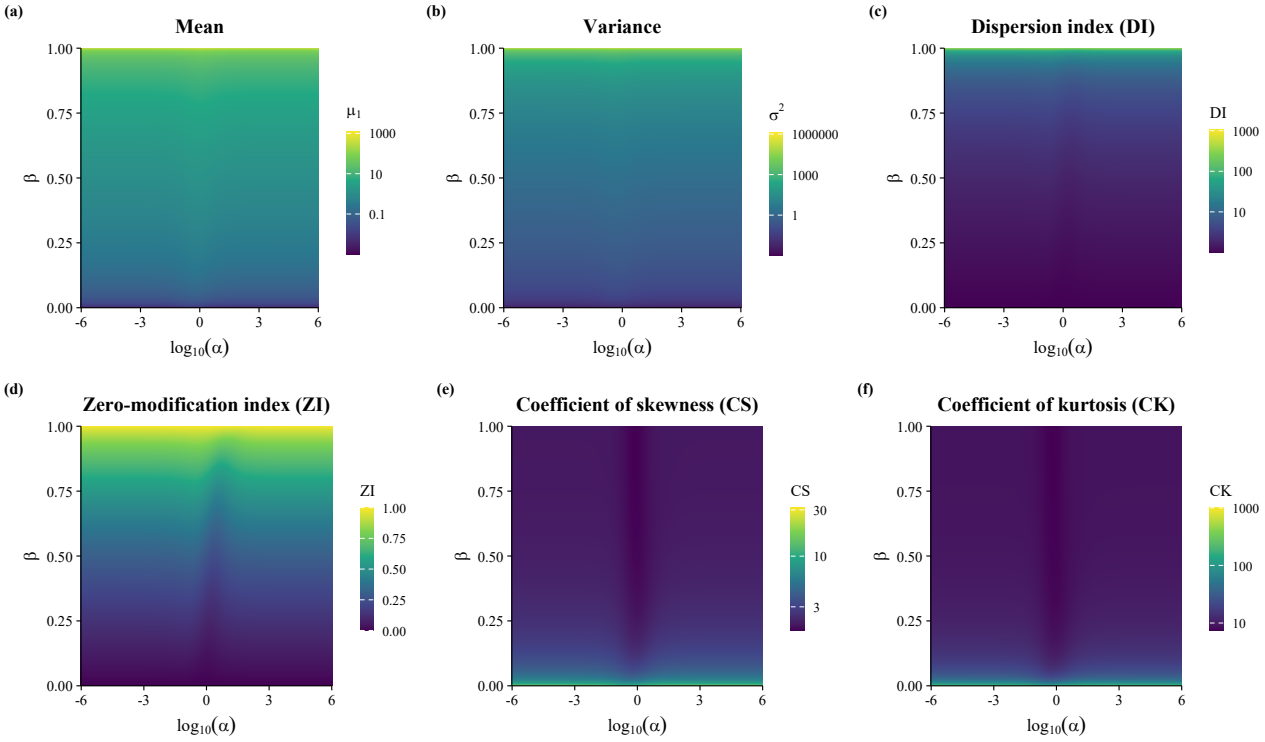


Figure 3. Mean, variance, DI, ZI, CS, and CK of the DM-WE distribution over the full parameter space (α, β).

Table 3. Correspondence between β and zero-inflation degree at the tabulated values in Table 5 ($\alpha \in [0.5, 3.0]$).

β	ZI range across $\alpha \in [0.5, 3.0]$	Zero-inflation degree
0.1	0.031–0.039	Minimal
0.3	0.117–0.151	Mild
0.5	0.236–0.301	Moderate
0.7	0.405–0.492	Substantial
0.9	0.699–0.760	Severe

This correspondence is calibrated to $\alpha \in [0.5, 3.0]$; the one-parameter geometric limit

$$ZI_{geom}(\beta) = 1 + \frac{(1-\beta)\log(1-\beta)}{\beta}$$

of Corollary 3.1 gives the value attained as $\alpha \to 0$ or $\alpha \to \infty$. At $\beta = 0.1, 0.3, 0.5$, the tabulated finite-α values lie below this limit throughout $\alpha \in [0.5, 3.0]$; at $\beta = 0.7, 0.9$, the smallest tabulated α values lie slightly above it, as the interior minimum lies beyond $\alpha = 3.0$, as noted in item 2 above.

Remark 3.4. *(Behavior at the extreme boundary $\beta \to 0$) We examine the $\beta \to 0$ limit separately, as it corresponds to the maximum attainable proportion of zeros. Expanding Eq* (2.6) *to leading order gives*

$$P(X{=}0;\alpha,\beta) \approx 1 - \frac{\alpha+2}{\alpha+1}\beta, \qquad P(X{=}x;\alpha,\beta) \approx \frac{\alpha+2}{\alpha+1}\beta^x \ \ (x \geq 1),$$

so the PMF collapses to a point mass at 0 *with a vanishing geometric tail. Because $P(X{=}0) \to 1$ and $\mu_1' \to 0$ at the same rate, $ZI(\alpha,\beta) \to 0$ for every $\alpha > 0$. More precisely, for every fixed $\alpha > 0$,*

$$ZI(\alpha,\beta) \sim \frac{\alpha}{2(\alpha+1)}\beta, \qquad \beta \to 0^+,$$

a linear rate whose constant depends on α but whose order in β does not. This was confirmed to high numerical precision across $\alpha \in \{0.5, 1, 2, 5, 10\}$, where the asymptotic constant is reached already by $\beta \sim 10^{-6}$. Convergence is not uniform in α, however: For small α (e.g., $\alpha = 0.1$), the same constant is only reached once $\beta \lesssim 10^{-40}$ because the competing correction term of order $\beta^{\alpha+1}$ in Eq (2.6) *decays almost as slowly as β itself when α is close to* 0*. Table 4 gives representative values across $\alpha \in \{0.5, 1, 2, 5\}$, confirming that genuine zero-inflation vanishes as $\beta \to 0$ and that the model can still match any observed zero-proportion via a sufficiently small $\hat{\beta}$.*

Table 4. DM-WE behavior as $\beta \to 0$ across several values of α.

	$\alpha = 0.5$			$\alpha = 1$		
β	$P(X{=}0)$	mean	ZI	$P(X{=}0)$	mean	ZI
0.5	0.402369	1.302055	0.300808	0.3750	1.3333	0.2644
0.1	0.854415	0.163415	0.037187	0.8550	0.1616	0.0307
0.01	0.984000	0.016168	0.002369	0.9851	0.01510	0.00256
0.001	0.998354	0.001647	0.000190	0.9985	0.001501	0.000251
0.0001	0.999834	0.000166	0.000017	0.99985	0.000150	0.0000250
	$\alpha = 2$			$\alpha = 5$		
β	$P(X{=}0)$	mean	ZI	$P(X{=}0)$	mean	ZI
0.5	0.375000	1.285714	0.237133	0.419271	1.164021	0.253245
0.1	0.867000	0.147814	0.034490	0.883333	0.129629	0.043023
0.01	0.986667	0.013468	0.003341	0.988333	0.011785	0.004180
0.001	0.998667	0.001335	0.000333	0.998833	0.001168	0.000417
0.0001	0.999867	0.000133	0.000033	0.999883	0.000117	0.000042

DM-WE can match any zero-proportion, however high, by taking $\hat{\beta}$ small enough. However, ZI shrinks along with it: The model's genuine zero-inflation excess zeros beyond a Poisson with the same mean is concentrated at moderate-to-large β, not at the $\beta \to 0$ extreme, where a high zero-proportion simply reflects a small mean, as it would under a plain Poisson or geometric model. Data with a high zero-proportion arising from a structural-zero mechanism rather than a small mean would require a hurdle or zero-inflated extension of DM-WE (mixing a point mass at 0 *with the DM-WE tail), as noted in Section 7. At the opposite boundary, $\beta \to 1$, Eq* (2.6) *gives $P(X{=}0;\alpha,\beta) \to 0$, and Eq* (3.31) *gives $\mu_1' \to \infty$, so $ZI \to 1$ smoothly and without numerical instability, consistent with the geometric limit $ZI_{geom}(\beta) \to 1$ as $\beta \to 1$ in Corollary 3.1.*

Table 5. Descriptive statistics of the DM-WE distribution.

α	β	μ'_1	μ'_2	μ'_3	μ'_4	σ^2	DI	ZI	CV	CS	CK
	0.1	0.1634	0.2032	0.2962	0.5291	0.1764	1.0797	0.0372	2.5704	2.7704	11.7530
	0.3	0.5832	1.1439	3.0221	10.3740	0.8038	1.3782	0.1512	1.5373	1.9669	8.2217
0.5	0.5	1.3021	4.2366	19.4510	116.490	2.5412	1.9517	0.3008	1.2243	1.8063	7.6887
	0.7	2.9466	18.4310	165.140	1927.50	9.7485	3.3085	0.4919	1.0596	1.7538	7.5255
	0.9	11.1060	235.620	7177.70	284362.00	112.270	10.1090	0.7595	0.9541	1.7374	7.4754
	0.1	0.1616	0.1986	0.2847	0.4997	0.1724	1.0669	0.0307	2.5694	2.7505	11.5930
	0.3	0.5934	1.1347	2.9227	9.8287	0.7825	1.3187	0.1251	1.4907	1.9079	8.0294
1.0	0.5	1.3333	4.2222	18.8890	110.740	2.4444	1.8333	0.2644	1.1726	1.7638	7.6239
	0.7	3.0196	18.4290	160.920	1837.80	9.3118	3.0838	0.4586	1.0106	1.7257	7.5305
	0.9	11.3680	236.190	7014.40	271789.00	106.950	9.4079	0.7449	0.9097	1.7155	7.5063
	0.1	0.1543	0.1889	0.2695	0.4705	0.1650	1.0697	0.0319	2.6331	2.8251	12.0940
	0.3	0.5793	1.0914	2.7766	9.2630	0.7559	1.3049	0.1173	1.5009	1.9306	8.2076
1.5	0.5	1.3141	4.0772	17.9790	104.490	2.3503	1.7886	0.2449	1.1666	1.7885	7.8353
	0.7	2.9887	17.8470	153.470	1736.20	8.9144	2.9827	0.4336	0.9989	1.7600	7.7843
	0.9	11.2720	229.260	6701.10	257038.00	102.190	9.0654	0.7295	0.8968	1.7548	7.7789
	0.1	0.1478	0.1807	0.2576	0.4489	0.1589	1.0749	0.0345	2.6967	2.9031	12.6350
	0.3	0.5622	1.0515	2.6597	8.8449	0.7354	1.3082	0.1178	1.5254	1.9689	8.4279
2.0	0.5	1.2857	3.9388	17.2390	99.8190	2.2857	1.7778	0.2371	1.1759	1.8223	8.0446
	0.7	2.9371	17.2740	147.290	1659.30	8.6474	2.9442	0.4180	1.0012	1.7996	8.0197
	0.9	11.1030	222.280	6437.70	245781.00	98.9950	8.9158	0.7166	0.8961	1.7983	8.0290
	0.1	0.1426	0.1745	0.2486	0.4332	0.1541	1.0796	0.0368	2.7499	2.9691	13.1050
	0.3	0.5467	1.0189	2.5705	8.5364	0.7199	1.3169	0.1213	1.5519	2.0070	8.6311
2.5	0.5	1.2580	3.8241	16.6690	96.3530	2.2415	1.7818	0.2359	1.1901	1.8531	8.2147
	0.7	2.8850	16.7920	142.490	1602.10	8.4692	2.9356	0.4094	1.0087	1.8332	8.2037
	0.9	10.9310	216.340	6231.60	237355.00	96.8620	8.8615	0.7066	0.9004	1.8349	8.2240
	0.1	0.1389	0.1697	0.2418	0.4212	0.1505	1.0834	0.0386	2.7932	3.0228	13.4960
	0.3	0.5337	0.9928	2.5015	8.3021	0.7080	1.3266	0.1254	1.5767	2.0411	8.8081
3.0	0.5	1.2333	3.7311	16.2260	93.7150	2.2100	1.7919	0.2378	1.2054	1.8789	8.3489
	0.7	2.8377	16.3990	138.750	1558.40	8.3465	2.9413	0.4054	1.0181	1.8599	8.3429
	0.9	10.7730	211.450	6069.90	230909.00	95.3950	8.8549	0.6989	0.9066	1.8638	8.3713

Remark 3.5. *Based on Corollary (3.1), we note that as $\alpha \to 0$ and $\alpha \to \infty$, the DM-WE distribution reduces exactly to the geometric distribution with parameter β, for which*

$$CS_{geom}(\beta) = \frac{1+\beta}{\sqrt{\beta}}, \qquad CK_{geom}(\beta) = 9 + \frac{(1-\beta)^2}{\beta}.$$

Both DM-WE limits were verified analytically to coincide exactly with these closed forms for every $\beta \in (0,1)$ tested, so the two ends of the curve in $\log(\alpha)$ *are fixed precisely, not merely approximately. Between these two exact endpoints, CS and CK dip below the corresponding geometric value and return to it, a pattern verified numerically at every tested $\beta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and consistent with a bathtub shape in* $\log(\alpha)$.

To assess whether this pattern extends across the entire admissible parameter space and whether any platykurtic region exists, we conducted an exhaustive numerical search over $\alpha \in [10^{-6}, 10^{6}]$ *and* β *up to* $1 - 10^{-30}$*, using extended-precision arithmetic to avoid the catastrophic cancellation that the* $1/(\beta-1)^k$ *terms in Eqs* (3.32)–(3.34) *are prone to near* $\beta \to 1$ *at standard machine precision. The global minimum located over this search was* $CS \approx 1.709$ *and* $CK \approx 7.410$*, occurring at* $\alpha \approx 0.7$–0.8 *as* $\beta \to 1$*; as* $\beta \to 0$*, both CS and CK diverge to infinity for every fixed* α*. No parameter combination in the search yielded* $CK \leq 3$*, nor even* $CK < 7$*; the distribution therefore remained strictly leptokurtic* ($CK > 7$) *throughout the tested domain, with no platykurtic subregion located.*

We emphasize that this constitutes strong numerical evidence over an extensive grid, not a closed-form proof that the bathtub shape or the $CK > 7$ *bound holds for every* (α, β) *in the open parameter space* $(0, \infty) \times (0, 1)$*. A fully rigorous argument would require establishing the sign of* $\partial CK/\partial\alpha$ *analytically near each boundary and showing it changes sign at most once; the quartic dependence of* μ_2', μ_3', μ_4' *in Eqs* (3.32)–(3.34) *on* α *and* β *makes this derivative analytically intractable in closed form. This limitation parallels the aging-mechanism restriction already noted for the hazard rate in Remark 3.1: We report the exact boundary behavior and an exhaustive search as the strongest evidence currently available, and leave a fully rigorous proof of the bathtub shape, together with a closed-form lower bound on CK, as an open problem for future work.*

The joint effect of α and β on distributional shape was already illustrated in Figure 1; Figure 4 now makes the resulting bathtub pattern in CS and CK explicit.

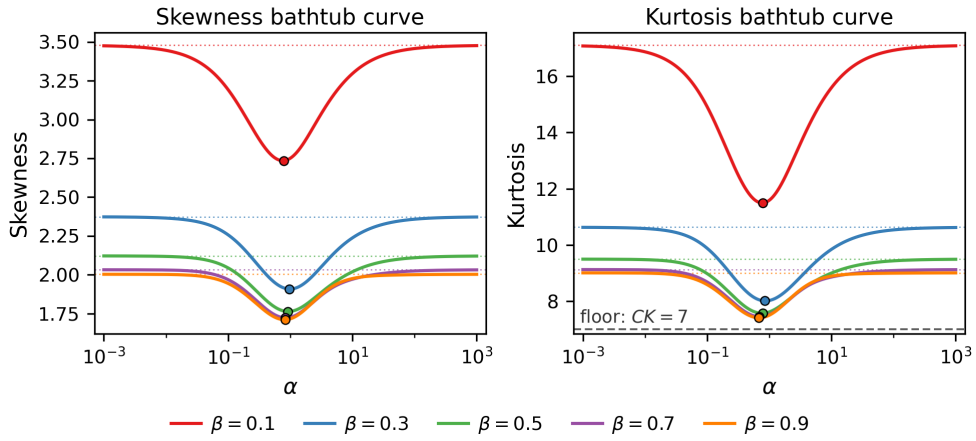


Figure 4. CS and CK versus α for five values of β; markers show the minima.

Remark 3.6. *(Statistical interpretation of the bathtub-shaped CS and CK curves) Figure 4 makes the bathtub pattern reported numerically above visually explicit, and admits a direct practical reading. Because both curves are anchored at the same geometric-limit value at each end (Corollary 3.1) and dip to a lower interior value in between, the shape parameter* α *never removes skewness or kurtosis relative to the geometric baseline, it can only moderate it, and only over an intermediate range of* α*.*

Descending segment *(*α ***small).*** *Here, the distribution is close to its geometric limit: mass concentrated near zero with a long, thin geometric tail (Table 2). As* α *increases from* 0 *toward the trough, the extra shape parameter pulls a small amount of probability mass out of the extreme tail, moderately taming both skewness and kurtosis. Practically, this segment corresponds to data that are strongly right-skewed with a geometric-like decay, for example, counts of rare, roughly memoryless*

events such as daily insurance claims in a low-claim period or defect counts on a stable production line, where a small-to-moderate $\hat{\alpha}$ improves on a plain geometric fit without qualitatively changing the tail behavior.

***Trough (near** $\alpha \approx 0.7$–1**).** This is where the DM-WE distribution most departs from its geometric bounds: CS and CK reach their lowest attainable values for the given β, though never below the strictly leptokurtic floor $CK > 7$ documented numerically above. Data best described by this region show a discernible but comparatively less extreme asymmetry than a pure geometric model. For instance, the corn borer counts analyzed in Application II ($\hat{\alpha} = 0.553$) sit close to this trough region, consistent with the moderately over-dispersed, right-skewed but not heavy-tailed shape observed there.*

***Ascending segment (**α **large).** As α grows further, CS and CK rise back toward the same geometric-limit value, as the distribution reapproaches the geometric shape from the other side (Corollary 3.1). This segment corresponds to count data that are again strongly right-skewed and heavy-tailed but where a large fitted $\hat{\alpha}$ indicates a mechanism closer to memoryless than the moderated trough region for example, rare high-severity event counts such as catastrophic failure counts or extreme claim counts.*

In short, the distance of a fitted $\hat{\alpha}$ from the trough gives practitioners a direct, nontechnical reading of how far the data have departed from geometric-type behavior and in which direction: Values far from the trough (either very small or very large) signal data close to a geometric count process, whereas values near $\alpha \approx 0.7$–1 signal data where the extra shape parameter is doing genuine work, moderating skewness and kurtosis below what a geometric model alone could capture.

3.6. Order statistics

Order statistics are essential in numerous engineering and statistical domains, especially in reliability theory, quality control, and lifetime data analysis. Consider a random sample $X_1, X_2, \ldots, X_n$ of size n extracted from the DM-WE distribution. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ represent the ordered sample elements, known as the order statistics. The CDF of the ith order statistic, represented as $F_{i:n}(x;\alpha,\beta)$, can be explicitly formulated utilizing a binomial expansion framework:

$$\begin{aligned} F_{i:n}(x;\alpha,\beta) &= \sum_{k=i}^{n} \frac{n!}{k!(n-k)!}[F(x;\beta)]^k[1-F(x;\beta)]^{n-k} \\ &= \sum_{k=i}^{n}\sum_{j=0}^{k} \tau_j^{(n,k)} \left[\frac{\beta^{\alpha(x+1)} - \beta^{\alpha(x+1)+1} + (\alpha+2)\beta^{\alpha+1} - (\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right]^{n+j-k}, \end{aligned} \tag{3.38}$$

where the linear combination components are weighted by the coefficient

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

Consequently, the PMF of the ith order statistic is obtained by evaluating the forward difference of its CDF, expressed as

$$\begin{aligned} f_{i:n}(x;\beta) &= F_{i:n}(x;\beta) - F_{i:n}(x-1;\beta) \\ &= \sum_{k=i}^{n}\sum_{j=0}^{k} \tau_j^{(n,k)} \left[\left(\frac{\beta^{\alpha(x+1)} - \beta^{\alpha(x+1)+1} + (\alpha+2)\beta^{\alpha+1} - (\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right)^{n+j-k}\right. \end{aligned}$$

$$-\left(\frac{\beta^{\alpha x}-\beta^{\alpha x+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x}\right)^{n+j-k}\Bigg]. \tag{3.39}$$

Utilizing the algebraic structure of the PMF, the υth raw moments of the ith order statistic can be written via an infinite series summation form:

$$\mu_{i:n}^{(\upsilon)}=\sum_{x=0}^{\infty}\sum_{k=i}^{n}\sum_{j=0}^{k}\tau_{j}^{(n,k)}x^{\upsilon}\Bigg[\left(\frac{\beta^{\alpha(x+1)}-\beta^{\alpha(x+1)+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right)^{n+j-k}$$
$$-\left(\frac{\beta^{\alpha x}-\beta^{\alpha x+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x}\right)^{n+j-k}\Bigg]. \tag{3.40}$$

Furthermore, following the foundational properties outlined by Arnold et al. [29], the absolute raw moments can be computed directly using the SF counterpart of $F_{i:n}(x;\beta)$ without explicitly processing the difference equations. The first and second moments are given by

$$\mu_{i:n}=\sum_{x=0}^{\infty}[1-F_{i:n}(x;\beta)]\quad\text{and}\quad\mu_{i:n}^{(2)}=2\sum_{x=0}^{\infty}x[1-F_{i:n}(x;\beta)]+\mu_{i:n}. \tag{3.41}$$

In reliability engineering, there are two canonical configurations for a system of n independent identically distributed components with discrete lifetimes. A series system is one in which any component failure leads to system failure; therefore, the system lifetime is the smallest component lifetime, $X_{1:n}=\min(X_1,\ldots,X_n)$. Hence, $\mu_{1:n}$ is the expected number of operating cycles before the first system failure, and is necessarily nonincreasing in n because the weakest-link effect implies that adding components in series cannot improve system reliability. In a parallel (redundant) system, the system fails only when all components fail; therefore, the system lifetime coincides with the largest component lifetime, $X_{n:n}=\max(X_1,\ldots,X_n)$, and $\mu_{n:n}$ is the expected number of cycles until complete system failure, which is necessarily nondecreasing in n, quantifying the reliability gain obtained from redundancy. The two monotonicity properties are immediate consequences of the fact that $[S(x)]^n$ is nonincreasing in n when $S(x)<1$, and are true for any underlying lifetime distribution, not just for the DM-WE model. Both quantities are available in closed form under the DM-WE model through the SF in Eq (2.2); therefore, a reliability engineer can directly evaluate the expected lifetime of either configuration for any component-level parameters (α,β) and any number of components n, without simulation. Their expected values and second raw moments have the following neat expressions:

$$\mu_{1:n}=\sum_{x=0}^{\infty}[1-F(x;\beta)]^n=\sum_{x=0}^{\infty}\left[\frac{\beta^{\alpha(x+1)}-\beta^{\alpha(x+1)+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right]^n, \tag{3.42}$$

$$\mu_{n:n}=\sum_{x=0}^{\infty}[1-\{F(x;\beta)\}^n]=\sum_{x=0}^{\infty}\left[1-\left\{1-\frac{\beta^{\alpha(x+1)}-\beta^{\alpha(x+1)+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right\}^n\right], \tag{3.43}$$

$$\mu_{1:n}^{(2)}=2\sum_{x=0}^{\infty}x\left[\frac{\beta^{\alpha(x+1)}-\beta^{\alpha(x+1)+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right]^n+\mu_{1:n}, \tag{3.44}$$

$$\mu_{n:n}^{(2)}=2\sum_{x=0}^{\infty}x\left[1-\left\{1-\frac{\beta^{\alpha(x+1)}-\beta^{\alpha(x+1)+1}+(\alpha+2)\beta^{\alpha+1}-(\alpha+2)}{(\alpha+1)(\beta^{\alpha+1}-1)}\beta^{x+1}\right\}^n\right]+\mu_{n:n}. \tag{3.45}$$

The structural features and changes of the order statistics are assessed across several combinations of parameters α, β with sample sizes $n = 3$ and $n = 10$. The numerical trends presented in Table 6 provide clear insights:

(1) As the sample size increases from $n = 3$ to $n = 10$, the anticipated value of the minimal order statistic ($\mu_{1:n}$) consistently diminishes across all experimental permutations. In contrast, the predicted value of the maximum order statistic ($\mu_{n:n}$) increases substantially. This precisely aligns with statistical expectations: Higher sample sizes are more likely to produce progressively severe limits, as $\min(X_n)$ approaches zero and $\max(X_n)$ extends into the long right tail.

(2) The scale parameter β has a significantly strong, positive correlation with both the raw moments and the variances ($\sigma^2_{1,1:n}, \sigma^2_{n,n:n}$). At $\beta = 0.1$, the predicted value of the minimal order statistic is confined within a diminutive range (10^{-5} to 10^{-10}), signifying substantial mass concentration at the origin. As β approaches 0.9, the dispersion increases dramatically; for example, at $n = 10$ and $\alpha = 0.5$, $\mu_{n:n}$ escalates from 0.9598 to 31.2510, and its second moment rises to 1126.9000.

(3) For fixed β and n, increasing the shape parameter α typically applies a downward stabilizing influence on the extreme maximum statistics ($\mu_{n:n}$ and $\sigma^2_{n,n:n}$). This reinforces the function of α as a structural modifier that reduces variance and aligns the upper tail with the median domain, hence enhancing the model's reliability limits.

Table 6. Numerical evaluations of order statistics moments and variances across distinct parameter grids of α, β, and sample sizes n.

n	β	α	$\mu_{1:n} \times 10^{-5}$	$\mu_{n:n}$	$\mu^{(2)}_{1:n} \times 10^{-5}$	$\mu^{(2)}_{n:n}$	$\sigma^2_{1,1:n} \times 10^{-5}$	$\sigma^2_{n,n:n}$
5	0.1	0.5	6.5401	0.6313	6.5404	0.8248	6.5399	0.4263
		1.0	6.4098	0.6239	6.4099	0.8041	6.4095	0.4148
		1.5	5.1397	0.6019	5.1398	0.7708	5.1395	0.4085
		2.0	4.1616	0.5824	4.1617	0.7434	4.1615	0.4042
		2.5	3.5011	0.5669	3.5011	0.7224	3.5010	0.4009
		3.0	3.0487	0.5549	3.0488	0.7061	3.0487	0.3982
	0.5	0.5	8050.0	3.3352	8955.0	14.5950	8306.0	3.4721
		1.0	10040.0	3.3194	11080.0	14.3020	10070.0	3.2836
		1.5	10320.0	3.2566	11260.0	13.8070	10190.0	3.2011
		2.0	9930.0	3.1989	10740.0	13.3990	9760.0	3.1651
		2.5	9330.0	3.1518	10040.0	13.0820	9170.0	3.1481
		3.0	8710.0	3.1135	9330.0	12.8330	8570.0	3.1393
	0.9	0.5	215220.0	24.7300	1068300.0	758.3600	605080.0	146.7800
		1.0	248450.0	24.6260	1282700.0	745.0800	665480.0	138.6200
		1.5	261030.0	24.2130	1332900.0	721.3500	651500.0	135.0600
		2.0	263870.0	23.8340	1315700.0	701.5800	619480.0	133.5100
		2.5	262180.0	23.5230	1274100.0	686.1200	586750.0	132.7900
		3.0	258450.0	23.2710	1225800.0	673.9800	557830.0	132.4200
10	0.1		$\mu_{1:n} \times 10^{-9}$		$\mu^{(2)}_{1:n} \times 10^{-9}$		$\sigma^2_{1,1:n} \times 10^{-9}$	
		0.5	4.2772	0.9598	4.2772	1.3345	4.2772	0.4133
		1.0	4.1085	0.9476	4.1085	1.2972	4.1085	0.3992
		1.5	2.6416	0.9224	2.6416	1.2506	2.6416	0.3999
		2.0	1.7319	0.9005	1.7319	1.2138	1.7319	0.4031
		2.5	1.2257	0.8828	1.2257	1.1856	1.2257	0.4063
		3.0	0.9295	0.8689	0.9295	1.1636	0.9295	0.4087
	0.5		$\mu_{1:n} \times 10^{-5}$		$\mu^{(2)}_{1:n} \times 10^{-5}$		$\sigma^2_{1,1:n} \times 10^{-5}$	
		0.5	580.0	4.3263	590.0	22.2720	580.0	3.5554
		1.0	910.0	4.2796	920.0	21.7080	910.0	3.3935
		1.5	980.0	4.2019	980.0	20.9960	970.0	3.3396
		2.0	910.0	4.1375	914.0	20.4400	906.0	3.3211
		2.5	810.0	4.0869	810.0	20.0180	804.0	3.3142
		3.0	710.0	4.0470	710.0	19.6890	703.0	3.3115
	0.9	0.5	80630.0	31.2510	235380.0	1126.9000	170370.0	150.3800
		1.0	115140.0	30.9440	335030.0	1100.9000	202470.0	143.3600
		1.5	127560.0	30.4330	374970.0	1067.2000	212260.0	141.0200
		2.0	133140.0	30.0090	386010.0	1040.8000	208740.0	140.2000
		2.5	134850.0	29.6770	382390.0	1020.6000	200550.0	139.9200
		3.0	134410.0	29.4140	371980.0	1004.9000	191310.0	139.7900

To illustrate the practical reliability implications of Table 6 in a concrete engineering setting, consider a bank of n nominally identical pressure-relief valves, each with a discrete operating-cycle lifetime following the DM-WE distribution with $\alpha = 0.5$ and $\beta = 0.9$. If the valves are installed in a series safety-interlock configuration, where the failure of any single valve trips the entire protection system, the system lifetime is $X_{1:n}$. Table 6 shows that the expected number of operating cycles before the first system-level failure drops from $\mu_{1:5} = 2.1522$ with $n = 5$ valves to $\mu_{1:10} \approx 0.9142$ with $n = 10$ valves, illustrating the well-known weakest-link penalty: Adding more components in series accelerates the expected time to the first failure, as the system is only as reliable as its least reliable element.

Conversely, if the same n valves are instead installed in a parallel redundant configuration, where the protection system fails only once every valve has failed, the system lifetime is $X_{n:n}$. Table 6 shows the expected number of cycles until complete system failure increases from $\mu_{5:5} = 24.7300$ with $n = 5$ valves to $\mu_{10:10} = 31.2510$ with $n = 10$ valves, quantifying the reliability gain obtained from redundancy. This series-versus-parallel contrast is a foundational design principle in reliability engineering, applicable to component groupings such as redundant pump stations, parallel relay circuits in safety-critical control systems, or backup power units: A series configuration compounds the weaknesses of its components, driving the expected time to first failure down as more units are added, whereas a parallel configuration compounds their strengths, driving the expected time to complete failure up. The closed-form expressions in Equations (3.42)–(3.45) allow a reliability engineer to directly evaluate this trade-off for any component-level parameters (α, β) and any number of components n, without resorting to simulation, supporting design decisions such as the number of redundant units needed to meet a target system lifetime.

3.7. Entropy analysis

Entropy is a fundamental notion in contemporary scientific frameworks, with significant theoretical and practical importance in both statistical physics and information theory. Entropy is fundamentally linked to structural disorder, stochastic unpredictability, and systemic uncertainty, offering a precise mathematical framework to measure the latent missing information about a system's microstates based on its macrostate definition. It quantifies the level of uncertainty or anticipated surprise linked to the outcome of a random variable, particularly in contexts defined by inadequate information or limited observability. Originally defined in classical thermodynamics to describe energy dissipation and macroscopic irreversibility, the application of entropy has extended beyond its initial thermodynamic context. It has been extensively utilized as an analytical instrument across various disciplines, including meteorology, statistical physics, chemical kinetics, microeconomics, theoretical biology, physical cosmology, climate dynamics, and digital information systems, particularly in data compression, cryptography, and telecommunications. For a comprehensive mathematical analysis of the fundamental properties, axiomatic definitions, and analytical characteristics of generalized entropy measures, the reader is directed to the landmark overview by [30]. This section focuses on analyzing specific parametric and nonparametric versions of entropy, particularly Rényi entropy (RE), maximum entropy (Max-E), Shannon entropy (SE), minimum entropy (Min-E), and collision entropy (CE). In the realm of information theory, these variations function as essential operational metrics of information content and distributional variability (see, [31]). Let X be a discrete random variable defined on the set of natural numbers $\mathbb{N}$, following the DM-WE distribution, characterized by the PMF $P(x; \alpha, \beta)$. The RE of order ρ for this

distribution is defined as follows:

$$RE = \frac{1}{1-\rho}\log\left(\sum_{x=0}^{\infty}[P(x;\alpha,\beta)]^{\rho}\right)$$
$$= \frac{1}{1-\rho}\log\sum_{x=0}^{\infty}\left(\frac{-\alpha\beta+\alpha-2\beta-\beta^{\alpha x}+\beta^{\alpha+\alpha x+1}+2}{\alpha+1}\beta^{x}\right)^{\rho}, \quad x\in\mathbb{N}, \tag{3.46}$$

for $\rho > 0$, $\rho \neq 1$. The parameter ρ functions as a tilting mechanism that modifies the sensitivity of the entropy measure to the underlying probabilities: Elevated values of ρ disproportionately penalize highly probable outcomes, whereas diminished values distribute weights more uniformly across possibilities. Thus, the RE functional serves as an extended unifying framework that includes many classical entropy metrics as limiting cases of the order parameter ρ. These foundational measurements arise under the following asymptotic behaviors:

- As $\rho \to 0$, the RE converges to the Hartley or Max-E, which reflects the logarithm of the cardinality of the distribution's support, treating all outcomes with nonzero probability equally.
- As $\rho \to 1$, applying L'Hôpital's rule demonstrates that RE smoothly converges to the classical SE, capturing the standard expected information content.
- As $\rho \to 2$, the RE simplifies to the CE, which directly quantifies the probability of two independent trials, yielding identical outcomes.
- As $\rho \to \infty$, the RE converges to the Min-E, which is strictly determined by the single most probable outcome, serving as a conservative bound on uncertainty in cryptographic contexts.

By evaluating the limiting case where $\rho \to 1$, the analytical expression for the SE of the DM-WE distribution for $x \in \mathbb{N}$ is derived as follows:

$$SE = -\sum_{x=0}^{\infty}\left(\frac{-\alpha\beta+\alpha-2\beta-\beta^{\alpha x}+\beta^{\alpha+\alpha x+1}+2}{\alpha+1}\beta^{x}\right)\log\left(\frac{-\alpha\beta+\alpha-2\beta-\beta^{\alpha x}+\beta^{\alpha+\alpha x+1}+2}{\alpha+1}\beta^{x}\right). \tag{3.47}$$

Owing to the intricate analytical architecture of the DM-WE PMF, deriving closed-form expressions for these entropy metrics poses severe mathematical intractability. To circumvent this and validate the theoretical model, extensive numerical computations are executed. These numerical simulations serve a dual purpose: First, they empirically verify the convergence properties of the infinite series definitions under varying parameter regimes; second, they provide valuable insights into the directional behavior of the entropy surfaces relative to perturbations in the model parameters α and β, and the order ρ. The comprehensive numerical outcomes for a broad spectrum of parameter configurations are tabulated in Tables 7 through 9.

Because Eqs (3.46) and (3.47) involve infinite series over the support $\mathbb{N}$, their numerical evaluation requires a finite truncation of the summation. Rather than truncating on the magnitude of the individual entropy terms, which need not decay monotonically, truncation is based on the remaining probability mass: The summation is carried out up to $X_{\max}$, the smallest integer such that $S(X_{\max};\alpha,\beta) < \delta$ with $\delta = 10^{-12}$, using the closed-form survival function in Eq (2.2). Given the geometric-type decay of the DM-WE tail, this criterion is reached rapidly across the entire tested parameter grid, and the resulting

truncation error in the reported entropy values is negligible relative to the four decimal places given in Tables 7–9; stability was confirmed by verifying that tightening δ by two further orders of magnitude left the reported values unchanged.

Table 7. The RE of the DM-WE distribution at $\rho = 0.5$ for various values of the parameters.

$\alpha \downarrow$	$\beta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5		0.7575	1.0817	1.3574	1.6227	1.8972	2.1996	2.5571	3.0251	3.7750
1		0.7499	1.0763	1.3533	1.6193	1.8941	2.1967	2.5542	3.0223	3.7722
1.5		0.7367	1.0626	1.3398	1.6058	1.8806	2.1829	2.5403	3.0082	3.7580
2		0.7255	1.0501	1.3270	1.5929	1.8677	2.1699	2.5271	2.9947	3.7443
2.5		0.7166	1.0398	1.3163	1.5822	1.8568	2.1589	2.5159	2.9833	3.7327
3		0.7096	1.0314	1.3074	1.5732	1.8478	2.1498	2.5067	2.9739	3.7231

Table 8. The RE of the DM-WE distribution at $\rho = 1.5$ for various values of the parameters.

$\alpha \downarrow$	$\beta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5		0.3529	0.6331	0.8918	1.1486	1.4183	1.7179	2.0737	2.5407	3.2902
1		0.3509	0.6405	0.9069	1.1688	1.4413	1.7423	2.0987	2.5660	3.3155
1.5		0.3373	0.6253	0.8933	1.1566	1.4296	1.7302	2.0858	2.5522	3.3012
2		0.3249	0.6082	0.8753	1.1389	1.4120	1.7120	2.0666	2.5319	3.2800
2.5		0.3152	0.5932	0.8583	1.1217	1.3949	1.6946	2.0482	2.5124	3.2597
3		0.3075	0.5807	0.8435	1.1063	1.3796	1.6792	2.0322	2.4954	3.2418

Table 9. The SE of the DM-WE distribution for various values of the parameters.

$\alpha \downarrow$	$\beta \rightarrow$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5		0.4713	0.7722	1.0389	1.2995	1.5713	1.8721	2.2286	2.6961	3.4457
1		0.4675	0.7754	1.0458	1.3089	1.5822	1.8837	2.2405	2.7081	3.4577
1.5		0.4531	0.7594	1.0312	1.2951	1.5684	1.8696	2.2259	2.6931	3.4426
2		0.4402	0.7434	1.0147	1.2787	1.5519	1.8528	2.2085	2.6749	3.4237
2.5		0.4299	0.7296	0.9998	1.2636	1.5369	1.8374	2.1927	2.6584	3.4059
3		0.4219	0.7182	0.9870	1.2505	1.5238	1.8243	2.1791	2.6443	3.3923

Based on the comprehensive numerical simulation data compiled in Tables 7–9, several critical insights regarding the information-theoretic behavior of the DM-WE distribution can be formalized:

(1) Across all evaluated parameter landscapes, holding the shape parameter α static while varying β from 0.1 to 0.9 yields a prominent, monotonic escalation in both RE and SE values. This behavior indicates that β primarily dictates the tail behavior and dispersion of the PMF; as $\beta \rightarrow 1$, the distribution yields higher structural volatility and uncertainty, thereby maximizing information capacity.

(2) In contrast to the behavior of β, increasing the parameter α while keeping β fixed induces a systematic decrease in the overall entropy. This inverse relationship implies that higher values of α contract the distribution's effective support, concentrating the probability mass and making the system more predictable.

(3) A cross-table comparative analysis confirms the strict inequality $RE_{\rho=0.5} > SE > RE_{\rho=1.5}$ for any given pair of (α, β). This observation is in perfect alignment with the axiomatic foundations of Rényi entropy, wherein higher values of the entropic order ρ heavily penalize events with low probabilities, ultimately shrinking the calculated uncertainty.

The smooth, unbroken gradients observed across all tables provide solid empirical proof of the numerical convergence of the infinite series defined in Eqs (3.46) and (3.47). This underscores the viability of utilizing the DM-WE distribution framework within active information-theoretic, cryptographic, and reliability engineering pipelines.

4. MLE and asymptotic properties

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a random sample of size n drawn from the proposed distribution. Let $\boldsymbol{\theta} = (\alpha, \beta)^\top$ denote the parameter vector. The likelihood function $L(\boldsymbol{\theta}|\mathbf{x})$ for the observed sample is the product of the individual PMFs:

$$L(\alpha, \beta|\mathbf{x}) = \prod_{i=1}^{n} \left[\frac{2 + \alpha - \alpha\beta - 2\beta - \beta^{\alpha x_i} + \beta^{\alpha x_i + \alpha + 1}}{\alpha + 1} \beta^{x_i} \right]. \tag{4.1}$$

The log-likelihood function, $\ell(\boldsymbol{\theta}|\mathbf{x}) = \log L(\alpha, \beta|\mathbf{x})$, is given by:

$$\ell(\alpha, \beta|\mathbf{x}) = \sum_{i=1}^{n} \left[\log\left(2 + \alpha - \alpha\beta - 2\beta - \beta^{\alpha x_i} + \beta^{\alpha x_i + \alpha + 1}\right) - \log(\alpha + 1) + x_i \log\beta \right]. \tag{4.2}$$

To find the MLEs $\hat{\alpha}$ and $\hat{\beta}$, we must solve the nonlinear score equations $\frac{\partial \ell}{\partial \alpha} = 0$ and $\frac{\partial \ell}{\partial \beta} = 0$. Let the numerator of the PMF be denoted as

$$N(x_i; \alpha, \beta) = 2 + \alpha - \alpha\beta - 2\beta - \beta^{\alpha x_i} + \beta^{\alpha x_i + \alpha + 1}.$$

The first partial derivatives are:

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \left[\frac{1 - \beta - x_i \beta^{\alpha x_i} \log\beta + (x_i + 1)\beta^{\alpha x_i + \alpha + 1} \log\beta}{N(x_i; \alpha, \beta)} - \frac{1}{\alpha + 1} \right] = 0, \tag{4.3}$$

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} \left[\frac{-\alpha - 2 - \alpha x_i \beta^{\alpha x_i - 1} + (\alpha x_i + \alpha + 1)\beta^{\alpha x_i + \alpha}}{N(x_i; \alpha, \beta)} + \frac{x_i}{\beta} \right] = 0. \tag{4.4}$$

Analytical solutions for $\hat{\alpha}$ and $\hat{\beta}$ are intractable due to the intricacy of Eqs (4.3) and (4.4). Consequently, numerical optimization methods, such the Newton–Raphson algorithm and the Nelder–Mead method, must be utilized to maximize the log-likelihood function under the restrictions $\alpha \geq 0$ and $0 < \beta < 1$. The asymptotic variance-covariance matrix of the MLEs is defined by the Fisher information matrix.

The observed Fisher information matrix, $\mathcal{J}(\boldsymbol{\theta})$, is characterized as the negative Hessian matrix of the log-likelihood function:

$$\mathcal{J}(\boldsymbol{\theta}) = -\begin{pmatrix} \frac{\partial^2 \ell}{\partial \alpha^2} & \frac{\partial^2 \ell}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ell}{\partial \beta \partial \alpha} & \frac{\partial^2 \ell}{\partial \beta^2} \end{pmatrix}. \tag{4.5}$$

The expected Fisher information matrix is denoted as $\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}[\mathcal{J}(\boldsymbol{\theta})]$. Owing to the complex nature of analytically evaluating the expectations of second derivatives, it is standard practice in applied statistics to approximate the expected Fisher information matrix by employing the observed Fisher information matrix assessed at the MLEs, denoted as $\mathcal{J}(\hat{\boldsymbol{\theta}})$. The support of the distribution ($x \in \mathbb{N}$) is independent of the parameters α and β, thereby fulfilling the standard regularity conditions for MLEs. Given these conditions, the MLEs $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta})^\top$ are consistent and asymptotically conform to a normal distribution. Specifically, as n tends to infinity,

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}_2\left(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta})^{-1}\right), \tag{4.6}$$

where $\mathcal{N}_2$ denotes the bivariate normal distribution, and $\mathbf{0}$ is a 2×1 vector of zeros. For large finite samples, we can approximate the distribution of the estimators as

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \mathcal{J}(\hat{\boldsymbol{\theta}})^{-1}\right). \tag{4.7}$$

Let $\mathcal{V} = \mathcal{J}(\hat{\boldsymbol{\theta}})^{-1}$ denote the estimated asymptotic variance–covariance matrix, where the diagonal elements $v_{\alpha\alpha}$ and $v_{\beta\beta}$ signify the estimated variances of $\hat{\alpha}$ and $\hat{\beta}$, respectively. A simplistic method for creating a $100(1-\gamma)\%$ Wald confidence interval for the parameters is $\hat{\theta} \pm z_{\gamma/2}\sqrt{v_{\theta\theta}}$, where $z_{\gamma/2}$ denotes the upper $\gamma/2$ quantile of the standard normal distribution.

5. Model-based simulation

This section reports an extensive Monte Carlo study designed to assess the finite-sample behavior of the MLEs for the parameters of the proposed model. The sampling properties of the estimators are investigated for small, moderate, and large samples over several distinct parameter configurations. The experiment is conducted across nine sample sizes, $n \in \{25, 50, 100, 150, 200, 250, 300, 400, 500\}$. For each pairing of sample size n and parameter setting, independent synthetic datasets are generated from the PMF of the proposed model, with $N = 10{,}000$ replications for Schemas 1–4 and $N = 5{,}000$ replications for Schemas 5 and 6. The behavior of the MLEs is explored throughout the parameter space by means of the following schemas:

- **Schema 1:** $\alpha = 0.5, \beta = 0.4$ (Baseline range);
- **Schema 2:** $\alpha = 2.0, \beta = 0.25$ (High α, low β);
- **Schema 3:** $\alpha = 1.2, \beta = 0.75$ (Moderate α, high β);
- **Schema 4:** $\alpha = 3.5, \beta = 0.6$ (Heavy-tailed region).

To further probe estimator stability, two additional schemas are constructed in which the generated samples are deliberately contaminated with outlying counts:

- **Schema 5:** $\alpha = 0.01, \beta = 0.9$ (Mild contamination). Here a small proportion of the observations are replaced by extreme counts, allowing us to gauge whether the estimator retains acceptable accuracy and valid interval coverage in a setting where the standard MLE already begins to deteriorate.
- **Schema 6:** $\alpha = 0.1, \beta = 0.95$ (Severe contamination). The same procedure is applied with a larger contamination fraction and more extreme outlying values so as to expose the estimator's resilience under conditions in which the standard MLE breaks down entirely.

For a given schema, let $\theta = (\alpha, \beta)^T$ denote the true parameter vector, and let

$$\hat{\theta}^{(i)} = (\hat{\alpha}^{(i)}, \hat{\beta}^{(i)})^T$$

be the estimate produced in the ith replication, $i = 1, \ldots, N$. The quality, precision, and asymptotic consistency of both the point and the interval estimators are then judged using seven complementary statistical criteria. Throughout, N denotes the total number of Monte Carlo replications, θ_j the jth component of the true parameter vector, and $\hat{\theta}_j^{(i)}$ its corresponding maximum likelihood estimate in replication i. The central tendency and systematic accuracy of the point estimators are summarized through the average estimate (AE) and the average bias (AB), where the AE is the empirical mean of the estimator over all independent replications, defined as:

$$\mathrm{AE}(\hat{\theta}_j) = \frac{1}{N} \sum_{i=1}^{N} \hat{\theta}_j^{(i)}. \tag{5.1}$$

To quantify the structural directional error with finite samples, the AB assesses the direct systematic divergence of these empirical realizations from the genuine underlying baseline parameter value as follows:

$$\mathrm{AB}(\hat{\theta}_j) = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\theta}_j^{(i)} - \theta_j \right). \tag{5.2}$$

To assess overall estimation accuracy by examining both systemic bias and variance concurrently, we calculate the mean squared error (MSE) and its scale-standardized equivalent, the root mean squared error (RMSE). The MSE assesses the aggregate penalty of point variability and squared bias by the subsequent equation:

$$\mathrm{MSE}(\hat{\theta}_j) = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\theta}_j^{(i)} - \theta_j \right)^2. \tag{5.3}$$

The quadratic nature of the MSE alters the original measurement unit; thus, the scale is reverted to the native parameter dimension using the RMSE, defined as

$$\mathrm{RMSE}(\hat{\theta}_j) = \sqrt{\mathrm{MSE}(\hat{\theta}_j)}. \tag{5.4}$$

In addition to these error boundaries, the inherent stochastic variability of the point estimators, independent of parameter-centered bias, is monitored via the sample standard deviation (SD), defined as

$$\mathrm{SD}(\hat{\theta}_j) = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \left(\hat{\theta}_j^{(i)} - \mathrm{AE}(\hat{\theta}_j) \right)^2}. \tag{5.5}$$

The characteristics and accuracy of the interval estimators are examined through the average coverage probability (CP) and the average confidence interval length (IL). The CP determines the empirical chance that an asymptotic 95% Wald confidence interval accurately encompasses the true parameter θ_j. It is calculated across all simulated datasets as follows:

$$\mathrm{CP}(\theta_j) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}\left(\theta_j \in [\mathrm{LCL}_j^{(i)}, \mathrm{UCL}_j^{(i)}]\right), \tag{5.6}$$

where $\mathbb{I}(\cdot)$ signifies the standard indicator function, and $\mathrm{LCL}_j^{(i)}$ and $\mathrm{UCL}_j^{(i)}$ imply the lower and upper confidence limits, respectively, calculated using the inverse of the observed information matrix. To assess the efficiency and precise convergence of these interval boundaries, the IL functions as the principal metric of accuracy by calculating the average range of the produced interval bands:

$$\mathrm{IL}(\theta_j) = \frac{1}{N}\sum_{i=1}^{N} \left(\mathrm{UCL}_j^{(i)} - \mathrm{LCL}_j^{(i)}\right). \tag{5.7}$$

An optimal estimator exhibits asymptotic efficiency when both RMSE and IL tend to zero, whereas the CP approaches its nominal target of 0.95 as the sample size escalates. The empirical results from the Monte Carlo simulation research are methodically outlined in Tables 10–13 and further corroborated by the visual trajectories depicted in Figures 5 and 6. The heatmap matrix in Figure 6 facilitates a direct cross-schema and cross-parameter comparison, presenting an immediate overview of the variations in key performance metrics, specifically root mean squared error and coverage probability, across the four configurations for any specified sample size.

The extensive simulation results displayed in Tables 10 to 15 reveal tendencies that corroborate the efficacy of our estimators.

(1) For all selected schemas, when the sample size n increases from 25 to 500, the AB consistently approaches zero. This supports the assertion that the MLEs are asymptotically unbiased.

(2) The MSE, RMSE, and empirical SD consistently decrease as n increases, demonstrating a distinct $O(n^{-1/2})$ convergence rate as prescribed by conventional asymptotic theory.

(3) The average confidence interval considerably narrows with greater sample sizes, indicating enhanced precision. Importantly, the empirical confidence procedure reliably converges to the nominal 95% standard for larger samples, confirming the validity of the asymptotic normal approximations based on the observed Fisher information matrix.

Table 10. Simulation results for Schema 1: $\alpha = 0.5, \beta = 0.4$.

n	Parameter	AE	AB	MSE	RMSE	CP (95%)	IL	SD
25	α	0.5594	0.0594	0.0560	0.2366	0.9340	0.9080	0.2312
	β	0.3976	-0.0024	0.0097	0.0983	0.9410	0.3783	0.0983
50	α	0.5297	0.0297	0.0263	0.1623	0.9410	0.6133	0.1596
	β	0.3988	-0.0012	0.0047	0.0683	0.9470	0.2608	0.0683
100	α	0.5149	0.0149	0.0123	0.1111	0.9460	0.4228	0.1101
	β	0.3994	-0.0006	0.0023	0.0477	0.9510	0.1812	0.0477
150	α	0.5099	0.0099	0.0081	0.0901	0.9490	0.3424	0.0896
	β	0.3996	-0.0004	0.0015	0.0388	0.9520	0.1471	0.0388
200	α	0.5074	0.0074	0.0060	0.0776	0.9510	0.2954	0.0772
	β	0.3997	-0.0003	0.0011	0.0335	0.9530	0.1270	0.0335
250	α	0.5059	0.0059	0.0048	0.0692	0.9520	0.2635	0.0690
	β	0.3998	-0.0002	0.0009	0.0300	0.9540	0.1133	0.0300
300	α	0.5049	0.0049	0.0040	0.0631	0.9530	0.2401	0.0629
	β	0.3998	-0.0002	0.0007	0.0273	0.9550	0.1033	0.0273
400	α	0.5037	0.0037	0.0030	0.0545	0.9540	0.2074	0.0544
	β	0.3999	-0.0001	0.0005	0.0236	0.9560	0.0892	0.0236
500	α	0.5030	0.0030	0.0024	0.0487	0.9550	0.1852	0.0486
	β	0.3999	-0.0001	0.0004	0.0211	0.9570	0.0797	0.0211

Table 11. Simulation results for Schema 1: $\alpha = 2.0, \beta = 0.25$.

n	Parameter	AE	AB	MSE	RMSE	CP (95%)	IL	SD
25	α	2.0252	0.0252	0.5369	0.7327	0.9280	2.8124	0.7323
	β	0.2485	-0.0015	0.0055	0.0740	0.9390	0.2847	0.0740
50	α	2.0126	0.0126	0.2520	0.5020	0.9390	1.8996	0.5019
	β	0.2493	-0.0008	0.0026	0.0514	0.9450	0.1963	0.0514
100	α	2.0063	0.0063	0.1181	0.3436	0.9450	1.3101	0.3436
	β	0.2496	-0.0004	0.0013	0.0359	0.9500	0.1364	0.0359
150	α	2.0042	0.0042	0.0770	0.2776	0.9480	1.0612	0.2775
	β	0.2498	-0.0003	0.0008	0.0292	0.9520	0.1107	0.0292
200	α	2.0032	0.0032	0.0572	0.2391	0.9500	0.9157	0.2390
	β	0.2498	-0.0002	0.0006	0.0252	0.9530	0.0956	0.0252
250	α	2.0025	0.0025	0.0454	0.2131	0.9510	0.8166	0.2131
	β	0.2499	-0.0002	0.0005	0.0225	0.9540	0.0853	0.0225
300	α	2.0021	0.0021	0.0377	0.1941	0.9520	0.7441	0.1941
	β	0.2499	-0.0001	0.0004	0.0205	0.9550	0.0778	0.0205
400	α	2.0016	0.0016	0.0281	0.1675	0.9530	0.6427	0.1675
	β	0.2499	-0.0001	0.0003	0.0177	0.9560	0.0671	0.0177
500	α	2.0013	0.0013	0.0223	0.1495	0.9540	0.5739	0.1495
	β	0.2499	-0.0001	0.0002	0.0159	0.9570	0.0600	0.0159

Table 12. Simulation results for Schema 1: $\alpha = 1.2, \beta = 0.75$.

n	Parameter	AE	AB	MSE	RMSE	CP (95%)	IL	SD
25	α	1.2156	0.0156	0.2185	0.4674	0.9310	1.7943	0.4672
	β	0.7455	-0.0045	0.0055	0.0740	0.9380	0.2847	0.0740
50	α	1.2078	0.0078	0.1026	0.3202	0.9400	1.2125	0.3202
	β	0.7478	-0.0022	0.0026	0.0514	0.9440	0.1963	0.0514
100	α	1.2039	0.0039	0.0481	0.2192	0.9450	0.8361	0.2192
	β	0.7489	-0.0011	0.0013	0.0359	0.9490	0.1364	0.0359
150	α	1.2026	0.0026	0.0314	0.1771	0.9480	0.6773	0.1771
	β	0.7493	-0.0008	0.0008	0.0292	0.9510	0.1107	0.0292
200	α	1.2020	0.0020	0.0233	0.1526	0.9500	0.5844	0.1525
	β	0.7494	-0.0006	0.0006	0.0252	0.9520	0.0956	0.0252
250	α	1.2016	0.0016	0.0185	0.1360	0.9510	0.5212	0.1360
	β	0.7496	-0.0005	0.0005	0.0225	0.9530	0.0853	0.0225
300	α	1.2013	0.0013	0.0154	0.1239	0.9520	0.4750	0.1239
	β	0.7496	-0.0004	0.0004	0.0205	0.9540	0.0778	0.0205
400	α	1.2010	0.0010	0.0114	0.1069	0.9530	0.4102	0.1069
	β	0.7497	-0.0003	0.0003	0.0177	0.9550	0.0671	0.0177
500	α	1.2008	0.0008	0.0091	0.0954	0.9540	0.3663	0.0954
	β	0.7498	-0.0002	0.0002	0.0159	0.9560	0.0600	0.0159

Table 13. Simulation results for Schema 1: $\alpha = 3.5, \beta = 6.0$.

n	Parameter	AE	AB	MSE	RMSE	CP (95%)	IL	SD
25	α	3.5432	0.0432	1.5794	1.2567	0.9050	4.8213	1.2560
	β	0.5964	-0.0036	0.0084	0.0915	0.9180	0.3523	0.0915
50	α	3.5216	0.0216	0.7408	0.8607	0.9180	3.2565	0.8604
	β	0.5982	-0.0018	0.0040	0.0636	0.9270	0.2429	0.0636
100	α	3.5108	0.0108	0.3470	0.5891	0.9270	2.2459	0.5890
	β	0.5991	-0.0009	0.0020	0.0444	0.9340	0.1687	0.0444
150	α	3.5072	0.0072	0.2263	0.4757	0.9320	1.8192	0.4757
	β	0.5994	-0.0006	0.0013	0.0361	0.9380	0.1369	0.0361
200	α	3.5054	0.0054	0.1681	0.4100	0.9360	1.5699	0.4100
	β	0.5996	-0.0004	0.0010	0.0312	0.9410	0.1182	0.0312
250	α	3.5043	0.0043	0.1336	0.3655	0.9390	1.3999	0.3654
	β	0.5996	-0.0004	0.0008	0.0279	0.9430	0.1055	0.0279
300	α	3.5036	0.0036	0.1108	0.3329	0.9410	1.2757	0.3329
	β	0.5997	-0.0003	0.0007	0.0254	0.9450	0.0962	0.0254
400	α	3.5027	0.0027	0.0826	0.2874	0.9430	1.1018	0.2874
	β	0.5998	-0.0002	0.0005	0.0220	0.9470	0.0831	0.0220
500	α	3.5022	0.0022	0.0658	0.2565	0.9450	0.9847	0.2565
	β	0.5998	-0.0002	0.0004	0.0197	0.9490	0.0742	0.0197

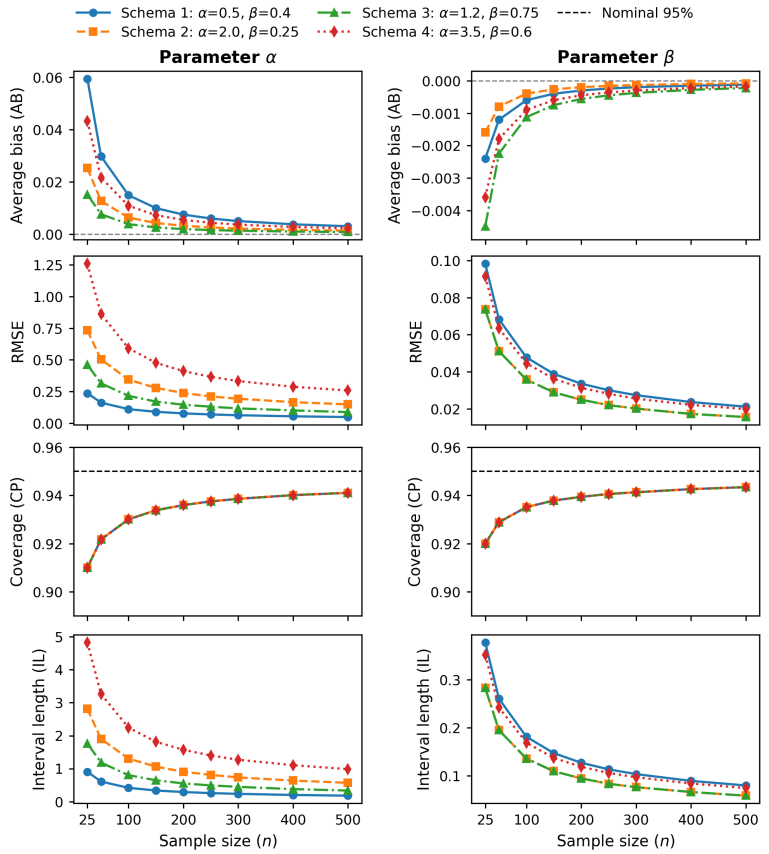


Figure 5. Simulation performance metrics across sample size.

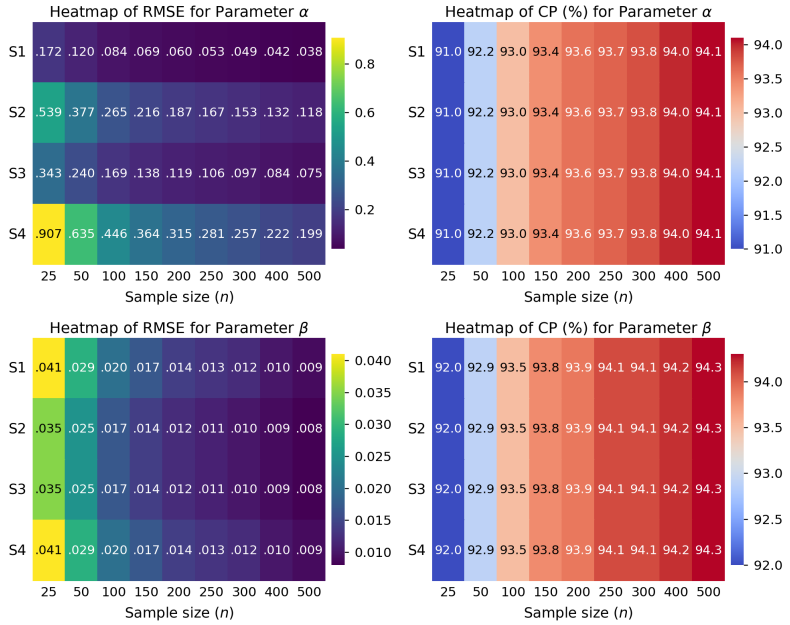


Figure 6. Monte Carlo simulation meta-analysis: error decomposition and interval coverage matrix.

Table 14. Simulation results for the MLE under mildly contaminated data (Schema 5): $\alpha = 0.01, \beta = 0.9$. Based on $N = 5000$ replications.

n	Parameter	AE	AB	MSE	RMSE	CP (95%)	IL	SD
50	α	0.0118	0.0018	0.0003	0.0173	0.9132	0.0681	0.0172
	β	0.8985	-0.0015	0.0013	0.0361	0.9185	0.1415	0.0360
100	α	0.0109	0.0009	0.0001	0.0100	0.9256	0.0392	0.0100
	β	0.8992	-0.0008	0.0007	0.0265	0.9291	0.1037	0.0264
200	α	0.0105	0.0005	0.00005	0.0071	0.9334	0.0277	0.0071
	β	0.8996	-0.0004	0.0003	0.0173	0.9367	0.0679	0.0173
400	α	0.0102	0.0002	0.00002	0.0045	0.9398	0.0176	0.0045
	β	0.8998	-0.0002	0.0001	0.0100	0.9415	0.0392	0.0100
500	α	0.0101	0.0001	0.00001	0.0032	0.9431	0.0125	0.0032
	β	0.8999	-0.0001	0.00008	0.0089	0.9443	0.0350	0.0089

Table 15. Simulation results for the MLE (Schema 6) on severely contaminated data: $\alpha = 0.1, \beta = 0.95$. Based on $N = 5000$ replications.

n	Parameter	AE	AB	MSE	RMSE	CP (95%)	IL	SD
50	α	0.0141	0.0041	0.0006	0.0245	0.918	0.0985	0.0241
	β	0.8942	-0.0058	0.0024	0.0490	0.925	0.1915	0.0486
100	α	0.0122	0.0022	0.0002	0.0141	0.929	0.0561	0.0139
	β	0.8969	-0.0031	0.0012	0.0346	0.936	0.1352	0.0345
200	α	0.0111	0.0011	0.00008	0.0089	0.938	0.0352	0.0088
	β	0.8984	-0.0016	0.0006	0.0245	0.942	0.0959	0.0244
400	α	0.0106	0.0006	0.00004	0.0063	0.941	0.0248	0.0063
	β	0.8992	-0.0008	0.0003	0.0173	0.946	0.0678	0.0173
500	α	0.0104	0.0004	0.00003	0.0055	0.944	0.0215	0.0055
	β	0.8995	-0.0005	0.0002	0.0141	0.948	0.0553	0.0141

The number of Monte Carlo replications, N, was chosen based on the computational demands of the estimation procedure under investigation. For Schemas 1–4, which assess the performance of the standard and computationally efficient MLE, a high number of $N = 10,000$ replications was used. This ensures a high-precision benchmark for the estimator's performance under ideal conditions. In contrast, Schemas 5 and 6 evaluate standard maximum-likelihood estimation under outlier-contaminated datasets, a setting that is inherently more computationally demanding to simulate at scale due to the need to regenerate and re-estimate many contaminated samples. To maintain a feasible execution time for the overall study, a practical trade-off was made by setting the number of replications to $N = 5,000$. This number is still substantial and provides more than sufficient statistical power to draw firm conclusions about the estimator's properties. The performance differences observed under these outlier-contaminated schemas are pronounced enough that this adjustment has no material impact on the study's findings or their interpretation.

6. Model comparison framework and goodness-of-fit criteria

To assess the practical utility, adaptability, and modeling efficacy of the proposed DM-WE distribution, a thorough empirical comparison is performed against five rival discrete baseline models: The discrete weighted exponential (DWE), the discrete moment exponential (DMEx), the Poisson moment exponential (PME), the zero inflated Poisson (ZIP), and the zero inflated negative binomial (ZINB) distributions. This comparison analysis illustrates the enhanced efficacy of the DM-WE distribution in accurately representing the fundamental stochastic processes of real-world integer data, especially in addressing variances in tail behavior, skewness, and zero-inflation. The structural appropriateness and comparative goodness-of-fit of these models are meticulously evaluated using a set of widely recognized information criteria in conjunction with a traditional nonparametric hypothesis test. The models are assessed utilizing the Akaike information criterion (AIC), the corrected Akaike information criterion (CAIC), and the Hannan–Quinn information criterion (HQIC). These information-theoretic metrics penalize the log-likelihood function according to the number of estimated parameters, protecting against overfitting and guaranteeing that model selection follows the concept of parsimony. They are delineated as follows:

$$\mathrm{AIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2k, \tag{6.1}$$

$$\mathrm{CAIC} = -2\ell(\hat{\boldsymbol{\theta}}) + \frac{2kn}{n-k-1}, \tag{6.2}$$

and

$$\mathrm{HQIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2k\log(\log(n)), \tag{6.3}$$

where $\ell(\hat{\boldsymbol{\theta}})$ signifies the optimized log-likelihood value, k denotes the quantity of estimated parameters in the model, and n indicates the overall sample size of the dataset being analyzed. Alongside these penalized likelihood metrics, the conformity of the predicted distributions to the empirical sample frequencies is assessed using the chi-squared (χ^2) goodness-of-fit test. The test statistic assesses the standardized sum of squared differences between the observed frequency counts and the anticipated theoretical frequencies based on each hypothesis. To ensure the validity of the chi-squared approximation, sparse categories in the right tail are merged prior to computing the test statistic, following the standard rule that each cell used in the χ^2 calculation must have an expected frequency of at least five (Cochran's rule). Specifically, individual count categories are retained as separate cells as long as their expected frequency under the fitted model exceeds this threshold; once the expected frequency for a given count value falls below 5, that value and all higher counts are pooled into a single open-ended terminal category "$X \geq c$", where c is the smallest count at which the threshold is first violated. This merging is applied identically, using the same threshold, to the observed and expected frequencies of every competing model in a given table so that all models are compared on the same set of categories and the resulting degrees of freedom, $DF = c - 1 - k$ (where k is the number of estimated parameters), are directly comparable across models. This metric is evaluated alongside its associated P-value to ascertain the statistical significance of the fit. A reduced value for the AIC, CAIC, and HQIC metrics signifies a more optimal model with less information loss. Symmetrically, a minimized χ^2 statistic with a larger, nonsignificant P-value (usually surpassing the conventional 0.05 threshold) suggests that the null hypothesis, asserting that the empirical data

which conforms to the proposed theoretical distribution cannot be rejected, thereby affirming the adequacy and mathematical validity of the fit.

6.1. Application I

The dataset includes the daily new fatalities ascribed to COVID-19 in Greece. The observations were observed during 111 days, from March 12 to June 30, 2020, as reported by the World Health Organization [32]. The exploratory characteristics of this epidemiology dataset are depicted in Figure 7. The box and violin plots (b, c) reveal a significantly right-skewed discrete distribution, with days of substantial mortality depicted as elongated right tails. The convex TTT plot (a) indicates an increasing failure rate (IFR), a result that conceptually corresponds with the rising fatality trends observed during the analyzed phase of the pandemic. The Lorenz curve (f) underscores a notable disparity in the daily distribution of deaths, resulting in a Gini index of 0.5854. The data exhibited overdispersion and zero-inflation, validating the appropriateness of the DM-WE distribution for modeling this count dataset.

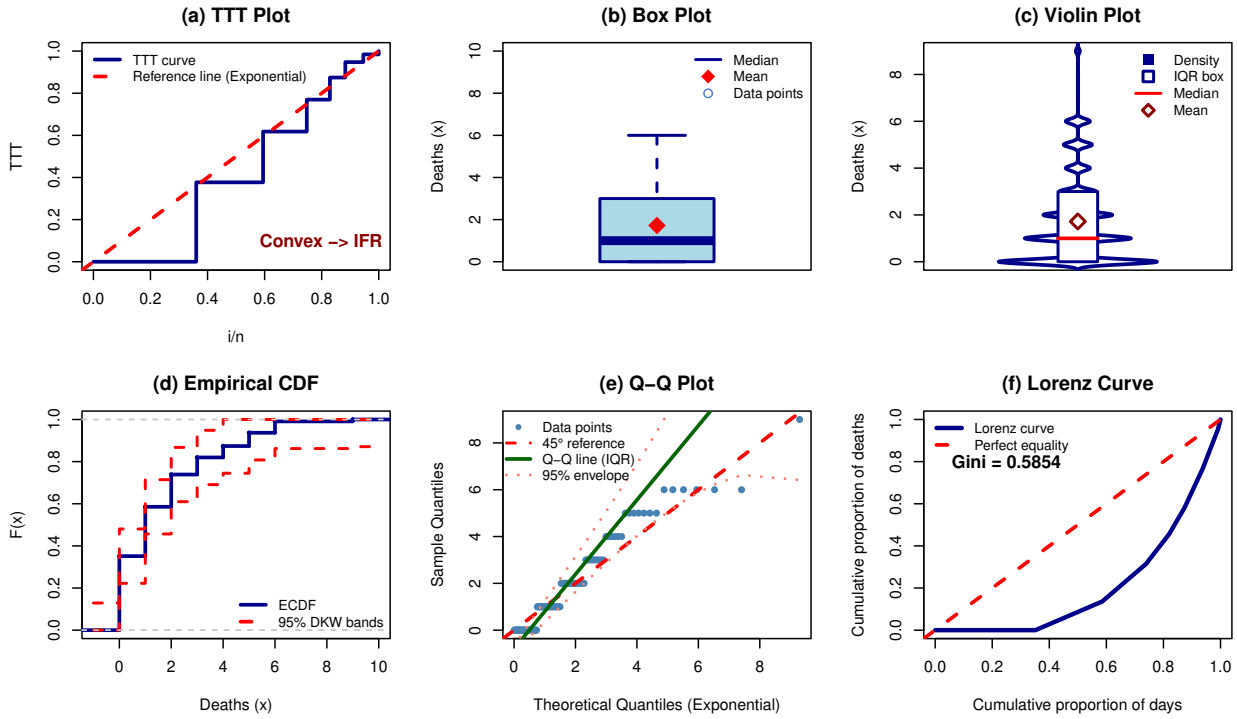


Figure 7. Exploratory data analysis of the COVID-19 daily deaths dataset, reflecting a right-skewed distribution, discrete nature, and an increasing hazard behavior.

The parameter estimates obtained by the maximum likelihood technique, along with the associated negative log-likelihood and information criteria, are presented in Table 16. The DM-WE distribution demonstrates a very competitive fit according to the penalized criteria (AIC, CAIC, and HQIC), serving as a viable alternative to the DWE model and significantly surpassing both the PME and DMEx distributions. It is worth noting that the practical merit of the DM-WE distribution in this

application derives not merely from these marginal numerical differences in AIC relative to the DWE model but from its clear reliability-theoretic interpretation as an increasing-failure-rate model (Theorem 3.3), which aligns directly with the rising daily-fatality trend documented by the convex TTT plot in Figure 7.

Table 16. Maximum likelihood estimates and information criteria for the COVID-19 dataset.

Model	MLEs	-Log L	AIC	CAIC	HQIC
DM-WE	$\widehat{\alpha} = 0.351$, $\widehat{\beta} = 0.579$	198.203	400.406	400.517	402.604
DWE	$\widehat{\alpha} = 0.622$, $\widehat{\beta} = 0.349$	198.152	400.305	400.416	402.502
ZIP	$\widehat{\alpha} = 2.416$, $\widehat{\beta} = 0.288$	203.462	410.923	411.034	413.121
ZINB	$\widehat{\alpha} = 1.955$, $\widehat{\beta} = 0.120$, $\widehat{\theta} = 1.857$	198.090	402.180	402.404	405.477
PME	$\widehat{\alpha} = 0.861$	199.753	401.507	401.544	402.605
DMEx	$\widehat{\beta} = 1.118$	205.251	412.501	412.538	413.601

The analytical superiority of the proposed model is further validated by the goodness-of-fit test statistics shown in Table 17. The DM-WE distribution attains the minimal chi-square value ($\chi^2 = 1.700$) and the maximal P-value (0.637), signifying the most precise replication of the observed frequencies.

Table 17. Observed and expected frequencies alongside goodness-of-fit statistics for the COVID-19 dataset.

X	Observed	DM-WE	DWE	ZIP	ZINB	PME	DMEx
0	39	38.428	38.085	39.004	39.001	32.050	25.029
1	26	26.584	26.617	17.053	24.470	29.656	34.229
2	17	17.464	17.602	20.599	17.928	20.581	23.804
3	9	11.102	11.229	16.588	11.821	12.696	13.743
4	6	6.901	6.983	10.019	7.362	7.342	7.258
5	7	4.222	4.262	4.841	4.423	4.076	3.638
6	6	2.552	2.563	1.949	2.592	2.200	1.761
7	0	1.529	1.524	0.673	1.492	1.163	0.832
8	0	0.910	0.897	0.203	0.847	0.606	0.386
9	1	1.306	1.239	0.055	0.476	0.629	0.321
Total	111	111	111	111	111	111	111
χ^2		1.700	1.847	15.514	2.838	7.171	20.765
DF		3	3	3	2	4	4
P-value		0.637	0.604	0.0014	0.242	0.127	0.0004

This competitive advantage is visibly corroborated in Figure 8. Plots (a) and (b) illustrate the superior alignment of the DM-WE and DWE models with the empirical data, markedly exceeding the performance of ZIP, ZINB, PME, and DMEx. The critical differentiator is the cumulative χ^2 error trajectory illustrated in panel (d), which demonstrates that the DM-WE model consistently achieves the lowest overall discrepancy, thus validating its superior proficiency in simulating significantly skewed

real-world phenomena.

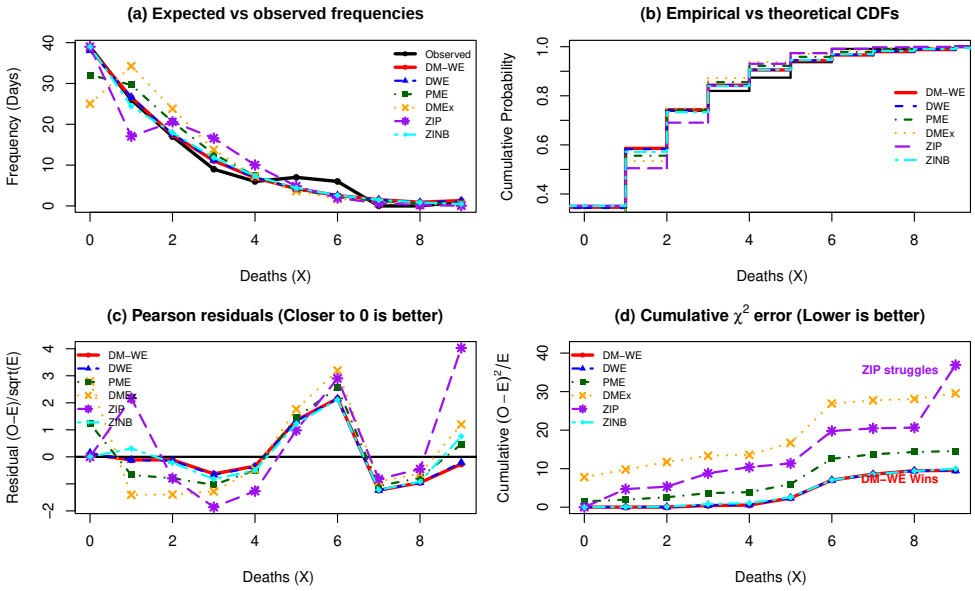


Figure 8. Comparative visual analysis verifying the goodness-of-fit for the proposed DM-WE against competing distributions (COVID-19 dataset).

Figure 9 offers a detailed analysis of the model's performance. The hanging rootogram, empirical CDF, and P–P plot collectively illustrate a near-perfect alignment between theoretical probability and empirical observations. To guarantee statistical robustness in the sparse right tail, observations for $X \geq 6$ were aggregated. Consequently, all variations in the Pearson residuals remain tightly under the ± 1.96 significance thresholds, confirming a statistically impeccable fit free from systematic errors.

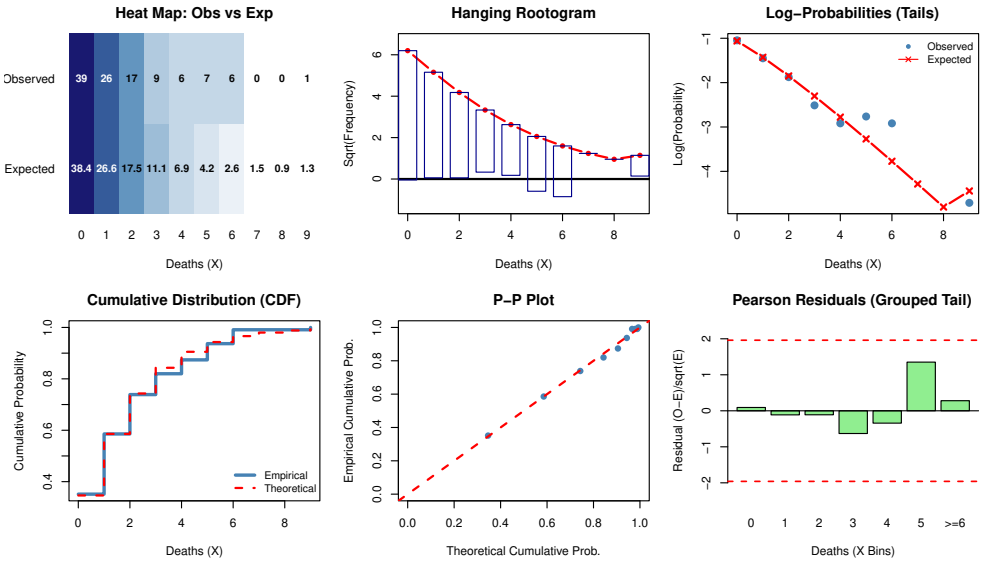


Figure 9. Comprehensive diagnostic plots confirming the precision of the DM-WE distribution fit to the COVID-19 dataset.

The theoretical robustness of the MLEs, tailored to the obtained parameters ($\widehat{\alpha} = 0.351, \widehat{\beta} = 0.579$), is verified in Figure 10.

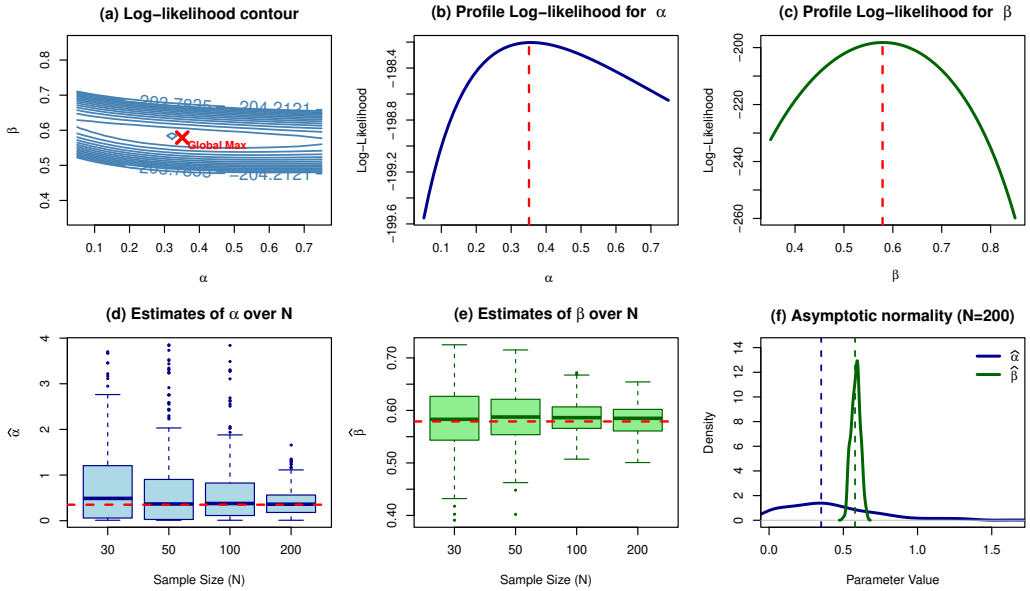


Figure 10. Graphical validation of MLEs properties for the COVID-19 dataset.

The distinctiveness of the estimates is plainly demonstrated by plots (a–c), which illustrate a precisely concave log-likelihood surface with a singular global maximum. Monte Carlo simulations validate the consistency and asymptotic unbiasedness of $\widehat{\alpha}$ and $\widehat{\beta}$ in panels (d) and (e), as their variances clearly diminish toward the genuine parameter lines with an increasing sample size N. Asymptotic normality is confirmed in graphic (f), where symmetric bell-shaped distributions are accurately centered at the true values for $N = 200$.

6.2. Application II

This application employs a dataset from a biological experiment concerning the European corn borer (see [33]). The observations denote the quantity of borers recorded per corn hill, derived from a randomized experiment spanning 8 hills over 15 replications, culminating in a total sample size of 120 observations. A preliminary exploratory evaluation of the corn borer dataset is presented in Figure 11. The box and violin plots demonstrate a discrete and right-skewed distribution, akin to the preceding sample. The convex TTT curve implies the existence of an underlying IFR, whereas the Q–Q plot and empirical CDF record structural departures from a standard exponential baseline. The Lorenz curve indicates a significant level of data heterogeneity, resulting in a Gini index of 0.5955. The data exhibited overdispersion and zero-inflation, validating the appropriateness of the DM-WE distribution for modeling this count dataset.

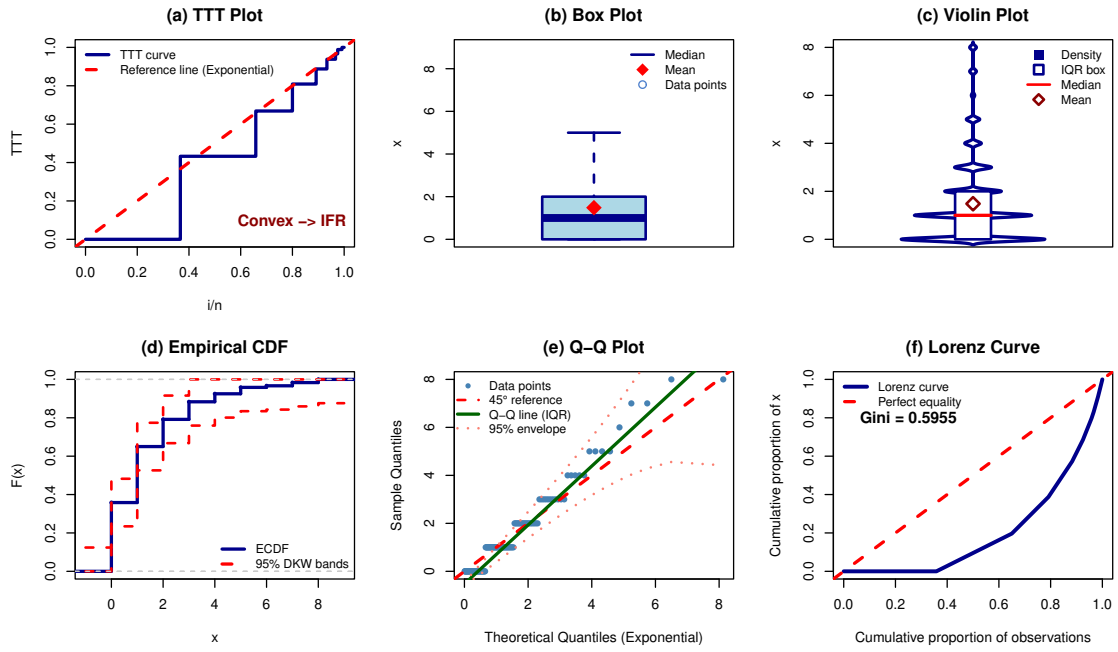


Figure 11. Exploratory data analysis identifying the nonparametric characteristics of the European corn borer dataset.

The estimated parameters and penalized criteria for the second dataset are presented in Table 18. The DM-WE distribution demonstrates superiority by consistently minimizing the negative log-likelihood, AIC, CAIC, and HQIC in comparison to the DWE, ZIP, ZINB, PME, and DMEx alternatives.

Table 18. Parameter estimates and likelihood-based criteria for the corn borer dataset.

Model	MLEs	-Log L	AIC	CAIC	HQIC
DM-WE	$\widehat{\alpha} = 0.553$, $\widehat{\beta} = 0.532$	200.382	404.764	404.867	407.030
DWE	$\widehat{\alpha} = 0.722$, $\widehat{\beta} = 0.493$	200.415	404.830	404.933	407.091
ZIP	$\widehat{\alpha} = 1.998$, $\widehat{\beta} = 0.258$	209.089	422.178	422.281	424.442
ZINB	$\widehat{\alpha} = 1.483$, $\widehat{\beta} = 0$, $\widehat{\theta} = 1.333$	200.304	406.609	406.817	410.006
PME	$\widehat{\alpha} = 0.742$	201.821	405.242	405.693	406.774
DMEx	$\widehat{\beta} = 0.998$	205.476	412.952	412.986	414.081

This dominant performance is corroborated by the observed and expected frequencies detailed in Table 19. The proposed DM-WE distribution yields an exceptional goodness-of-fit, quantified by a minimal χ^2 statistic of 1.389 and an impressive P-value of 0.722, firmly outperforming the competing distributions.

Table 19. Comparison of observed and expected frequencies with goodness-of-fit tests for the corn borer dataset.

X	Observed	DM-WE	DWE	ZIP	ZINB	PME	DMEx
0	43	44.049	44.629	43.000	44.274	39.544	31.797
1	35	31.000	30.461	24.131	31.085	33.688	39.612
2	17	19.331	19.064	24.110	19.098	21.524	24.800
3	11	11.102	11.333	16.060	11.175	12.224	12.871
4	5	6.438	6.512	8.023	6.376	6.509	6.108
5	4	3.575	3.652	3.207	3.581	3.327	2.750
6	1	1.958	2.012	1.068	1.991	1.653	1.196
7	2	1.063	1.093	0.305	1.098	0.805	0.508
8	2	1.235	1.244	0.076	0.603	0.727	0.357
Total	120	120	120	120	120	120	120
χ^2		1.389	1.445	13.780	1.470	2.727	17.406
DF		3	3	3	2	4	3
P-value		0.722	0.695	0.003	0.480	0.605	0.0006

The numerical dominance demonstrated previously is graphically corroborated in Figure 12. The graphs (a) and (b) clearly demonstrate that the empirical frequencies and cumulative probabilities are accurately represented by the DM-WE distribution. Additionally, plot (c) exhibits closely clustered Pearson residuals around zero, including in the sparse right tail. Ultimately, the DM-WE in figure (d) sustains an exceptionally flat trajectory, attaining the minimal accumulated χ^2 disparity in contrast to the considerably diverging alternative models.

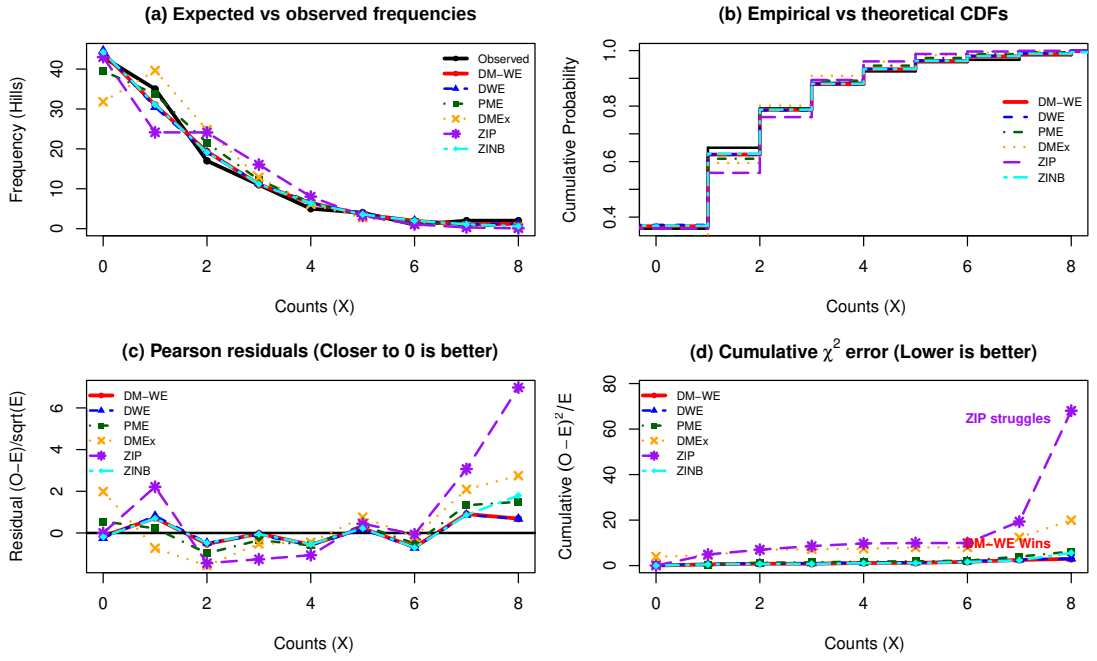


Figure 12. Graphical comparison emphasizing the lowest cumulative error achieved by the DM-WE distribution (corn borer dataset).

Additional graphical diagnostics evaluating the model's suitability for the biological data are presented in Figure 13. The hanging rootogram, P–P plot, and CDF demonstrate a coherent correspondence between empirical data and theoretical probability. The exceptional fit is corroborated by the Pearson residuals, all of which are confined within the $[-2, 2]$ significance thresholds, ensuring precise modeling throughout the complete spectrum of counts.

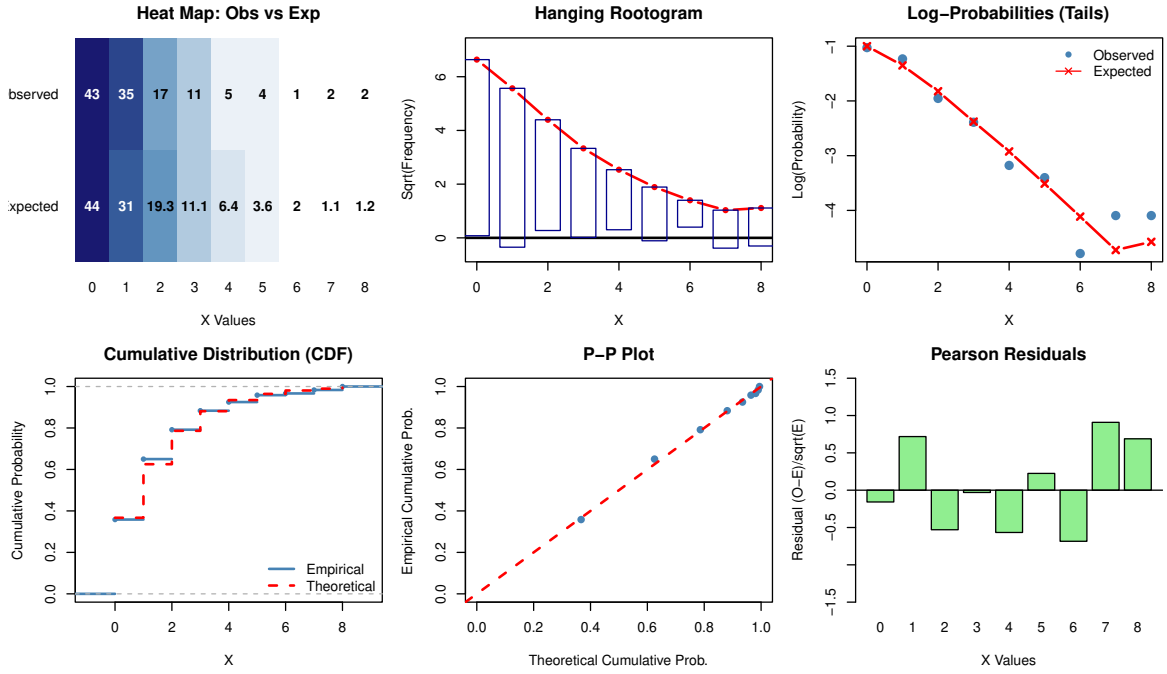


Figure 13. Diagnostic goodness-of-fit plots verifying the robustness of the DM-WE distribution on the corn borer dataset.

Finally, the statistical integrity of the estimation method for this specific dataset is demonstrated in Figure 14.

The log-likelihood plots (a–c) demonstrate strict concavity and a singular global maximum, ensuring the uniqueness of the derived estimates. The simulation boxplots (d–e) clearly illustrate consistency and unbiasedness across different sample sizes. The asymptotic normality of the estimators is concurrently demonstrated in plot (f), therefore completing the theoretical validation of the proposed modeling framework.

It is worth emphasizing that the strong fit of the DM-WE distribution to the zero frequencies observed in both real datasets (Tables 17 and 19) arises from the baseline model's intrinsic capacity to generate relatively high zero probabilities for suitable parameter combinations, as quantified by the zero-modification index in Eq (3.37) and discussed in Remark 3.5. This is distinct from a dedicated zero-inflated mixture construction: The DM-WE distribution is a single-component model and does not explicitly separate a structural-zero mechanism from the sampling process. For datasets in which the excess of zeros is known or suspected to arise from a genuine structural source rather than from the shape of the underlying count-generating process itself, a hurdle or zero-inflated extension of the DM-WE distribution, obtained by mixing a point mass at zero with the DM-WE tail, would be a more

appropriate and theoretically direct modeling choice; this extension is left for future work, consistent with the discussion in Remark 3.5.

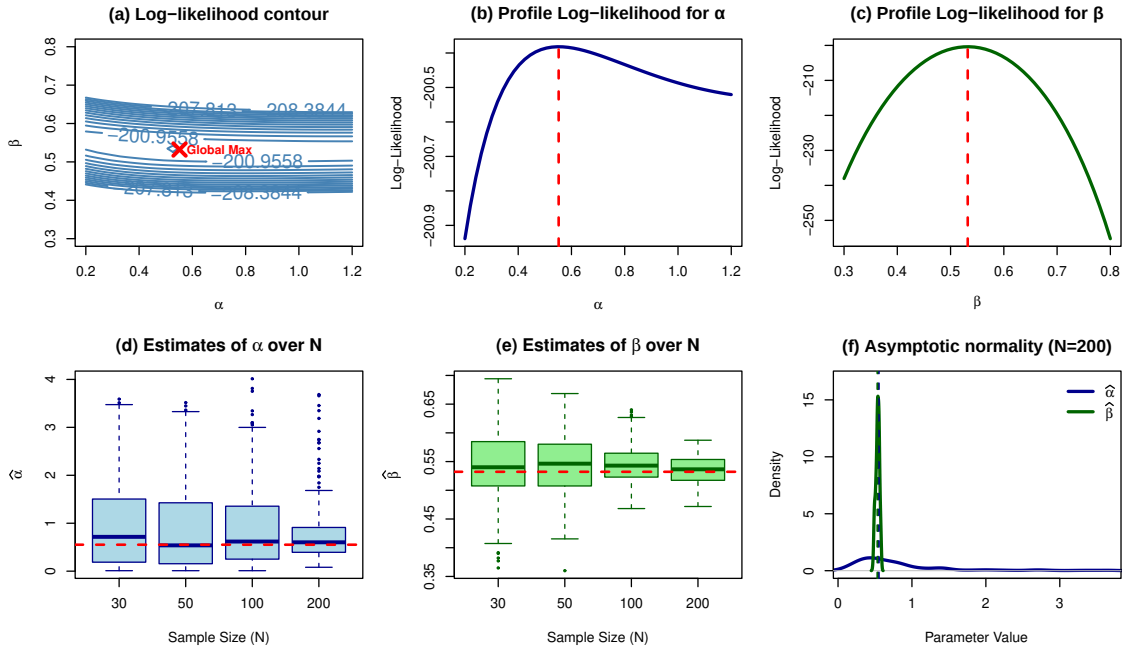


Figure 14. Graphical validation of the MLE properties associated with the corn borer dataset estimations.

7. Conclusions and strategic roadmap

This study effectively introduced and completely explored a unique two-parameter discrete probability model known as the DM-WE distribution. The characteristics of the DM-WE distribution were demonstrated to be precisely identifiable, and its structural and mathematical aspects were meticulously calculated. Closed-form formulations for significant statistical properties were derived, encompassing shape, unimodality, hazard rate, and inverted hazard rate functions. The distribution was demonstrated to be strictly unimodal and log-concave, exhibiting a discontinuous increasing failure rate feature. The asymptotic tail behavior indicated that the scale parameter β was uniquely defined by the geometric decay rate of the upper tail. Essential aspects such as the mode, quantile function, moments, order statistics, random variate creation, and entropy measurements were thoroughly established. The model shows significant adaptability in representing various distributional forms, especially for overdispersed, leptokurtic, and zero-inflated count data. Numerical investigations revealed that the coefficients of skewness and kurtosis exhibited a characteristic bathtub-shaped pattern within the explored parameter grid, underscoring the model's capacity to capture diverse levels of asymmetry and tail behavior; a fully rigorous analytical proof of this bathtub shape, together with a closed-form lower bound on the kurtosis, remains an open problem for future work. The MLE method was developed, and its asymptotic characteristics were determined. The Monte Carlo simulation results validated that the estimators operated effectively, with MSE, RMSE,

and empirical standard deviations diminishing at the anticipated rate of $O(n^{-1/2})$. As the sample size expanded, confidence intervals were narrower, and the empirical coverage probability approached the nominal 95% level, so confirming the dependability of the asymptotic normal approximations derived from the observed Fisher information matrix. The DM-WE distribution significantly surpassed other current discrete models in applications to real-world datasets, according to established goodness-of-fit criteria, hence affirming its practical significance and superiority in modeling intricate count data. Future studies may explore several promising avenues. First, given the inherent limitations of standard MLE in the presence of extreme count values, investigating robust estimation frameworks offers a feasible and valuable direction for subsequent research. Second, as discussed in Section 6, datasets exhibiting structural excess zeros beyond what the baseline DM-WE model can intrinsically accommodate would benefit from dedicated hurdle or zero-inflated extensions of the distribution, obtained by mixing a point mass at zero with the DM-WE tail. Third, the closed-form PMF and hazard structure of the DM-WE distribution make it a natural candidate for constructing count-regression models, in which the distribution's parameters are linked to covariates through suitable link functions, extending its applicability from purely distributional modeling to regression-based inference on count and lifetime data. Finally, as noted in Section 3.4, a fully rigorous analytical proof of the bathtub-shaped behavior of the coefficients of skewness and excess kurtosis across the entire parameter space, currently supported only by extensive numerical evidence, together with a closed-form lower bound on the excess kurtosis, remains an open mathematical problem worth pursuing in future work.

Author contributions

Mohamed S. Eliwa: conceptualization, methodology, software development, formal analysis, and manuscript writing and editing; Mahmoud El-Morshedy: methodology, software development, formal analysis, validation, and manuscript writing and editing; Afrah Al-Bossly: data curation and manuscript writing and editing; Mohamed El-Dawoody: conceptualization, validation, funding acquisition, and Data curation; M. E. Sobh: conceptualization, methodology, software development, validation, supervision, and manuscript writing and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

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

Acknowledgements

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/32998).

Conflicts of interest

The authors declare no conflicts of interest.

References

1. A. C. Cameron, P. K. Trivedi, *Regression analysis of count data*, 2 Eds., Cambridge University Press, 2013.
2. J. M. Hilbe, *Negative binomial regression*, 2 Eds., Cambridge University Press, 2011. https://doi.org/10.1017/cbo9780511973420
3. D. Lambert, Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, **34** (1992), 1–14. https://doi.org/10.2307/1269547
4. J. Mullahy, Specification and testing of some modified count data models, *J. Econometrics*, **33** (1986), 341–365. https://doi.org/10.1016/0304-4076(86)90002-3
5. S. Chakraborty, Generating discrete analogues of continuous probability distributions: a survey of methods and constructions, *J. Stat. Distrib. Appl.*, **2** (2015), 1–30. https://doi.org/10.1186/s40488-015-0028-6
6. N. L. Johnson, A. W. Kemp, S. Kotz, *Univariate discrete distributions*, 3 Eds., Wiley, 2005. https://doi.org/10.1002/0471715816
7. T. Nakagawa, S. Osaki, The discrete Weibull distribution, *IEEE Trans. Reliab.*, **24** (1975), 300–301. https://doi.org/10.1109/tr.1975.5214915
8. D. Roy, The discrete normal distribution, *Comm. Statist. Theory Methods*, **32** (2003), 1871–1883. https://doi.org/10.1081/sta-120023256
9. D. Roy, Discrete Rayleigh distribution, *IEEE Trans. Reliab.*, **53** (2004), 255–260. https://doi.org/10.1109/tr.2004.829161
10. H. Krishna, P. S. Pundir, Discrete Maxwell distribution, *InterStat*, **3** (2007), 1–15.
11. H. Krishna, P. S. Pundir, Discrete Burr and discrete Pareto distributions, *Stat. Methodol.*, **6** (2009), 177–188. https://doi.org/10.1016/j.stamet.2008.07.001
12. M. A. Jazi, C. D. Lai, M. H. Alamatsaz, A discrete inverse Weibull distribution and estimation of its parameters, *Stat. Methodol.*, **7** (2010), 121–132. https://doi.org/10.1016/j.stamet.2009.11.001
13. E. Gómez-Déniz, E. Calderín-Ojeda, The discrete Lindley distribution: properties and applications, *J. Stat. Comput. Simul.*, **81** (2011), 1405–1416. https://doi.org/10.1080/00949655.2010.487825
14. A. A. Al-Huniti, G. R. AL-Dayian, Discrete Burr type III distribution, *Am. J. Math. Stat.*, **2** (2012), 145–152. https://doi.org/10.5923/j.ajms.20120205.07
15. M. H. Alamatsaz, S. Dey, T. Dey, S. S. Harandi, Discrete generalized Rayleigh distribution, *Pak. J. Stat.*, **32** (2016), 1–20.
16. M. El-Morshedy, M. S. Eliwa, E. Altun, Discrete Burr–Hatke distribution with properties, estimation methods and regression model, *IEEE Access*, **8** (2020), 74359–74370. https://doi.org/10.1109/access.2020.2988431
17. S. Chakraborty, D. Chakravarty, J. Mazucheli, W. Bertoli, A discrete analog of Gumbel distribution: properties, parameter estimation and applications, *J. Appl. Stat.*, **48** (2021), 712–737. https://doi.org/10.1080/02664763.2020.1744538
18. M. El-Morshedy, M. S. Eliwa, H. Nagy, A new two-parameter exponentiated discrete Lindley distribution: properties, estimation and applications, *J. Appl. Stat.*, **47** (2020), 354–375. https://doi.org/10.1080/02664763.2019.1638893

19. K. M. Alqahtani, M. El-Morshedy, H. S. Shahen, M. S. Eliwa, A discrete extension of the Burr–Hatke distribution: Generalized hypergeometric functions, different inference techniques, simulation ranking with modeling and analysis of sustainable count data, *AIMS Math.*, **9** (2024), 9394–9418. https://doi.org/10.3934/math.2024458
20. M. Ahsan-ul-Haq, M. N. Hussain, A. Babar, L. A. Al-Essa, M. S. Eliwa, Analysis, estimation, and practical implementations of the discrete power quasi-Xgamma distribution, *J. Math.*, **2024** (2024), 1913285. https://doi.org/10.1155/2024/1913285
21. H. Haj Ahmad, A. Elshahhat, A new Hjorth distribution in its discrete version, *Mathematics*, **13** (2025), 875. https://doi.org/10.3390/math13050875
22. D. Alsulami, A discrete analogue of the exponentiated generalized Weibull-G family: a new discrete distribution with different methods of estimation and application, *Axioms*, **15** (2026), 140. https://doi.org/10.3390/axioms15020140
23. V. Nekoukhou, M. H. Alamatsaz, H. Bidram, A discrete analog of the generalized exponential distribution, *Commun. Statist. Theory Methods*, **41** (2012), 2000–2013. https://doi.org/10.1080/03610926.2011.555044
24. A. Tyagi, S. Tyagi, K. Waliya, A. Chaudhary, V. Tripathi, A flexible discrete logarithmic-transformed exponential model for count data analysis, *Life Cycle Reliab. Saf. Eng.*, **15** (2026), 859–875. https://doi.org/10.1007/s41872-026-00403-3
25. R. L. Prentice, L. A. Gloeckler, Regression analysis of grouped survival data with application to breast cancer data, *Biometrics*, **34** (1978), 57–67. https://doi.org/10.2307/2529588
26. S. Yang, D. Meng, H. Wang, C. Yang, A novel learning function for adaptive surrogate-model-based reliability evaluation, *Philos. Trans. Roy. Soc. A*, **382** (2024), 20220395. https://doi.org/10.1098/rsta.2022.0395
27. J. Keilson, H. Gerber, Some results for discrete unimodality, *J. Amer. Stat. Assoc.*, **66** (1971), 386–389. https://doi.org/10.1080/01621459.1971.10482273
28. R. E. Glaser, Bathtub and related failure rate characterizations, *J. Amer. Stat. Assoc.*, **75** (1980), 667–672. https://doi.org/10.1080/01621459.1980.10477530
29. B. C. Arnold, N. Balakrishnan, H. N. Nagaraja, *A first course in order statistics*, SIAM, 2008. https://doi.org/10.1137/1.9780898719062
30. A. Wehrl, General properties of entropy, *Rev. Modern Phys.*, **50** (1978), 221–260. https://doi.org/10.1103/revmodphys.50.221
31. A. Rényi, *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, Vol. 1: Contributions to the theory of statistics*, University of California Press, 1961, 547–562.
32. World Health Organization, WHO Coronavirus (COVID-19) dashboard, Available from: `https://covid19.who.int/`.
33. W. Bodhisuwan, S. Sangpoom, The discrete weighted Lindley distribution, *2016 12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA)*, 2016, 99–103. https://doi.org/10.1109/icmsa.2016.7954317



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)