AIMS Mathematics

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

AIMS Mathematics, 11(9): 29488–29528.
DOI: 10.3934/math.20261171
Received: 22 June 2026
Revised: 22 July 2026
Accepted: 27 July 2026
Published: 14 September 2026

Research article

The unit-Weibull geometric distribution: Mathematical properties, identifiability, and asymptotic analysis with applications to data modeling

Magrisha Namsaw[1], **Bhanita Das**[1], **Mohamed S. Eliwa**[2,3] **and Mohamed F. Abouelenein**[4,*]

[1] Department of Statistics, North Eastern Hill University, Shillong, Meghalaya, India
[2] Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia
[3] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
[4] Department of Insurance and Risk Management, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Riyadh, Saudi Arabia

* **Correspondence:** Email: MABOUELENEIN@imamu.edu.sa.

Abstract: This study introduces the unit-Weibull geometric (UWG) distribution, a novel bounded compound probability distribution defined on the unit interval $(0, 1)$. Constructed by compounding the unit-Weibull and geometric distributions, the UWG framework is physically motivated by reliability theory and systems engineering, specifically modeling series systems where the minimum component lifetime dictates overall system failure. The proposed model offers exceptional flexibility for analyzing bounded data, including proportions, percentages, and fractions. Notably, the distribution can accommodate both symmetric and asymmetric (positively or negatively skewed) data under varying degrees of kurtosis (mesokurtic, leptokurtic, and platykurtic behaviors). A comprehensive mathematical and statistical characterization of the UWG distribution is provided. This includes an identifiability analysis based on the asymptotic matching of tail exponents, derivation of the moment-generating structure, an examination of modality, and a formal log-convexity and log-concavity analysis to assess its shape flexibility and boundary behavior. Furthermore, key reliability and information-theoretic measures are derived, including quantile-based descriptive statistics, incomplete and conditional moments, mean residual life, measures of inequality, uncertainty metrics, and the distributions of order statistics. Crucially, the hazard rate function of the UWG distribution can model diverse risk behaviors over time, exhibiting highly versatile shapes, including monotonic (increasing or decreasing), unimodal, and bathtub-shaped profiles. Model parameters are estimated via the method of maximum likelihood. The finite-sample properties, consistency, and asymptotic efficiency of the resulting estimators are validated through an extensive Monte Carlo simulation study. Finally, the empirical adequacy and superiority of the UWG distribution are demonstrated using two real-world Datasets, where it consistently outperforms established baseline and competing compound models for

bounded data.

Keywords: statistical model; identifiability; series system reliability; risk curves; simulation; decision-making
Mathematics Subject Classification: 60E05, 62E15, 62N05, 62P30

1. Introduction

In recent years, statistical literature has witnessed a burgeoning interest in the development of novel univariate distributions specifically bounded on the unit interval $(0, 1)$. These models are indispensable for accurately capturing the stochastic behavior of variables constrained to a finite range, such as proportions, rates, indices, and percentages, which are routinely encountered in fields ranging from economics to reliability engineering. While the classical beta and Kumaraswamy distributions have historically served as the benchmarks for unit-response modelling, they often exhibit strict structural limitations, particularly when fitting data characterized by highly skewed profiles, heavy tails, or complex, non-monotonic hazard rate shapes. Consequently, contemporary researchers frequently resort to transforming standard continuous distributions onto the unit interval via appropriate mapping functions. A prominent and mathematically elegant methodology for expanding the flexibility of baseline distributions is the compounding framework. Pioneered by Marshall and Olkin [1], this approach blends a continuous lifetime distribution with a truncated discrete distribution. From a physical and structural reliability perspective, compounding is inherently motivated by the operational architecture of complex systems. Consider a system comprising of N homogeneous components organized in a specific structural topology (e.g., a series or parallel configuration), where the total failure time of the system depends directly on this arrangement. If N is treated as a latent, discrete random variable representing the fluctuating number of active components or defects, and the individual failure times X_i $(i = 1, 2, \ldots, N)$ are independent and identically distributed (i.i.d.) continuous random variables, the resulting marginal distribution of the system's lifetime is a compounded distribution.

Mathematically, if the system operates under a series configuration, the lifetime is governed by the minimum order statistic, $Y = \min(X_1, \ldots, X_N)$. To provide a simple conceptual understanding of the series-system framework, consider an electronic device consisting of two independent components connected in series, such as a sensor and a control unit. In a series configuration, the entire system functions only if both components operate successfully. Consequently, the system fails as soon as either component fails, making the system's lifetime equal to the minimum of the components' lifetimes. This example is presented solely to illustrate the underlying reliability concept that motivates the proposed distribution and does not represent a specific application considered in this study. Conversely, a parallel configuration is governed by the maximum order statistic, $Y = \max(X_1, \ldots, X_N)$. By evaluating the conditional cumulative distribution function (CDF) given N and compounding it with the probability mass function (PMF) of N, one derives a new marginal CDF. This structural mechanism introduces additional shape parameters that drastically broaden the tail behavior, skewness, and kurtosis of the baseline model. Crucially, it yields a highly flexible hazard rate function (HRF) capable of taking increasing, decreasing, bathtub, or unimodal (upside-down bathtub) shapes, which are vital for survival analysis. The literature features an extensive array of compounded distributions utilizing unbounded baselines. Notable examples include the exponential-geometric (EG) distribution [2], which exhibits a decreasing hazard rate, alongside the exponential-Poisson (EP) [3] and exponential-logarithmic (EL) [4] models. This paradigm has been broadly generalized using the Weibull and related baselines, yielding the Weibull-geometric [5], Weibull-Poisson [6], exponential-Weibull [7], modified Weibull-geometric [8], additive Weibull-geometric [9], exponentiated transmuted Weibull-geometric [10], exponentiated inverse Weibull-geometric [11], Weibull-Lindley [12], and inverse Weibull-geometric [13] distributions. More recent power-series extensions include the Ishita-power series [14], inverse Lindley-power series [15], Pareto-Poisson [16], Zhang-power series [17], and the Alpha power modified Weibull-geometric [18] distributions. Despite the vast volume of research dedicated to compounding unbounded baseline models, a significant gap exists in the statistical literature regarding the compounding of distributions strictly restricted to the unit interval. While bounded data are pervasive, the direct application of unbounded compound models to $(0, 1)$ data requires ad-hoc transformations that often destroy the physical interpretation of the system's underlying series-parallel reliability structure. To date, only a sparse selection of unit-bound compound families have been proposed, such as the unit-Burr XII-power series [19], unit exponentiated half Logistic-power series [20], Topp Leone-Gompertz-G power series [21], unit Gompertz-power series [22], odd extension Weibull G-class [23], new Kumaraswamy-G class [24], log-XLindley [25], log-Bilal [26], and a new Kumaraswamy generalized-G class [27] distributions.

To bridge this literature gap, this paper introduces a novel bounded compound distribution by pairing the unit-Weibull distribution originally developed by Mazucheli et al. [28] with a truncated geometric distribution under a series system configuration. The CDF of the baseline unit-Weibull distribution is given by:

$$F(x; \alpha, \beta) = e^{-\alpha(-\log x)^{\beta}}, \quad 0 < x < 1, \quad \alpha, \beta > 0, \tag{1.1}$$

where α and β are the scale and shape parameters, respectively. The number of components N is assumed to follow a geometric distribution with the PMF

$$P(N = n) = (1 - p)p^{n-1}, \quad n = 1, 2, 3, \ldots, \quad 0 < p < 1. \tag{1.2}$$

By synthesizing these two components, we construct the UWG distribution. The primary scientific motivations for proposing the unit-Weibull geometric (UWG) model are twofold:

(i) It provides a mathematically rigorous, bounded distribution that accommodates data in the $(0, 1)$ domain without losing the physical interpretability of a stochastic series system (e.g., modeling the time to the first failure or the minimum proportion of structural integrity across a random number of components).

(ii) It resolves the inflexibility limitations of the classical unit-Weibull baseline by introducing the compounding parameter p, which drastically alters the kurtosis and allows the model to capture highly non-monotonic, bathtub, and upside-down bathtub failure risk behaviors that competing models fail to fit adequately.

The remainder of this paper is structured systematically to provide a comprehensive analysis of the proposed model. Section 2 formally introduces the structural framework of the UWG distribution, establishing its probability density function (PDF) and CDF alongside visual representations of its versatile shapes. Section 3 investigates the structural identifiability of the UWG parameters, utilizing the asymptotic matching of tail exponents to prove model uniqueness. Section 4 conducts a rigorous mathematical analysis of the UWG density, focusing on its log-convexity, log-concavity, modality, and limiting behaviors at the boundaries of the unit interval. Section 5 derives the fundamental mathematical and structural properties of the distribution, including raw and conditional moments, quantile functions, mean residual life, and information-theoretic uncertainty measures. In Section 6, the point estimation framework for the unknown model parameters is formulated using the method of maximum likelihood estimation. Section 7 presents a comprehensive Monte Carlo simulation study designed to empirically evaluate the finite-sample performance, consistency, and asymptotic efficiency of the resulting estimators. Section 8 showcases the practical utility and superior fitting capability of the UWG model by applying it to two bounded, real-world Datasets and benchmarking its performance against established baseline distributions. Finally, Section 9 offers concluding remarks and outlines potential avenues for future research.

2. The UWG distribution: Theory and visualization

In this section, we formally introduce the UWG distribution, establishing its foundational stochastic mechanism through a reliability-based compounding framework. Let us consider a engineering or biological system comprising N homogeneous components organized in a strict series topology, where N is a latent, discrete random variable representing the fluctuating number of components or latent flaws. We assume that the individual component lifetimes, denoted X_i ($i = 1, 2, \ldots, N$), are i.i.d. continuous random variables operating conditioned on N. Suppose that the stochastic behavior of the latent component count N follows a truncated discrete geometric distribution with the PMF previously stated in Eq (1.2)

$$P(N = n) = (1 - p)p^{n-1}, \quad n = 1, 2, 3, \ldots, \quad 0 < p < 1. \tag{2.1}$$

Furthermore, assume that the individual component lifetimes X_i are governed by the baseline unit-Weibull distribution [28] with the CDF given in Eq (1.1)

$$F(x; \alpha, \beta) = e^{-\alpha(-\log x)^{\beta}}, \quad 0 < x < 1, \quad \alpha, \beta > 0. \tag{2.2}$$

Because the system is arranged in a series configuration, the overall system failure time, denoted by the random variable X, is governed by the minimum order statistic of the sample, such that:

$$X = \min(X_1, X_2, \ldots, X_N). \tag{2.3}$$

Conditioned on a fixed sample size $N = n$, the conditional CDF of X is mathematically formulated via survival concepts as follows

$$F(x \mid N = n; \alpha, \beta) = 1 - [1 - F(x)]^n = 1 - \left[1 - e^{-\alpha(-\log x)^{\beta}}\right]^n. \tag{2.4}$$

By applying the law of total probability and compounding the conditional distribution with the discrete distribution of N, the unconditional, marginal CDF of the proposed UWG model is derived as follows:

$$F_{\mathrm{UWG}}(x; \alpha, \beta, p) = \sum_{n=1}^{\infty} P(N = n) F(x \mid N = n; \alpha, \beta). \tag{2.5}$$

Evaluating this infinite series yields the following closed-form expression:

$$F_{\mathrm{UWG}}(x;\alpha,\beta,p)=\frac{e^{-\alpha(-\log x)^{\beta}}}{1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)},\tag{2.6}$$

where $0<x<1$, $\alpha>0$ represents the scale parameter, $\beta>0$ denotes the shape parameter, and $0<p<1$ acts as the compounding parameter. By taking the first derivative of Eq (2.6) with respect to x, the corresponding PDF of the UWG distribution is obtained as follows

$$f_{\mathrm{UWG}}(x;\alpha,\beta,p)=\frac{1}{x}\frac{(1-p)\alpha\beta(-\log x)^{\beta-1}e^{-\alpha(-\log x)^{\beta}}}{\left[1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)\right]^{2}}.\tag{2.7}$$

To fully capture the dynamic reliability traits of the UWG model, we derive its survival function ($S_{\mathrm{UWG}}(x)$) and HRF ($h_{\mathrm{UWG}}(x)$), given, respectively, by

$$S_{\mathrm{UWG}}(x;\alpha,\beta,p)=\frac{(1-p)\left(1-e^{-\alpha(-\log x)^{\beta}}\right)}{1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)}\tag{2.8}$$

and

$$h_{\mathrm{UWG}}(x;\alpha,\beta,p)=\frac{\alpha\beta(-\log x)^{\beta-1}e^{-\alpha(-\log x)^{\beta}}}{x\left[1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)\right]\left(1-e^{-\alpha(-\log x)^{\beta}}\right)},\tag{2.9}$$

where $0<x<1$, $\alpha>0$ represents the scale parameter, $\beta>0$ denotes the shape parameter, and $0<p<1$ acts as the compounding parameter. To evaluate the structural capabilities and visual properties of the newly proposed model, Figures 1 and 2 provide graphical illustrations of the PDF and HRF across four distinct parameter vectors (α,β,p). These visual trajectories highlight the profound impact that the compounding parameter p exerts on the overall flexibility of the distribution, enabling it to fit diverse data topologies.

As depicted in Figure 1, the density curves display an exceptionally broad range of geometric profiles. Depending on the values assigned to the shape parameter β and the geometric parameter p, the PDF can be strictly decreasing, highly right-skewed, symmetrical, or heavily left-skewed. When $\beta\leq 1$, the distribution is capable of capturing highly J-shaped or reversed-J trajectories, which are essential for modeling fraction data concentrated near the boundaries. For $\beta>1$, the model manifests various unimodal structures. The parameter p modulates the peak height (kurtosis) and tail thickness of the density functions. This behavioral elasticity proves that the UWG model can accommodate bounded observations that conventional unit distributions struggle to model without risking severe specification biases.

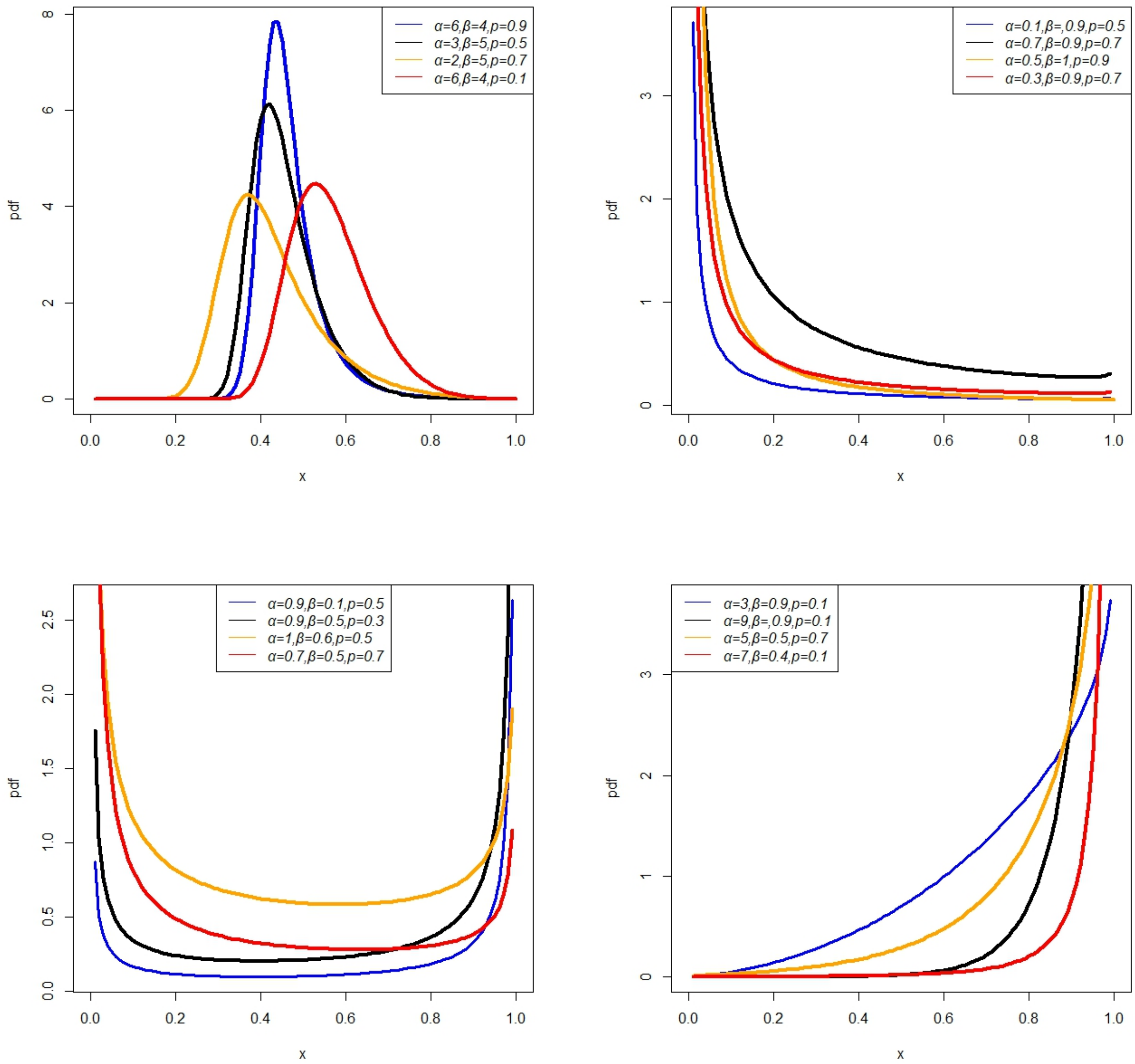

Figure 1. The PDF plots of the UWG distribution.

The true advantage of the compounding mechanism shows in the HRF profiles illustrated in Figure 2. While the classical baseline unit-Weibull distribution has notable constraints regarding its hazard shapes, the compounded UWG distribution overcomes this limitation entirely. Figure 2 confirms that the $h_{\mathrm{UWG}}(x)$ yields classic monotonic shapes, such as strictly increasing or strictly decreasing trends; more importantly, it accommodates intricate non–monotonic risk profiles. Specifically, the model can capture a classic bathtub-shaped profile (high risk at the infant stage, stabilizing during mid-life, and escalating near the upper boundary of 1). Furthermore, it exhibits complex increasing–decreasing–increasing patterns, alongside upside-down bathtub (unimodal) trajectories. This capability is exceptionally vital when modeling proportions or rates in survival systems that experience fluctuating risk periods over time. The capacity to represent such varied hazard structures verifies that the UWG model provides a highly adaptable theoretical alternative for lifetime and empirical rate analysis.

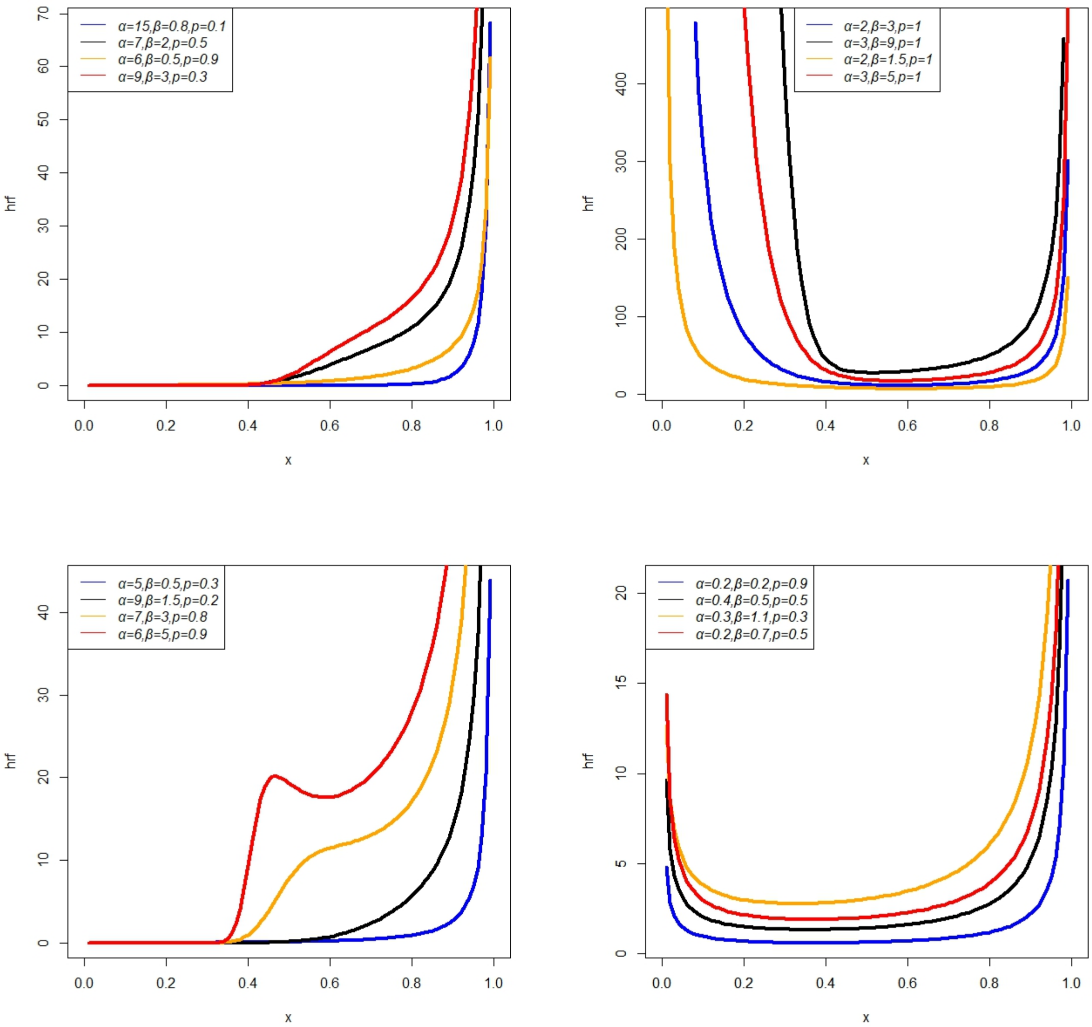

Figure 2. The HRF plots of the UWG distribution.

3. Identifiability analysis of the UWG distribution

The property of identifiability is fundamental to statistical inference, ensuring that distinct parameter vectors correspond to distinct probability distributions and thereby guaranteeing the uniqueness of maximum likelihood estimates. In this section, we establish the identifiability of the UWG distribution with parameter vector $\boldsymbol{\theta} = (\alpha, \beta, p)^\top \in \Theta$, where the parameter space is $\Theta = \mathbb{R}^+ \times \mathbb{R}^+ \times (0, 1)$, and the density $f_{\mathrm{UWG}}(x; \boldsymbol{\theta})$ is given in Eq (2.7). We provide both a rigorous analytical proof and a discussion of the underlying structural mechanisms that ensure identifiability. Formally, a parametric family of distributions $\{F_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ is said to be identifiable if the mapping $\boldsymbol{\theta} \mapsto F_{\boldsymbol{\theta}}$ is injective; that is, for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta$, we have

$$F_{\boldsymbol{\theta}_1}(x) = F_{\boldsymbol{\theta}_2}(x) \quad \text{for all } x \in (0, 1) \quad \Longrightarrow \quad \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2. \tag{3.1}$$

Equivalently, it suffices to establish the implication at the level of the PDF, since densities that are equal almost everywhere with respect to Lebesgue measure must share identical parameter values. We therefore aim to prove that

$$f_{\mathrm{UWG}}(x;\alpha_1,\beta_1,p_1) = f_{\mathrm{UWG}}(x;\alpha_2,\beta_2,p_2), \quad x \in (0,1) \quad \Longrightarrow \quad (\alpha_1,\beta_1,p_1) = (\alpha_2,\beta_2,p_2). \tag{3.2}$$

To facilitate the identifiability argument, we adopt the logarithmic transformation $y = -\log x \in (0,\infty)$, under which the UWG density assumes the form

$$g(y;\alpha,\beta,p) = \frac{(1-p)\alpha\beta y^{\beta-1} e^{-\alpha y^{\beta}}}{\left[1 - p\left(1 - e^{-\alpha y^{\beta}}\right)\right]^2}, \quad y > 0. \tag{3.3}$$

Consider two parameter vectors (α_1,β_1,p_1) and (α_2,β_2,p_2) such that $g(y;\alpha_1,\beta_1,p_1) = g(y;\alpha_2,\beta_2,p_2)$ for almost all $y > 0$. Taking the ratio and equating to unity, we obtain the functional equation

$$\frac{(1-p_1)\alpha_1\beta_1 y^{\beta_1-1} e^{-\alpha_1 y^{\beta_1}}}{\left[1 - p_1 + p_1 e^{-\alpha_1 y^{\beta_1}}\right]^2} = \frac{(1-p_2)\alpha_2\beta_2 y^{\beta_2-1} e^{-\alpha_2 y^{\beta_2}}}{\left[1 - p_2 + p_2 e^{-\alpha_2 y^{\beta_2}}\right]^2}, \quad \text{a.e. } y > 0. \tag{3.4}$$

3.1. Identifiability via asymptotic matching of tail exponents

The identifiability of the shape parameter β can be established through a careful comparison of the asymptotic behavior of the density ratio as $y \to 0^+$ and $y \to \infty$. Since the equality in Eq (3.4) holds for almost all y, it must hold in particular in arbitrarily small neighborhoods of the boundaries.

Step 1: Identifiability of β via the origin behavior.

As $y \to 0^+$, we use the expansion $e^{-\alpha y^{\beta}} = 1 - \alpha y^{\beta} + O(y^{2\beta})$. The denominator in Eq (3.3) satisfies

$$1 - p + p e^{-\alpha y^{\beta}} = 1 - p\alpha y^{\beta} + O(y^{2\beta}). \tag{3.5}$$

Consequently, the density behaves asymptotically as

$$g(y;\alpha,\beta,p) \sim (1-p)\alpha\beta y^{\beta-1}, \quad y \to 0^+. \tag{3.6}$$

If $g(y;\alpha_1,\beta_1,p_1) = g(y;\alpha_2,\beta_2,p_2)$ for all y in a neighborhood of zero, then their leading asymptotic terms must coincide, yielding

$$(1-p_1)\alpha_1\beta_1 y^{\beta_1-1} \sim (1-p_2)\alpha_2\beta_2 y^{\beta_2-1}, \quad y \to 0^+. \tag{3.7}$$

Taking logarithms of the ratio and passing to the limit $y \to 0^+$, we obtain

$$\lim_{y\to 0^+} \left[(\beta_1 - \beta_2)\log y + \log\left(\frac{(1-p_1)\alpha_1\beta_1}{(1-p_2)\alpha_2\beta_2}\right)\right] = 0. \tag{3.8}$$

For this limit to be finite, the coefficient of $\log y$ must vanish, forcing $\beta_1 = \beta_2$. Thus, the shape parameter β is uniquely determined by the boundary exponent of the density at the origin in the transformed domain.

Step 2: Identifiability of α via the tail behavior.

Having established $\beta_1 = \beta_2 = \beta$, we now exploit the tail behavior as $y \to \infty$ to identify the scale parameter α. For a large y, the exponential term $e^{-\alpha y^\beta}$ decays to zero, and the density in Eq (3.3) is dominated by

$$g(y; \alpha, \beta, p) \sim \frac{\alpha\beta}{1-p} y^{\beta-1} e^{-\alpha y^\beta}, \quad y \to \infty. \tag{3.9}$$

With β already identified, equating the asymptotic forms for the two parameter vectors gives

$$\frac{\alpha_1}{1-p_1} e^{-\alpha_1 y^\beta} \sim \frac{\alpha_2}{1-p_2} e^{-\alpha_2 y^\beta}, \quad y \to \infty. \tag{3.10}$$

Taking logarithms, dividing by y^β, and taking the limit $y \to \infty$ yields

$$-\alpha_1 + \lim_{y\to\infty} \frac{1}{y^\beta} \log\left(\frac{\alpha_1(1-p_2)}{\alpha_2(1-p_1)}\right) = -\alpha_2. \tag{3.11}$$

Since the logarithmic term vanishes in the limit, we immediately obtain $\alpha_1 = \alpha_2$. Hence, the scale parameter α is identifiable from the exponential tail decay rate.

Step 3: Identifiability of p via exact functional equality.

With $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 = \beta_2 = \beta$ established, the equality of densities in Eq (3.4) reduces to an algebraic condition on the compounding parameter. Canceling the common factors yields the simplified identity

$$\frac{1-p_1}{\left[1-p_1+p_1 e^{-\alpha y^\beta}\right]^2} = \frac{1-p_2}{\left[1-p_2+p_2 e^{-\alpha y^\beta}\right]^2}, \quad \text{a.e. } y > 0. \tag{3.12}$$

Define the function $h(u; p) = (1-p)/(1-p+pu)^2$ for $u \in (0, 1]$, where $u = e^{-\alpha y^\beta}$. Equation (3.12) asserts that $h(u; p_1) = h(u; p_2)$ for all u in an interval (since $y \mapsto e^{-\alpha y^\beta}$ is continuous and non–constant). The function $h(u; p)$ is strictly monotone in p for each fixed $u \in (0, 1)$. To see this, compute the partial derivative

$$\begin{aligned}\frac{\partial h}{\partial p} &= \frac{-(1-p+pu)^2 - 2(1-p)(1-p+pu)(u-1)}{(1-p+pu)^4} \\ &= \frac{[-(1-p+pu) - 2(1-p)(u-1)]}{(1-p+pu)^3} = \frac{1-p+pu-2u}{(1-p+pu)^3}.\end{aligned} \tag{3.13}$$

Since $0 < u < 1$ and $0 < u < 1$, we see that the denominator is always positive; hence, the sign of $\partial h/\partial p$ is determined by the numerator. A straightforward algebraic simplification shows that $\partial h/\partial p < 0$ for all $p \in (0, 1)$ and $u \in (0, 1)$, establishing strict monotonicity of $p \mapsto h(u; p)$. Consequently, the equality $h(u; p_1) = h(u; p_2)$ for u in a set of positive measure forces $p_1 = p_2$.

Thus, all three parameters are uniquely determined, completing the proof of identifiability.

3.2. Alternative proof via moment generating structure

An alternative and perhaps more elegant proof of identifiability can be constructed by examining the structure of the CDF. From the definition of the UWG distribution, the CDF takes the form

$$F_{\mathrm{UWG}}(x; \alpha, \beta, p) = \frac{1 - e^{-\alpha(-\log x)^\beta}}{1 - p\left(1 - e^{-\alpha(-\log x)^\beta}\right)}. \tag{3.14}$$

This can be recognized as a composition $F_{\mathrm{UWG}}(x) = \phi_p(F_W(x))$, where $F_W(x) = e^{-\alpha(-\log x)^\beta}$ is the CDF of the unit-Weibull distribution and $\phi_p(t) = 1 - t/(1 - p + pt)$ is the odds-generating function of the geometric distribution. Since $\phi_p : [0, 1] \to [0, 1]$ is strictly increasing and invertible for each $p \in (0, 1)$, we have

$$F_W(x) = \phi_p^{-1}(F_{\mathrm{UWG}}(x)) = \frac{1 - (1 - p)F_{\mathrm{UWG}}(x)}{1 + pF_{\mathrm{UWG}}(x)}. \tag{3.15}$$

If $F_{\mathrm{UWG}}(x; \alpha_1, \beta_1, p_1) = F_{\mathrm{UWG}}(x; \alpha_2, \beta_2, p_2)$ for all x, then applying the inverse transformation yields the equality of unit-Weibull CDFs, possibly with different parameters. The identifiability of the unit-Weibull distribution (which follows from the identifiability of the ordinary Weibull distribution under the logarithmic transformation) then ensures $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. Once α and β are fixed, the injectivity of $p \mapsto \phi_p$ guarantees $p_1 = p_2$. This structural decomposition provides an elegant confirmation of the result obtained through asymptotic analysis. It is noteworthy that the UWG distribution remains identifiable even as $p \to 0^+$ or $p \to 1^-$, where the model reduces to the unit-Weibull distribution or degenerates, respectively. This continuity of identifiability at the boundary of the parameter space ensures stable inferential behavior across the full range of admissible parameter values, a property not universally shared by all compound distributions.

4. Convexity and asymptotic properties of the UWG density

In this section, we examine the analytical shape properties of the PDF $f_{\mathrm{UWG}}(x; \alpha, \beta, p)$ given in Eq (2.7), with particular emphasis on its convexity characteristics and limiting behavior. Throughout our analysis, we restrict attention to the domain $x \in (0, 1)$, with the scale parameter $\alpha > 0$, the shape parameter $\beta > 0$, and the compounding parameter $p \in (0, 1)$. To facilitate our investigation, we introduce the logarithmic transformation $y = -\log x$, which maps the unit interval $(0, 1)$ bijectively onto the positive real line $(0, \infty)$. Under this transformation, the density takes the form

$$g(y; \alpha, \beta, p) = \frac{(1 - p)\alpha\beta y^{\beta-1} e^{-\alpha y^\beta}}{\left[1 - p\left(1 - e^{-\alpha y^\beta}\right)\right]^2}, \quad y > 0, \tag{4.1}$$

where the Jacobian $|dx/dy| = e^{-y} = x$ has been absorbed accordingly. The function $g(y)$ inherits the analytical structure of the original density while eliminating the logarithmic singularity at the origin, thereby simplifying the subsequent convexity analysis. We note that Eq (4.1) can be expressed compactly in terms of the Weibull survival function $\bar{F}_W(y) = e^{-\alpha y^\beta}$ as follows

$$g(y) = \frac{(1 - p)\alpha\beta y^{\beta-1} \bar{F}_W(y)}{\left[1 - p + p\bar{F}_W(y)\right]^2}. \tag{4.2}$$

4.1. Log-convexity and log-concavity analysis

The log-convexity or log-concavity of a density function governs the modality, tail behavior, and reliability properties of the corresponding distribution. We therefore compute the second derivative of $\log g(y)$ with respect to y.

Taking the natural logarithm of Eq (4.1) yields

$$\log g(y) = \log\left[(1 - p)\alpha\beta\right] + (\beta - 1)\log y - \alpha y^\beta - 2\log\left[1 - p\left(1 - e^{-\alpha y^\beta}\right)\right]. \tag{4.3}$$

Differentiating Eq (4.3) with respect to y, we obtain

$$\frac{d}{dy}\log g(y) = \frac{\beta-1}{y} - \alpha\beta y^{\beta-1} - \frac{2p\alpha\beta y^{\beta-1}e^{-\alpha y^{\beta}}}{1-p\left(1-e^{-\alpha y^{\beta}}\right)}. \tag{4.4}$$

A second differentiation yields the crucial quantity

$$\begin{aligned}\frac{d^2}{dy^2}\log g(y) = &-\frac{\beta-1}{y^2} - \alpha\beta(\beta-1)y^{\beta-2} \\ &- \frac{2p\alpha\beta y^{\beta-2}e^{-\alpha y^{\beta}}\left[(\beta-1-\alpha\beta y^{\beta})\left(1-p+pe^{-\alpha y^{\beta}}\right)+p\alpha\beta y^{\beta}e^{-\alpha y^{\beta}}\right]}{\left[1-p+pe^{-\alpha y^{\beta}}\right]^2}.\end{aligned} \tag{4.5}$$

Case $\beta \le 1$. When $\beta \le 1$, the first term $-\frac{\beta-1}{y^2}$ is non-negative or zero. However, the remaining terms are strictly negative for all $y > 0$, as $\alpha\beta(\beta-1)y^{\beta-2} \le 0$ and the final fraction is manifestly positive owing to $p \in (0,1)$ and all constituent factors being non-negative. Consequently, the sign of $\frac{d^2}{dy^2}\log g(y)$ is indeterminate in general, but for $\beta < 1$, the dominant term near the origin is $-\frac{\beta-1}{y^2} > 0$, implying local log-convexity near $y = 0$. As $y \to \infty$, the exponential decay dominates and the density becomes log-concave in the tail. Thus, for $\beta \le 1$, the density typically exhibits a log-convex region near the origin, transitioning to log-concavity in the extreme right tail, which permits at most one interior mode.

Case $\beta > 1$. For $\beta > 1$, all three components in Eq (4.5) are strictly negative for every $y > 0$. The first term $-\frac{\beta-1}{y^2} < 0$, the second term $-\alpha\beta(\beta-1)y^{\beta-2} < 0$, and the third term is also negative since all factors in the numerator are positive. Therefore, we have

$$\frac{d^2}{dy^2}\log g(y) < 0 \quad \text{for all } y > 0 \text{ when } \beta > 1. \tag{4.6}$$

This establishes that $g(y)$ is strictly log-concave on $(0,\infty)$ whenever $\beta > 1$. Moreover, the log-concavity result in Eq (4.6) holds uniformly for all $\alpha > 0$ and $p \in (0,1)$, indicating that the transformed density possesses a stable structural property throughout the admissible parameter space. However, owing to the nonlinear transformation $x = e^{-y}$, this property alone does not imply the corresponding log-concavity or monotonicity of the density and its associated reliability functions on the original x-scale. Therefore, the shapes of the density, hazard rate, and reversed hazard rate are investigated through numerical illustrations, which demonstrate that the proposed UWG distribution is capable of exhibiting a broad range of functional forms for different parameter configurations.

4.2. Asymptotic behavior at the boundaries

We now characterize the limiting behavior of $f_{\mathrm{UWG}}(x)$ as $x \to 0^+$ (equivalently $y \to \infty$) and $x \to 1^-$ (equivalently $y \to 0^+$).

4.2.1. Behavior near the origin ($x \to 0^+$)

As $x \to 0^+$, we have $y = -\log x \to \infty$. In this regime, $e^{-\alpha y^\beta} \to 0$, so the denominator in Eq (4.1) approaches $(1-p)^2$. The density therefore behaves asymptotically as

$$g(y) \sim \frac{\alpha\beta}{1-p} y^{\beta-1} e^{-\alpha y^\beta}, \quad y \to \infty. \tag{4.7}$$

Transforming back to the original scale, we obtain

$$f_{\mathrm{UWG}}(x) \sim \frac{\alpha\beta}{1-p} \frac{(-\log x)^{\beta-1}}{x} e^{-\alpha(-\log x)^\beta}, \quad x \to 0^+. \tag{4.8}$$

The limiting behavior of the density near the origin depends on the shape parameter β. Specifically, $f_{\mathrm{UWG}}(x) \to \infty$ as $x \to 0^+$ for $0 < \beta < 1$, whereas $f_{\mathrm{UWG}}(x) \to 0$ as $x \to 0^+$ for $\beta > 1$. Thus, the boundary singularity is present only when $0 < \beta < 1$. The geometric compounding parameter p further modulates this behavior through the factor $(1-p)^{-1}$. As $p \to 1^-$, the amplification becomes increasingly pronounced, whereas as $p \to 0^+$, the density converges to the baseline unit-Weibull distribution.

4.2.2. Behavior near the upper bound ($x \to 1^-$)

As $x \to 1^-$, we have $y = -\log x \to 0^+$. With the expansion $e^{-\alpha y^\beta} = 1 - \alpha y^\beta + \frac{1}{2}\alpha^2 y^{2\beta} + O(y^{3\beta})$, the denominator in Eq (4.1) satisfies

$$1 - p\left(1 - e^{-\alpha y^\beta}\right) = 1 - p\alpha y^\beta + O(y^{2\beta}). \tag{4.9}$$

Hence,

$$\left[1 - p\left(1 - e^{-\alpha y^\beta}\right)\right]^{-2} = 1 + 2p\alpha y^\beta + O(y^{2\beta}). \tag{4.10}$$

Substituting into Eq (4.1), we obtain

$$g(y) \sim (1-p)\alpha\beta y^{\beta-1} \left(1 - \alpha y^\beta\right)\left(1 + 2p\alpha y^\beta\right), \quad y \to 0^+. \tag{4.11}$$

Retaining the leading-order term, the asymptotic form simplifies to

$$g(y) \sim (1-p)\alpha\beta y^{\beta-1}, \quad y \to 0^+. \tag{4.12}$$

Returning to the original variable x via $y \sim 1 - x$ as $x \to 1^-$, we find

$$f_{\mathrm{UWG}}(x) \sim (1-p)\alpha\beta(1-x)^{\beta-1}, \quad x \to 1^-. \tag{4.13}$$

Equation (4.13) reveals three distinct behavioral regimes near the upper boundary, controlled entirely by the shape parameter β as follows:

- $\beta = 1$ (exponential-type boundary): $f_{\mathrm{UWG}}(x) \to (1-p)\alpha > 0$, indicating a finite non–zero density at $x = 1^-$;
- $0 < \beta < 1$ (pole at unity): $f_{\mathrm{UWG}}(x) \to +\infty$ as $x \to 1^-$, with the singularity of order $(1-x)^{\beta-1}$;
- $\beta > 1$ (vanishing at unity): $f_{\mathrm{UWG}}(x) \to 0$ as $x \to 1^-$, with polynomial decay rate $(1-x)^{\beta-1}$.

The compounding parameter p enters the leading asymptotic constant multiplicatively through $(1-p)$, uniformly attenuating the density near $x = 1$ without affecting the qualitative nature of the boundary behavior. This stands in contrast to its effect near the origin, where it appears as a denominator amplification factor.

4.3. Modality and shape flexibility

Combining the convexity and asymptotic analyses above, we can fully characterize the modality of the UWG's density. From Eq (4.4), the critical points satisfy

$$\frac{\beta-1}{y}-\alpha\beta y^{\beta-1}=\frac{2p\alpha\beta y^{\beta-1}e^{-\alpha y^{\beta}}}{1-p\left(1-e^{-\alpha y^{\beta}}\right)}. \tag{4.14}$$

For $\beta > 1$, log-concavity guarantees a unique solution to Eq (4.14), corresponding to the single mode. For $\beta \leq 1$, the density may exhibit a mode near the origin (due to the log-convex region transitioning to log-concavity) or remain strictly decreasing when the right-hand side of Eq (4.14) never intersects the left-hand side. The parameter p modulates the location of the mode through the compounding term on the right-hand side: As $p \to 0^+$, the equation reduces to that of the ordinary unit-Weibull distribution, while increasing p shifts the mode toward smaller values of y (equivalently, larger values of x). This rich parametric control over shape underscores the UWG distribution's utility for modeling proportion-type data with diverse density forms.

5. Statistical properties

5.1. Quantile function and related descriptive measures

The quantile function serves as an indispensable tool for generating random variables via the inverse transform method and for calculating robust descriptive statistics such as the median, quartile deviation, and measures of asymmetry. By definition, the u^{th} quantile function, denoted $x_u = F^{-1}_{\text{UWG}}(u)$ for $0 < u < 1$, is obtained by analytically inverting the CDF presented in Eq (2.6)

$$u=\frac{e^{-\alpha(-\log x_u)^{\beta}}}{1-p\left(1-e^{-\alpha(-\log x_u)^{\beta}}\right)}. \tag{5.1}$$

Solving explicitly for x_u yields the following functional form:

$$x_u=\exp\left\{-\left[-\frac{1}{\alpha}\log\left(\frac{u(1-p)}{1-up}\right)\right]^{\frac{1}{\beta}}\right\},\quad 0<u<1. \tag{5.2}$$

By setting $u = 0.50$ in Eq (5.2), the median (M) of the UWG distribution is expressed explicitly as:

$$M=\exp\left\{-\left[-\frac{1}{\alpha}\log\left(\frac{1-p}{2-p}\right)\right]^{\frac{1}{\beta}}\right\}, \tag{5.3}$$

where $\alpha > 0$ represents the scale parameter, $\beta > 0$ denotes the shape parameter, and $0 < p < 1$ acts as the compounding parameter.

Analogously, the first quartile ($x_{0.25}$) and the third quartile ($x_{0.75}$) are determined by substituting $u = 0.25$ and $u = 0.75$, respectively, into Eq (5.2). To evaluate the mathematical behavior of these quantiles under fluctuating parameter spaces, numerical estimates across 12 distinct parameter vectors are reported in Table 1.

Table 1. Quartile estimates of the UWG distribution.

Parameters			Quantiles		
α	β	p	$x_{0.25}$	$x_{0.50}$ (Median)	$x_{0.75}$
0.5	2.0	0.50	0.13907	0.22711	0.36394
1.0			0.24784	0.35059	0.48933
3.0			0.44692	0.54599	0.66190
4.0			0.49784	0.59210	0.69952
2.0	0.5	0.50	0.38804	0.73953	0.93685
	1.0		0.37796	0.57735	0.77460
	2.0		0.37292	0.47656	0.60327
	4.0		0.37040	0.42278	0.49120
3.0	2.0	0.01	0.50579	0.61729	0.73270
		0.10	0.49702	0.60709	0.72319
		0.60	0.42973	0.52403	0.63795
		0.90	0.34305	0.40900	0.49702

The numerical results tabulated in Table 1 unveil critical insights into the localized behavior of the UWG distribution. When β and p are held constant, an escalation in the scale parameter α shifts all three quartiles toward the upper boundary of 1, compressing the interquartile range. Conversely, tracking the impact of the shape parameter β shows that as it grows, the median and upper quartiles decrease significantly, indicating a rightward-to-leftward concentration of probability mass. Finally, the compounding parameter p shifts the quantiles downward as it approaches unity, demonstrating its direct utility in modeling highly skewed unit Datasets. Furthermore, the analytical determination of the mode is established by maximizing the log-likelihood function of a single observation, $\log f_{\mathrm{UWG}}(x)$. Differentiating with respect to x yields

$$\begin{aligned}\frac{d \log f_{\mathrm{UWG}}(x)}{dx} =& -\frac{1}{x} - \frac{1}{x\log x}\left[(\beta-1) + \alpha\beta(-\log x)^{\beta}\right] \\ &- \frac{2p\alpha\beta(-\log x)^{\beta-1} e^{-\alpha(-\log x)^{\beta}}}{x\left[1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)\right]} = 0. \end{aligned} \tag{5.4}$$

The mode is the roots of the non–linear equation above, which can be computed via standard numerical iterative routines such as the Newton–Raphson method.

5.2. *Ordinary and central moments*

Moments constitute foundational building blocks for evaluating central tendency, dispersion, and tail behaviors. To derive the r^{th} ordinary moment about the origin (μ_r') for the UWG distribution, we begin with the continuous definition:

$$\mu_r' = \int_0^1 x^r f_{\mathrm{UWG}}(x)\, dx = \int_0^1 x^{r-1} \frac{(1-p)\alpha\beta(-\log x)^{\beta-1} e^{-\alpha(-\log x)^{\beta}}}{\left[1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)\right]^2}\, dx. \tag{5.5}$$

To simplify the denominator term, we use the generalized power series expansion that is valid for $|z| < 1$

$$(1 - z)^{-2} = \sum_{i=0}^{\infty}(i + 1)z^i. \tag{5.6}$$

Substituting $z = p\left(1 - e^{-\alpha(-\log x)^\beta}\right)$ into Eq (5.6) transforms the integrand. Subsequent application of the standard binomial theorem to the term $\left(1 - e^{-\alpha(-\log x)^\beta}\right)^i$ yields

$$\left(1 - e^{-\alpha(-\log x)^\beta}\right)^i = \sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}e^{-\alpha(i-m)(-\log x)^\beta}. \tag{5.7}$$

By utilizing the transformation $y = -\log x$ (which maps the limits from $x \in (0, 1)$ to $y \in (\infty, 0)$), and integrating via the definition of the complete gamma function, the r^{th} raw moment reduces to the following infinite double series

$$\mu'_r = (1 - p)\alpha\beta\sum_{i=0}^{\infty}(i + 1)p^i\sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\int_0^\infty y^{\beta-1}\exp\left\{-ry - \alpha(1 + i - m)y^\beta\right\}dy. \tag{5.8}$$

Applying a series expansion to e^{-ry} yields the finalized analytical expression for the r^{th} raw moment:

$$\mu'_r = (1 - p)\alpha\sum_{i=0}^{\infty}(i + 1)p^i\sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\frac{\Gamma\left(1 + \frac{r}{\beta}\right)}{[\alpha(1 + i - m)]^{1+\frac{r}{\beta}}}. \tag{5.9}$$

Setting $r = 1$ in Eq (5.9) provides the expected value or mean of the distribution:

$$\mu'_1 = (1 - p)\alpha\sum_{i=0}^{\infty}(i + 1)p^i\sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\frac{\Gamma\left(1 + \frac{1}{\beta}\right)}{[\alpha(1 + i - m)]^{1+\frac{1}{\beta}}}. \tag{5.10}$$

To explore the shape parameters systematically, the relationships between the ordinary moments and the cumulants (C_r) are exploited. The r^{th} cumulant is recursively determined by:

$$C_r = \mu'_r - \sum_{j=1}^{r-1}\binom{r-1}{j-1}C_j\mu'_{r-j}. \tag{5.11}$$

Evaluating Eq (5.11) sequentially up to the fourth order yields:

$$\begin{aligned}
&C_1 = \mu'_1 \quad \text{mean}),\\
&C_2 = \mu'_2 - (\mu'_1)^2 \quad \text{(variance)},\\
&C_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3,\\
&C_4 = \mu'_4 - 4\mu'_3\mu'_1 - 3(\mu'_2)^2 + 12\mu'_2(\mu'_1)^2 - 6(\mu'_1)^4.
\end{aligned}$$

Using these derived components, the standardized coefficients of skewness (CS) and kurtosis (CK) are computed via

$$CS = \frac{C_3}{(C_2)^{3/2}}, \quad \text{and} \quad CK = \frac{C_4}{(C_2)^2}. \tag{5.12}$$

Table 2 presents the raw moments and descriptive indices for various parameter combinations of the three-parameter distribution under study. The table is structured into three distinct blocks, each isolating the effect of one parameter while holding the others constant.

Table 2. Quantitative characteristics of the UWG distribution.

Parameters			Raw moments				Descriptive indices		
α	β	p	μ_1'	μ_2'	μ_3'	μ_4'	Variance	Skewness	Kurtosis
0.5	2.0	0.50	0.2728	0.1060	0.0525	0.0305	0.0316	1.1365	3.9840
1.0			0.3826	0.1778	0.0962	0.0584	0.0314	0.7440	3.0868
3.0			0.5594	0.3359	0.2148	0.1450	0.0230	0.3188	2.5931
4.0			0.6021	0.3826	0.2553	0.1778	0.0201	0.2366	2.5574
2.0	0.5	0.50	0.6437	0.5179	0.4462	0.3981	0.1036	−0.6155	2.0295
	1.0		0.5708	0.3863	0.2876	0.2274	0.0605	−0.1301	2.0326
	2.0		0.4959	0.2728	0.1639	0.1060	0.0268	0.4531	2.6937
	4.0		0.4391	0.2020	0.0976	0.0496	0.0092	0.9369	4.0887
3.0	2.0	0.01	0.6192	0.4073	0.2818	0.2034	0.0239	0.0075	2.3999
		0.10	0.6108	0.3969	0.2718	0.1944	0.0239	0.0515	2.4088
		0.60	0.5404	0.3144	0.1955	0.1289	0.0224	0.4189	2.7213
		0.90	0.4327	0.2037	0.1046	0.0585	0.0165	1.0263	4.2638

The first block examines the influence of the shape parameter α, with $\beta = 2.0$ and the mixing proportion $p = 0.50$ held fixed. As α increases from 0.5 to 4.0, a clear monotonic trend emerges in the first four raw moments, all of which increase substantially. The mean μ_1' rises from 0.2728 to 0.6021, while the second raw moment μ_2' more than triples from 0.1060 to 0.3826. Despite this absolute growth in the moments, the variance exhibits an inverse relationship with α, declining from 0.0316 to 0.0201, indicating that the distribution becomes more concentrated around its mean at higher values of α. The shape characteristics undergo a pronounced transformation: Skewness decreases sharply from 1.1365 to 0.2366, reflecting a marked shift from strong positive asymmetry toward approximate symmetry. Concurrently, the kurtosis falls from 3.9840 to 2.5574, suggesting that the tails of the distribution become substantially lighter as α increases.

The second block investigates the effect of the parameter β, ranging from 0.5 to 4.0, while $\alpha = 2.0$ and $p = 0.50$ remain constant. The behavior observed here contrasts notably with that of α. All four raw moments decrease monotonically as β grows, with the mean dropping from 0.6437 to 0.4391 and the second raw moment declining from 0.5179 to 0.2020. The variance undergoes a dramatic reduction from 0.1036 at $\beta = 0.5$ to 0.0092 at $\beta = 4.0$, representing a decrease of over 90%. The skewness exhibits a striking sign reversal: At $\beta = 0.5$, the distribution displays moderate negative skewness (−0.6155), which approaches zero at $\beta = 1.0$ (−0.1301), crosses into positive territory at $\beta = 2.0$ (0.4531), and reaches 0.9369 at $\beta = 4.0$. The kurtosis follows a non-monotonic, U-shaped trajectory, starting at 2.0295 for $\beta = 0.5$, remaining nearly unchanged at $\beta = 1.0$ (2.0326), then rising to 2.6937 at $\beta = 2.0$ and ultimately reaching 4.0887 at $\beta = 4.0$. This pattern suggests that both low and high values of β produce distributions with distinct tail behaviors, while intermediate values yield lighter tails.

The third block explores the role of the mixing parameter p, which varies from 0.01 to 0.90 with

$\alpha = 3.0$ and $\beta = 2.0$ held fixed. As p increases, all raw moments decrease consistently, with the mean declining from 0.6192 to 0.4327 and the second raw moment falling from 0.4073 to 0.2037. The variance shows a modest reduction from 0.0239 to 0.0165, indicating that the dispersion of the distribution is relatively stable across the range of p when compared with the effects of α and β. Both skewness and kurtosis increase systematically with p. The skewness transitions from a value near zero (0.0075 at $p = 0.01$) to pronounced positive asymmetry (1.0263 at $p = 0.90$), while the kurtosis rises from 2.3999 to 4.2638 over the same interval. The near-zero skewness and moderate kurtosis at very low p suggest that one component dominates the mixture, producing a relatively symmetric, light-tailed distribution. As p increases and the second component gains weight, the distribution acquires heavier tails and stronger positive skewness, characteristic of a mixture where one component exhibits greater variability and right skew.

Taken together, these results reveal the distinct and interpretable roles played by each parameter. The parameter α primarily governs the central location and symmetry of the distribution, with higher values promoting symmetry and lighter tails. The parameter β exerts a powerful influence on both the dispersion and the direction of asymmetry, capable of shifting the distribution from negatively to positively skewed. The mixing parameter p controls the relative contribution of the underlying components, modulating tail weight and skewness without substantially altering the overall variance. This parametric flexibility makes the distribution well-suited for modelling heterogeneous data with varying degrees of asymmetry and tail behavior.

5.3. Moment generating function

The moment generating function (MGF), defined as $M_X(t) = E[e^{tX}]$, offers an integral transform perspective of the probability structure. Utilizing the foundational Maclaurin series expansion for the exponential term, $e^{tx} = \sum_{l=0}^{\infty} \frac{(tx)^l}{l!}$, we can write

$$M_X(t) = \int_0^1 \sum_{l=0}^{\infty} \frac{(tx)^l}{l!} f_{\mathrm{UWG}}(x)\,dx = \sum_{l=0}^{\infty} \frac{t^l}{l!} \left(\int_0^1 x^l f_{\mathrm{UWG}}(x)\,dx \right) = \sum_{l=0}^{\infty} \frac{t^l}{l!} \mu_l'. \tag{5.13}$$

By directly embedding the expression for the l^{th} ordinary raw moment derived in Eq (5.9), the closed-form representation of the MGF for the UWG distribution simplifies to

$$M_X(t) = (1-p)\alpha \sum_{l=0}^{\infty} \sum_{i=0}^{\infty} \sum_{m=0}^{i} \frac{(i+1)(-1)^{i-m} p^i t^l \binom{i}{m} \Gamma\left(1 + \frac{l}{\beta}\right)}{l!\,[\alpha(1+i-m)]^{1+\frac{l}{\beta}}}. \tag{5.14}$$

5.4. Mean deviations about the mean and median

The amount of scatter within a population is effectively measured using the mean deviations about central landmarks. Let $\mu = \mu_1'$ represent the mean (Eq (5.10)) and M denote the median (Eq (5.2)). The mean deviation about the mean ($D_\mu(X)$) and the mean deviation about the median ($D_M(X)$) are defined, respectively, as follows

$$D_\mu(X) = \int_0^1 |x-\mu| f_{\mathrm{UWG}}(x)dx = 2\mu F_{\mathrm{UWG}}(\mu) - 2\int_0^\mu x f_{\mathrm{UWG}}(x)dx \tag{5.15}$$

and

$$D_M(X) = \int_0^1 |x - M| f_{\mathrm{UWG}}(x)dx = \mu - 2\int_0^M x f_{\mathrm{UWG}}(x)dx. \tag{5.16}$$

For structural brevity, let us denote a generic threshold boundary by $\phi \in \{\mu, M\}$ and introduce the incomplete first moment function $\psi(\phi) = \int_0^{\phi} x f_{\mathrm{UWG}}(x)dx$. With the algebraic series expansions detailed in Section 3.2 and applying the definition of the upper incomplete gamma function, $\Gamma(a,z) = \int_z^{\infty} w^{a-1}e^{-w}dw$, the integral can be solved as

$$\psi(\phi) = (1-p)\alpha \sum_{i=0}^{\infty}(i+1)p^i \sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\frac{\Gamma\left(1+\frac{1}{\beta},\, \alpha(1+i-m)(-\log\phi)^{\beta}\right)}{[\alpha(1+i-m)]^{1+\frac{1}{\beta}}}. \tag{5.17}$$

By substituting the incomplete moment representation from Eq (5.17) into Eqs (5.15) and (5.16), the explicit expressions for the mean deviations are obtained. The mean deviation about the mean is given by

$$\begin{aligned} D_\mu(X) =& 2\mu\left[\frac{e^{-\alpha(-\log\mu)^{\beta}}}{1-p\left(1-e^{-\alpha(-\log\mu)^{\beta}}\right)}\right] \\ &- 2(1-p)\alpha\sum_{i=0}^{\infty}(i+1)p^i\sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\frac{\Gamma\left(1+\frac{1}{\beta},\, \alpha(1+i-m)(-\log\mu)^{\beta}\right)}{[\alpha(1+i-m)]^{1+\frac{1}{\beta}}}, \end{aligned} \tag{5.18}$$

and the mean deviation about the median is given by

$$D_M(X) = \mu - 2(1-p)\alpha\sum_{i=0}^{\infty}(i+1)p^i\sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\frac{\Gamma\left(1+\frac{1}{\beta},\, \alpha(1+i-m)(-\log M)^{\beta}\right)}{[\alpha(1+i-m)]^{1+\frac{1}{\beta}}}. \tag{5.19}$$

These equations can be readily computed using standard numerical software containing built-in incomplete gamma algorithms.

5.5. *Incomplete moments*

Incomplete moments are vital analytical tools for investigating vital structural features of a probability distribution, such as tail behavior, and for constructing key economic inequality measures. The h^{th} incomplete moment of a random variable X following the UWG distribution, denoted by $I_X(z;h)$, is defined over the bounded sub-interval $[0,z]$ as

$$I_X(z;h) = \int_0^z x^h f_{\mathrm{UWG}}(x)\,dx. \tag{5.20}$$

By utilizing the generalized binomial series expansion from Eq (5.6) to handle the denominator of the density function, substituting $y = -\log x$, and integrating via the upper incomplete gamma function, $\Gamma(a,w) = \int_w^{\infty} u^{a-1}e^{-u}du$, we obtain the following infinite series representation:

$$I_X(z;h) = (1-p)\alpha\sum_{i=0}^{\infty}(i+1)p^i\sum_{m=0}^{i}\binom{i}{m}(-1)^{i-m}\frac{\Gamma\left(1+\frac{h}{\beta},\, \alpha(1+i-m)(-\log z)^{\beta}\right)}{[\alpha(1+i-m)]^{1+\frac{h}{\beta}}}. \tag{5.21}$$

The first two incomplete moments (obtained by setting $h = 1$ and $h = 2$, respectively) provide essential descriptive information about the shape of the distribution. Crucially, the first incomplete moment ($h = 1$) underpins foundational econometric and lifetime indicators, including the Lorenz and Bonferroni curves, the mean residual life function, and the mean waiting time.

5.6. Conditional moments

In survival analysis and life testing frameworks, analyzing the residual risk behavior of a system, given that it has already survived up to an operational checkpoint t, is critically instructive. The n-th conditional moment of a random variable X following the UWG distribution is formally defined as

$$E\left[X^n \mid X > t\right] = \frac{1}{S_{\mathrm{UWG}}(t)} \int_t^1 x^n f_{\mathrm{UWG}}(x)\, dx. \tag{5.22}$$

Substituting the explicit mathematical forms of the density function $f_{\mathrm{UWG}}(x)$ (Eq (2.7)) and the survival function $S_{\mathrm{UWG}}(t)$ (Eq (2.8)) into this definition yields an integral over the survival window $[t, 1]$. Using a power series expansion on the denominator and applying the definition of the lower incomplete gamma function, $\gamma(a, w) = \int_0^w u^{a-1} e^{-u} du$, the n^{th} conditional moment reduces to:

$$E\left[X^n \mid X > t\right] = \frac{\alpha\left[1 - p\left(1 - e^{-\alpha(-\log t)^\beta}\right)\right]}{1 - e^{-\alpha(-\log t)^\beta}} \sum_{i=0}^{\infty} (i+1) p^i \sum_{m=0}^{i} \binom{i}{m} (-1)^{i-m} \frac{\gamma\left(1 + \frac{n}{\beta}, \alpha(1+i-m)(-\log t)^\beta\right)}{\left[\alpha(1+i-m)\right]^{1+\frac{n}{\beta}}}. \tag{5.23}$$

5.7. Mean residual life

In reliability analysis, the expected additional lifetime that a component will survive, given its survival up to time t, is a pivotal metric known as the mean residual life ($\mu_R(t)$). Mathematically, it is defined as

$$\mu_R(t) = E[X - t \mid X > t] = \frac{1}{S_{\mathrm{UWG}}(t)} \int_t^1 S_{\mathrm{UWG}}(x)\, dx, \quad t > 0. \tag{5.24}$$

Equivalently, $\mu_R(t)$ can be formulated using the total area under the density curve minus the accumulated past lifetime

$$\mu_R(t) = \frac{1}{S_{\mathrm{UWG}}(t)} \left(\int_t^1 x f_{\mathrm{UWG}}(x)\, dx - t \cdot S_{\mathrm{UWG}}(t) \right) = \frac{1}{S_{\mathrm{UWG}}(t)} \left(\mu_1' - I_X(t; 1)\right) - t, \tag{5.25}$$

where $S_{\mathrm{UWG}}(t)$ is the survival function defined in Eq (2.8), μ_1' is the global population mean derived in Eq (5.10), and $I_X(t; 1)$ represents the first incomplete moment evaluated at the truncation threshold t by setting $h = 1$ in Eq (5.21).

5.8. Measures of inequality and uncertainty

To establish the practical utility of the UWG model in socioeconomic studies, risk analysis, and information theory, this subsection derives its implied curves of inequality alongside its uncertainty and entropy profiles.

5.8.1. Lorenz and bonferroni curves

The Lorenz ($L(u)$) and Bonferroni ($B(u)$) curves are standard analytical tools utilized globally to quantify income concentration, wealth inequality, and industrial reliability dispersion. For a specified cumulative population proportion $u = F_{\mathrm{UWG}}(t)$, where $t = F^{-1}_{\mathrm{UWG}}(u)$ is the quantile function from Eq (5.2), these curves are mathematically defined as

$$L(u) = \frac{1}{\mu_1'}\int_0^t x f_{\mathrm{UWG}}(x)\,dx = \frac{I_X(t;1)}{\mu_1'} \quad \text{and} \quad B(u) = \frac{L(u)}{u} = \frac{I_X(t;1)}{u\mu_1'}, \tag{5.26}$$

where $I_X(t;1)$ is the first incomplete moment (Eq (5.21)) and μ_1' is the overall population expectation (Eq (5.10)).

5.8.2. Rényi entropy

The entropy of a random variable measures the inherent uncertainty, information content, or stochastic randomness embedded within a physical or statistical system. The Rényi entropy, parameterized by θ, is defined as

$$I_R(\theta) = \frac{1}{1-\theta}\log\left(\int_0^1 \left[f_{\mathrm{UWG}}(x)\right]^{\theta}\,dx\right), \quad \theta > 0,\ \theta \neq 1. \tag{5.27}$$

By embedding the explicit PDF of the UWG distribution from Eq (2.7) into this formulation, the core integral can be expressed as

$$\int_0^1 \left[f_{\mathrm{UWG}}(x)\right]^{\theta}\,dx = (1-p)^{\theta}\alpha^{\theta}\beta^{\theta}\int_0^1 x^{-\theta}\frac{(-\log x)^{\theta(\beta-1)}e^{-\theta\alpha(-\log x)^{\beta}}}{\left[1-p\left(1-e^{-\alpha(-\log x)^{\beta}}\right)\right]^{2\theta}}\,dx. \tag{5.28}$$

Applying a generalized binomial series expansion to expand the denominator term to the power of -2θ, and using the transform $y = -\log x$, the expression can be integrated analytically using gamma functions. Taking the logarithm and simplifying yields the finalized expression for the Rényi entropy of the UWG model

$$\begin{aligned} I_R(\theta) =& \frac{\theta}{1-\theta}\left[\log(1-p) + \log\alpha + \log\beta\right] \\ &+ \frac{1}{1-\theta}\log\left(\sum_{i=0}^{\infty}\sum_{m=0}^{i}\frac{\Gamma(2\theta+i)p^i\binom{i}{m}(-1)^{i-m}\Gamma\left(\theta(\beta-1)+1\right)}{i!\,\Gamma(2\theta)\left[\alpha(\theta+i-m)-(\theta-1)\right]^{\theta(\beta-1)+1}}\right). \end{aligned} \tag{5.29}$$

5.9. Order statistics

Order statistics are essential in reliability engineering for analyzing parallel, series, and k-out-of-n system configurations. Let $X_1, X_2, \ldots, X_r$ represent a random sample of size r drawn from the UWG distribution, and let $X_{1:r} \leq X_{2:r} \leq \cdots \leq X_{r:r}$ denote their corresponding ordered sequence. The PDF of the i^{th} order statistic, $X_{i:r}$, is conventionally formulated as

$$f_{i:r}(x) = \frac{1}{B(i, r-i+1)} f_{\mathrm{UWG}}(x)\left[F_{\mathrm{UWG}}(x)\right]^{i-1}\left[1 - F_{\mathrm{UWG}}(x)\right]^{r-i}. \tag{5.30}$$

Expanding the survival component via the binomial theorem yields

$$f_{i:r}(x) = \frac{1}{B(i, r-i+1)} \sum_{l=0}^{r-i} \binom{r-i}{l} (-1)^l f_{\mathrm{UWG}}(x) \left[F_{\mathrm{UWG}}(x)\right]^{l+i-1}. \tag{5.31}$$

By substituting the explicit mathematical definitions of $F_{\mathrm{UWG}}(x)$ and $f_{\mathrm{UWG}}(x)$ into Eq (5.31) and simplifying the interacting series expansions, the exact, computationally traceable expression for the PDF of the i^{th} order statistic reduces to

$$f_{i:r}(x) = \frac{(1-p)\alpha\beta}{x \cdot B(i, r-i+1)} \sum_{l=0}^{r-i} \sum_{j=0}^{\infty} \binom{r-i}{l} \binom{-(l+i+1)}{j} (-1)^{l+j} p^j (-\log x)^{\beta-1} e^{-\alpha(l+i+j)(-\log x)^{\beta}}. \tag{5.32}$$

Analogously, the CDF of the i^{th} order statistic, $X_{i:r}$, is expressed using the standard binomial distribution of success probabilities

$$F_{i:r}(x) = \sum_{j=i}^{r} \binom{r}{j} \left[F_{\mathrm{UWG}}(x)\right]^j \left[1 - F_{\mathrm{UWG}}(x)\right]^{r-j} = \sum_{j=i}^{r} \sum_{l=0}^{r-j} \binom{r}{j} \binom{r-j}{l} (-1)^l \left[F_{\mathrm{UWG}}(x)\right]^{j+l}. \tag{5.33}$$

Substituting the structural form of the UWG distribution's CDF yields the following finalized expression

$$F_{i:r}(x) = \sum_{j=i}^{r} \sum_{l=0}^{r-j} \sum_{k=0}^{\infty} \binom{r}{j} \binom{r-j}{l} \binom{-(j+l)}{k} (-1)^{l+k} p^k e^{-\alpha(j+l+k)(-\log x)^{\beta}}. \tag{5.34}$$

Importantly, the structural behavior of the system's operational extremes, namely, the minimum component failure time ($X_{1:r}$, which mimics series architecture) and the maximum component failure time ($X_{r:r}$, which mimics parallel architecture) are readily obtained by setting $i = 1$ and $i = r$ into Eqs (5.32) and (5.34).

6. Maximum likelihood estimation and asymptotic inference

In this section, we address the point and interval estimation of the unknown parameters governing the UWG distribution utilizing the method of MLEs. Let $X_1, X_2, \ldots, X_v$ be a random sample of size v drawn i.i.d. from the UWG distribution given by the PDF in Eq (2.7). Let $\mathbf{\Theta} = (\alpha, \beta, p)^T$ denote the vector of unknown parameters. The total log-likelihood function, $l(\mathbf{\Theta}) = \log L(\alpha, \beta, p; x_i)$, is obtained by taking the natural logarithm of the joint sample density product

$$\begin{aligned} l(\mathbf{\Theta}) =& v\log(1-p) + v\log\alpha + v\log\beta - \sum_{i=1}^{v} \log x_i + (\beta-1)\sum_{i=1}^{v} \log(-\log x_i) \\ & - \alpha \sum_{i=1}^{v} (-\log x_i)^{\beta} - 2\sum_{i=1}^{v} \log\left[1 - p\left(1 - e^{-\alpha(-\log x_i)^{\beta}}\right)\right]. \end{aligned} \tag{6.1}$$

To determine the MLEs, we take the first-order partial derivatives of Eq (6.1) with respect to each component parameter. Setting these score equations equal to zero yields the matching score functions

$$\frac{\partial l(\mathbf{\Theta})}{\partial \alpha} = \frac{v}{\alpha} - \sum_{i=1}^{v} (-\log x_i)^{\beta} + 2p \sum_{i=1}^{v} \frac{(-\log x_i)^{\beta} e^{-\alpha(-\log x_i)^{\beta}}}{1 - p\left(1 - e^{-\alpha(-\log x_i)^{\beta}}\right)} = 0, \tag{6.2}$$

$$
\begin{aligned}
\frac{\partial l(\boldsymbol{\Theta})}{\partial \beta} =& \frac{v}{\beta} + \sum_{i=1}^{v} \log(-\log x_i) - \alpha \sum_{i=1}^{v} (-\log x_i)^{\beta} \log(-\log x_i) \\
&+ 2p\alpha \sum_{i=1}^{v} \frac{(-\log x_i)^{\beta} \log(-\log x_i) e^{-\alpha(-\log x_i)^{\beta}}}{1 - p\left(1 - e^{-\alpha(-\log x_i)^{\beta}}\right)} = 0, \qquad (6.3)\\
\frac{\partial l(\boldsymbol{\Theta})}{\partial p} =& \frac{v}{p-1} + 2 \sum_{i=1}^{v} \frac{1 - e^{-\alpha(-\log x_i)^{\beta}}}{1 - p\left(1 - e^{-\alpha(-\log x_i)^{\beta}}\right)} = 0. \qquad (6.4)
\end{aligned}
$$

The maximum likelihood estimators $\hat{\boldsymbol{\Theta}} = (\hat{\alpha}, \hat{\beta}, \hat{p})^T$ are defined as the simultaneous solutions to the nonlinear system of equations ($\partial l/\partial \alpha = 0, \partial l/\partial \beta = 0, \partial l/\partial p = 0$). Because these equations cannot be solved analytically in closed-form, numerical optimization algorithms such as the Newton–Raphson method or quasi-Newton methods are required. These algorithms can be implemented using standard statistical software packages, such as the `maxLik` or `optim` routines in `R`. For large sample sizes, interval estimation and hypothesis testing are established using the asymptotic distribution of the MLEs. To derive the asymptotic variance–covariance framework, we construct the 3×3 observed Fisher information matrix, denoted $\boldsymbol{J}(\boldsymbol{\Theta})$, which contains the second-order partial derivatives of the log-likelihood function

$$
\boldsymbol{J}(\boldsymbol{\Theta}) = -\begin{pmatrix} \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \alpha^2} & \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \alpha \partial \beta} & \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \alpha \partial p} \\ \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \beta \partial \alpha} & \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \beta^2} & \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \beta \partial p} \\ \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial p \partial \alpha} & \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial p \partial \beta} & \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial p^2} \end{pmatrix}. \qquad (6.5)
$$

For mathematical brevity in writing the second derivatives, let us define the auxiliary functional components $w_i = (-\log x_i)^{\beta}$, $\Omega_i = e^{-\alpha w_i}$, and $\Lambda_i = 1 - p(1 - \Omega_i)$. The components of $\boldsymbol{J}(\boldsymbol{\Theta})$ are derived as follows:

$$
\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \alpha^2} =& -\frac{v}{\alpha^2} - 2p \sum_{i=1}^{v} \frac{w_i^2 \Omega_i \left[\Lambda_i - p\Omega_i\right]}{\Lambda_i^2}, \qquad (6.6)\\
\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \beta^2} =& -\frac{v}{\beta^2} - \alpha \sum_{i=1}^{v} w_i (\log(-\log x_i))^2 \\
&+ 2p\alpha \sum_{i=1}^{v} \frac{w_i (\log(-\log x_i))^2 \Omega_i \left[(1 - \alpha w_i)\Lambda_i + p\alpha w_i \Omega_i\right]}{\Lambda_i^2}, \qquad (6.7)\\
\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial p^2} =& -\frac{v}{(1-p)^2} + 2 \sum_{i=1}^{v} \frac{(1 - \Omega_i)^2}{\Lambda_i^2}, \qquad (6.8)\\
\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \alpha \partial \beta} =& \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \beta \partial \alpha} = -\sum_{i=1}^{v} w_i \log(-\log x_i) \\
&+ 2p \sum_{i=1}^{v} \frac{w_i \log(-\log x_i) \Omega_i \left[(1 - \alpha w_i)\Lambda_i + p\alpha w_i \Omega_i\right]}{\Lambda_i^2}, \qquad (6.9)\\
\frac{\partial^2 l(\boldsymbol{\Theta})}{\partial \alpha \partial p} =& \frac{\partial^2 l(\boldsymbol{\Theta})}{\partial p \partial \alpha} = 2 \sum_{i=1}^{v} \frac{w_i \Omega_i \left[\Lambda_i + p(1 - \Omega_i)\right]}{\Lambda_i^2} = 2 \sum_{i=1}^{v} \frac{w_i \Omega_i}{\Lambda_i^2}, \qquad (6.10)
\end{aligned}
$$

$$\frac{\partial^2 l(\mathbf{\Theta})}{\partial\beta\partial p} = \frac{\partial^2 l(\mathbf{\Theta})}{\partial p\partial\beta} = 2\alpha\sum_{i=1}^{v}\frac{w_i\log(-\log x_i)\Omega_i}{\Lambda_i^2}. \tag{6.11}$$

Under the standard regularity conditions required for maximum likelihood estimators (such as the interior positioning of true parameters within a compact space and the non-singularity of the information matrix), the joint distribution of $\hat{\mathbf{\Theta}}$ exhibits asymptotic multivariate normality as the sample size $v \to \infty$. Mathematically, this limiting distribution is expressed as

$$\sqrt{v}\left(\hat{\mathbf{\Theta}} - \mathbf{\Theta}\right) \xrightarrow{d} N_3\left(\mathbf{0},\, \boldsymbol{I}(\mathbf{\Theta})^{-1}\right), \tag{6.12}$$

where $\boldsymbol{I}(\mathbf{\Theta}) = E[\boldsymbol{J}(\mathbf{\Theta})]$ is the expected Fisher information matrix. Evaluating the exact mathematical expectation of the non–linear ratios in $\boldsymbol{J}(\mathbf{\Theta})$ can be analytically difficult. Therefore, we utilize the consistently estimated observed inverse Fisher information matrix, evaluated at the calculated parameter values $\hat{\mathbf{\Theta}}$

$$\boldsymbol{J}(\hat{\mathbf{\Theta}})^{-1} = \begin{pmatrix} \widehat{\mathrm{Var}}(\hat{\alpha}) & \widehat{\mathrm{Cov}}(\hat{\alpha},\hat{\beta}) & \widehat{\mathrm{Cov}}(\hat{\alpha},\hat{p}) \\ \widehat{\mathrm{Cov}}(\hat{\beta},\hat{\alpha}) & \widehat{\mathrm{Var}}(\hat{\beta}) & \widehat{\mathrm{Cov}}(\hat{\beta},\hat{p}) \\ \widehat{\mathrm{Cov}}(\hat{p},\hat{\alpha}) & \widehat{\mathrm{Cov}}(\hat{p},\hat{\beta}) & \widehat{\mathrm{Var}}(\hat{p}) \end{pmatrix}. \tag{6.13}$$

The asymptotic standard errors (SE) of the estimates are obtained by taking the square root of the diagonal entries of this matrix $SE(\hat{\alpha}) = \sqrt{\widehat{\mathrm{Var}}(\hat{\alpha})}$, $SE(\hat{\beta}) = \sqrt{\widehat{\mathrm{Var}}(\hat{\beta})}$, and $SE(\hat{p}) = \sqrt{\widehat{\mathrm{Var}}(\hat{p})}$. Consequently, we can construct the two-sided $100(1-\tau)\%$ asymptotic confidence intervals (ACIs) for the parameters under the normality assumption. For a significance level $\tau \in (0,1)$, the intervals are calculated as:

$$\hat{\alpha} \pm z_{\frac{\tau}{2}}\sqrt{\widehat{\mathrm{Var}}(\hat{\alpha})}, \quad \hat{\beta} \pm z_{\frac{\tau}{2}}\sqrt{\widehat{\mathrm{Var}}(\hat{\beta})}, \quad \text{and} \quad \hat{p} \pm z_{\frac{\tau}{2}}\sqrt{\widehat{\mathrm{Var}}(\hat{p})}, \tag{6.14}$$

where $z_{\frac{\tau}{2}}$ represents the upper $(\frac{\tau}{2})^{\text{th}}$ standard normal percentile, satisfying $P(Z > z_{\frac{\tau}{2}}) = \frac{\tau}{2}$. Under the standard regularity conditions for MLE, the MLEs of the parameter vector $\mathbf{\Theta}$ are consistent, asymptotically normal, and asymptotically efficient. Specifically, as $n \to \infty$

$$\sqrt{n}\left(\hat{\mathbf{\Theta}} - \mathbf{\Theta}\right) \xrightarrow{d} N\left(\mathbf{0},\, J^{-1}(\mathbf{\Theta})\right), \tag{6.15}$$

where $J(\mathbf{\Theta})$ denotes the Fisher information matrix. These asymptotic properties hold provided that the usual regularity conditions for MLE are satisfied, including parameter identifiability, differentiability of the log-likelihood function, and the existence of a positive definite Fisher information matrix (see, e.g., [29, 30]).

7. Simulation analysis and finite-sample evaluation

From a computational perspective, the MLE procedure is straightforward to implement using standard numerical optimization techniques. Throughout this study, the optimization routine converged reliably for both the simulated and real Datasets. To enhance numerical stability, multiple initial values were considered, with a global optimization algorithm providing suitable starting values that were subsequently refined using a local optimization routine. In practice, the use of reasonable initial values together with appropriate convergence tolerances is recommended to obtain stable and reproducible parameter estimates.

To further assess the performance of the proposed estimation procedure, we conduct a comprehensive Monte Carlo simulation study to evaluate the finite-sample behavior of the MLEs for the UWG distribution. The study examines the consistency and estimation accuracy of the MLEs under different sample sizes and parameter settings, thereby providing numerical evidence for the effectiveness and practical applicability of the estimation framework developed in Section 6. To simulate pseudo-random samples from the UWG distribution, we implement the inversion method (also known as the inverse transform sampling technique). This technique leverages the fundamental probability principle that if U is a continuous uniform random variable defined on the open standard interval $(0,1)$, then the transformed random variable $X = F_{\mathrm{UWG}}^{-1}(U)$ precisely follows the target distribution governed by the CDF $F_{\mathrm{UWG}}(x)$. By equating the structural CDF expression from Eq (2.6) to a uniform variate $u \in (0,1)$ and solving for x, the inverse transform algorithm generates random variates explicitly through the following closed-form expression:

$$x = \exp\left\{-\left[-\frac{1}{\alpha}\log\left(\frac{u(1-p)}{1-up}\right)\right]^{\frac{1}{\beta}}\right\}. \tag{7.1}$$

The step-by-step pseudo-code for the simulation routine is organized as follows:

(1) Specify the target vector of parameter values $\mathbf{\Theta} = (\alpha, \beta, p)^T$.
(2) Generate a random realization u from the continuous uniform distribution, such that $u \sim \mathrm{Uniform}(0,1)$.
(3) Compute the corresponding realization x via the analytical transformation defined in Eq (7.1).
(4) Repeat steps 2 and 3 sequentially a total of v times to yield a simulated random sample of size v.

To evaluate the asymptotic behavior and finite-sample properties of the calculated estimators, we investigate five distinct parameter combinations selected to represent diverse distribution shapes (including right-skewed, left-skewed, J-shaped, and non–monotonic hazard configurations). These configurations are defined as follows

- **Set 1:** $\alpha = 2.5, \beta = 3.0, p = 0.7$ (highly left-skewed, unimodal profile);
- **Set 2:** $\alpha = 0.9, \beta = 0.6, p = 0.3$ (reversed-J shaped density profile);
- **Set 3:** $\alpha = 7.0, \beta = 3.0, p = 0.1$ (highly compressed, symmetric-like boundary profile);
- **Set 4:** $\alpha = 6.0, \beta = 1.5, p = 0.9$ (heavy-tailed distribution near the upper bound);
- **Set 5:** $\alpha = 4.5, \beta = 1.5, p = 0.5$ (moderately skewed, flexible unimodal profile).

For each parameter set, random samples are generated across six distinct sample sizes: $v \in \{30, 50, 100, 150, 200, 500\}$, representing small, medium, and large sample environments. To ensure numerical stability and avoid localized trapping during optimization, the log-likelihood maximization is performed using the differential evolution global optimization algorithm via the `DEoptim` package in `R`. This optimization loop is executed over $M = 1000$ independent Monte Carlo replications for each unique experimental cell. Let $\theta \in \{\alpha, \beta, p\}$ denote a generic target parameter, and let $\hat{\theta}_j$ represent its corresponding maximum likelihood estimate obtained during the j^{th} simulation replication. The finite-sample performance is quantified by computing the average MLE, the average empirical bias, and the mean squared error (MSE), which are calculated using the following mathematical formulas

$$\text{Average MLE}(\hat{\theta}) = \frac{1}{M}\sum_{j=1}^{M}\hat{\theta}_j, \tag{7.2}$$

$$\text{Bias}(\hat{\theta}) = \frac{1}{M}\sum_{j=1}^{M}(\hat{\theta}_j - \theta), \tag{7.3}$$

$$\text{MSE}(\hat{\theta}) = \frac{1}{M}\sum_{j=1}^{M}(\hat{\theta}_j - \theta)^2. \tag{7.4}$$

The complete numerical findings generated from our Monte Carlo simulation study are organized in Table 3. A systematic analysis of the simulation metrics presented in Table 3 reveals consistent behavior across the tested parameter variations. As the sample size v increases from 30 to 500, the average maximum likelihood estimates consistently converge toward the pre–specified true target values. This convergence pattern is evident across all five simulation configurations, confirming that the numerical optimization successfully identifies the global maximum of the multi–parameter likelihood surface. The behavior of the estimated bias provides further evidence of the properties of the estimators. For small sample sizes ($v = 30$), moderate bias is observed due to the non–linear structure of the score functions. For example, in Set 1, the scale parameter α starts with a bias of 0.1515, and the shape parameter β exhibits a positive bias of 0.4063. However, as the sample size increases to $v = 150$ and eventually to $v = 500$, these biases decrease significantly, approaching zero (bias($\hat{\alpha}$) = 0.0493 and bias($\hat{\beta}$) = 0.0271 at $v = 500$). This pattern is consistent across all parameter sets, demonstrating that the estimators are asymptotically unbiased. Finally, the MSE values show a monotonic decline as a function of sample size v. In Set 4, for instance, the MSE for the scale parameter α drops from 3.4449 ($v = 30$) to 2.0924 ($v = 100$), and reaches 0.6438 at $v = 500$. Similarly, the MSE for the compounding parameter p drops from 0.2947 to 0.0207. These simulation findings confirm that the maximum likelihood method provides stable, consistent, and asymptotically efficient parameter estimates for the UWG distribution, justifying its application to real-world Datasets.

Furthermore, Table 4 presents the coverage probabilities (CP), average lengths (AL), and the corresponding 95% confidence intervals (95% CI) associated with the maximum likelihood estimates reported in Table 3. It can be observed that the coverage probabilities for all the parameters are consistently high, with most values exceeding 0.90 and some approaching 1. This indicates that the constructed confidence intervals successfully capture the true parameter values with high probability, demonstrating the reliability and accuracy of the estimation procedure.

In addition, the average lengths of the confidence intervals are reasonably small and decrease as the sample size increases, indicating the improved precision of the parameter estimates for larger samples. Narrower confidence intervals suggest greater stability and efficiency of the estimators obtained through the MLE method. The overall behavior of the coverage probabilities, together with the acceptable interval lengths, confirms that the MLE method performs adequately and provides reliable parameter estimation for the proposed distribution.

Table 3. Simulation results for various schemes.

		MLE			bias			MSE		
Sets	v	α	β	p	α	β	p	α	β	p
Set 1	30	2.5151	3.4063	0.4339	0.1515	0.4063	-0.2661	1.8296	1.2570	0.2232
	50	2.6356	3.2500	0.4824	0.1356	0.2500	-0.2176	2.0083	0.9903	0.1898
	100	2.6305	3.1185	0.5484	0.1305	0.1185	-0.1516	1.3288	0.6769	0.1422
	150	2.6115	3.0743	0.5730	0.1115	0.0743	-0.1270	1.0661	0.5535	0.1213
	200	2.5849	3.0674	0.5908	0.0849	0.0674	-0.1092	0.8604	0.4505	0.0999
	500	2.5493	3.0271	0.6439	0.0493	0.0271	-0.0561	0.4478	0.2398	0.0468
Set 2	30	1.3400	0.5819	0.3452	0.4400	-0.0181	0.0452	1.4830	0.0241	0.1429
	50	1.3031	0.5689	0.3586	0.4031	-0.0311	0.0586	1.1922	0.0205	0.1398
	100	1.2015	0.5759	0.3332	0.3015	-0.0241	0.0332	0.8712	0.0151	0.1151
	150	1.1016	0.5819	0.3128	0.2016	-0.0181	0.0128	0.4723	0.0109	0.1033
	200	1.0710	0.5854	0.3090	0.1710	-0.0146	0.0090	0.4106	0.0091	0.0922
	500	0.9685	0.5952	0.2893	0.0685	-0.0048	-0.0107	0.1486	0.0042	0.0611
Set 3	30	7.9886	2.7917	0.2886	0.9886	-0.2083	0.1886	3.9617	0.4762	0.1648
	50	7.8816	2.8041	0.2849	0.8816	-0.1959	0.1849	2.9132	0.3920	0.1540
	100	7.5840	2.8315	0.2397	0.5840	-0.1685	0.1397	1.8923	0.3045	0.1197
	150	7.4787	2.8295	0.2463	0.4787	-0.1705	0.1463	1.3585	0.2492	0.1137
	200	7.3898	2.8869	0.2069	0.3898	-0.1131	0.1069	1.0094	0.1541	0.0829
	500	7.1872	2.9309	0.1714	0.1872	-0.0691	0.0714	0.3797	0.0601	0.0501
Set 4	30	5.5637	2.0712	0.5373	-0.4363	0.5712	-0.3627	3.4449	0.9002	0.2947
	50	5.5447	1.9449	0.6001	-0.4553	0.4449	-0.2999	2.8210	0.6757	0.2358
	100	5.6850	1.7911	0.6918	-0.3150	0.2911	-0.2082	2.0924	0.4295	0.1503
	150	5.7385	1.7201	0.7340	-0.2615	0.2201	-0.1660	1.7511	0.3344	0.1156
	200	5.8313	1.6701	0.7688	-0.1687	0.1701	-0.1312	1.4669	0.2820	0.0889
	500	5.9259	1.5790	0.8490	-0.0741	0.0790	-0.0510	0.6438	0.1120	0.0207
Set 5	30	5.1296	1.5634	0.3864	0.6296	0.0634	-0.1136	3.1612	0.2047	0.1581
	50	4.9018	1.5111	0.4019	0.4018	0.0111	-0.0981	2.0931	0.1493	0.1466
	100	4.8449	1.4630	0.4372	0.3449	-0.0370	-0.0628	1.6549	0.1234	0.1336
	150	4.7069	1.4936	0.4268	0.2069	-0.0064	-0.0732	1.1338	0.0907	0.1130
	200	4.6666	1.4790	0.4338	0.1666	-0.0210	-0.0662	0.9762	0.0802	0.1051
	500	4.5300	1.4991	0.4512	0.0300	-0.0009	-0.0488	0.3992	0.0362	0.0616

Table 4. Coverage probability, average length, and 95% confidence intervals of the MLEs.

			CP			AL			95%	CI
Sets	v	α	β	p	α	β	p	α	β	p
Set 1	30	0.9400	0.9680	1.0000	5.3023	4.3949	1.8519	(-0.1360 , 5.1663)	(1.2089 , 5.6038)	(-0.4920 , 1.3599)
	50	0.9360	0.9660	1.0000	5.5553	3.9009	1.7078	(-0.1420 , 5.4132)	(1.2995 , 5.2004)	(-0.3715 , 1.3363)
	100	0.9350	0.9670	1.0000	4.5187	3.2252	1.4780	(0.3711 , 4.8899)	(1.5059 , 4.7311)	(-0.1906 , 1.2874)
	150	0.9460	0.9610	0.8590	4.0475	2.9165	1.3651	(0.5878 , 4.6352)	(1.6161 ,4.5326)	(-0.1095 ,1.2556)
	200	0.9450	0.9520	0.8830	3.6360	2.6312	1.2388	(0.7669 , 4.4029)	(1.7518 , 4.3830)	(-0.0286 , 1.2102)
	500	0.9460	0.9390	0.9300	2.6231	1.9195	0.8483	(1.2377 , 3.8608)	(2.0674 , 3.9868)	(0.2198 , 1.0680)
Set 2	30	0.9370	0.9380	1.0000	4.7737	0.6089	1.4816	(-1.0468 ,3.7268)	(0.2775 ,0.8864)	(-0.3956 ,1.0860)
	50	0.9400	0.9310	1.0000	4.2801	0.5618	1.4657	(-0.8370 , 3.4431)	(0.2880 ,0.8498)	(-0.3743 ,1.0914)
	100	0.9390	0.9240	0.9620	3.6588	0.4817	1.3297	(-0.627,9 3.0309)	(0.3351, 0.8168)	(-0.331,6 0.9980)
	150	0.9440	0.9410	0.9640	2.6941	0.4084	1.2597	(-0.245,4 2.4486)	(0.3777, 0.7861)	(-0.317,0 0.9427)
	200	0.9580	0.9540	0.9620	2.5120	0.3749	1.1900	(-0.185,0 2.3270)	(0.397,9 0.7728)	(-0.286,0 0.9040)
	500	0.9790	0.9650	0.9800	1.5112	0.2555	0.9692	(0.2129, 1.7241)	(0.4675, 0.7230)	(-0.195,3 0.7739)
Set 3	30	1.0000	0.9350	0.9210	7.8024	2.7051	1.5914	(4.0875, 11.8898)	(1.4392, 4.1442)	(-0.507,1 1.0843)
	50	1.0000	0.9280	0.9040	6.6907	2.4544	1.5386	(4.5363, 11.2269)	(1.5769, 4.0313)	(-0.484,4 1.0542)
	100	0.9140	0.9300	0.9050	5.3924	2.1631	1.3565	(4.8878, 10.2802)	(1.7499, 3.9130)	(-0.438,6 0.9179)
	150	0.9310	0.9340	0.9100	4.5689	1.9570	1.3221	(5.1942, 9.7631)	(1.8510, 3.8080)	(-0.414,7 0.9073)
	200	0.9420	0.9520	0.9120	3.9383	1.5391	1.1286	(5.4207, 9.3590)	(2.1173, 3.6564)	(-0.357,4 0.7712)
	500	0.9340	0.9340	0.9140	2.4156	0.9612	0.8772	(5.9794, 8.3950)	(2.4503, 3.4115)	(-0.267,2 0.6100)
Set 4	30	0.9740	0.9480	1.0000	7.2757	3.7193	2.1280	(1.9259, 9.2016)	(0.2116, 3.9308)	(-0.526,6 1.6013)
	50	0.9750	0.9530	1.0000	6.5839	3.2223	1.9034	(2.2528, 8.8367)	(0.3337, 3.5560)	(-0.351,6 1.5518)
	100	0.9700	0.9410	0.8690	5.6703	2.5689	1.5196	(2.8498, 8.5201)	(0.5067, 3.0756)	(-0.068,0 1.4516)
	150	0.9610	0.9360	0.8890	5.1873	2.2669	1.3329	(3.1449, 8.3321)	(0.5866, 2.8535)	(0.0676, 1.4005)
	200	0.9540	0.9180	0.9030	4.7477	2.0817	1.1689	(3.4575 , 8.2052)	(0.6292 , 2.7109)	(0.1844 , 1.3532)
	500	0.9450	0.9410	0.9410	3.1453	1.3116	0.5636	(4.3533 , 7.4986)	(0.9232 , 2.2348)	(0.5672 , 1.1308)
Set 5	30	0.9160	0.9500	1.0000	6.9697	1.7736	1.5588	(1.6447 , 8.6144)	(0.6766 , 2.4502)	(-0.3930, 1.1658)
	50	0.9410	0.9390	1.0000	5.6713	1.5148	1.5010	(2.0661 , 7.7374)	(0.7537 , 2.2685)	(-0.3486, 1.1524)
	100	0.9350	0.9340	1.0000	5.0429	1.3768	1.4329	(2.3235 , 7.3663)	(0.7746 , 2.1514)	(-0.2792, 1.1537)
	150	0.9380	0.9380	1.0000	4.1741	1.1806	1.3180	(2.6198 , 6.7939)	(0.9033 , 2.0839)	(-0.2322, 1.0858)
	200	0.9430	0.9480	1.0000	3.8731	1.1099	1.2707	(2.7301 , 6.6032)	(0.9240 , 2.0340)	(-0.2016, 1.0691)
	500	0.9600	0.9650	0.9170	2.4768	0.7463	0.9731	(3.2916 , 5.7684)	(1.1260 , 1.8723)	(-0.0354, 0.9378)

8. Data modeling

In this section, we analyze two different real-world data sets to assess the fitting performance and practical usefulness of the proposed distribution. To demonstrate its efficiency and flexibility, we also carry out a comparative study with several well-established distributions for modelling unit interval data. The selected benchmark models represent a range of commonly used bounded distributions with different structural characteristics and degrees of flexibility, thereby providing a meaningful basis for evaluating the performance of the proposed model. These include the unit-Weibull (UW) [28], unit modified Weibull (UMW) [31], power upper truncated Weibull (PUTW) [32], unit inverse exponentiated Weibull (UIEW) [33], and Beta (B) [34] distributions. This comparison is intended to evaluate how well the proposed model performs relative to other established models for unit interval data. For each fitted model, we calculate a range of goodness-of-fit measures and model selection criteria in order to provide a comprehensive comparison. These include the negative log-likelihood ($-\ell$), Akaike information criterion (AIC), corrected Akaike information criterion (AICC), Bayesian information criterion (BIC), and Hannan–Quinn information criterion (HQIC). In addition, we compute the Kolmogorov–Smirnov (K–S) statistic along with its associated p-value to further assess the adequacy of each model in capturing the underlying distribution of the observed data. The comparative evaluation is based on these criteria, where models with smaller values of $-\ell$, AIC, AICC, BIC, HQIC, and the K–S statistic are regarded as providing a better fit to the data. Furthermore, a higher p-value for the K–S test generally indicates stronger agreement between the fitted model and the empirical data. Therefore, the model that consistently achieves the most favorable values across these measures is considered to offer the best overall fit.

8.1. Application I

The first application considers a real-life data comprising the daily COVID-19 recovery rates recorded in Spain over the period from 3 March to 7 May 2020. This data, which is publicly available in [35], consists of 67 observations bounded within the unit interval, making it particularly suitable for analysis using unit distributions. The complete dataset is presented in Table 5. Prior to model fitting, a preliminary diagnostic analysis was conducted to examine the underlying hazard structure of the dataset. As illustrated in Figure 3, both the total time on test (TTT) plot and the empirical hazard rate plot consistently exhibit a monotonically increasing failure rate behavior. Specifically, the concave shape of the TTT plot provides strong graphical evidence of an increasing HRF, which is further corroborated by the upward trend observed in the empirical hazard rate plot. This increasing hazard structure implies that the probability of an event occurring grows over time, a characteristic that is inherently accommodated by the UWG distribution. Consequently, these diagnostic results suggest that the UWG distribution provides a good descriptive fit to the present Dataset, demonstrating its flexibility for modeling the observed data.

Table 5. COVID-19 recovery rates in Spain (3 March – 7 May 2020).

0.6670,	0.5000,	0.5000,	0.4286,	0.7500,	0.6531,	0.5161,	0.7895,	0.7689,
0.6873,	0.5200,	0.7251,	0.6375,	0.6078,	0.6289,	0.5712,	0.5923,	0.6061,
0.5924,	0.5921,	0.5592,	0.5954,	0.6164,	0.6455,	0.6725,	0.6838,	0.6850,
0.6947,	0.7210,	0.7315,	0.7412,	0.7508,	0.7519,	0.7547,	0.7645,	0.7715,
0.7759,	0.7807,	0.7838,	0.7847,	0.7871,	0.7902,	0.7934,	0.7913,	0.7962,
0.7971,	0.7977,	0.8007,	0.8038,	0.8289,	0.8322,	0.8354,	0.8371,	0.8387,
0.8456,	0.8490,	0.8535,	0.8547,	0.8564,	0.8580,	0.8604,	0.8628,	0.6586,
0.7070,	0.7963,	0.8516						

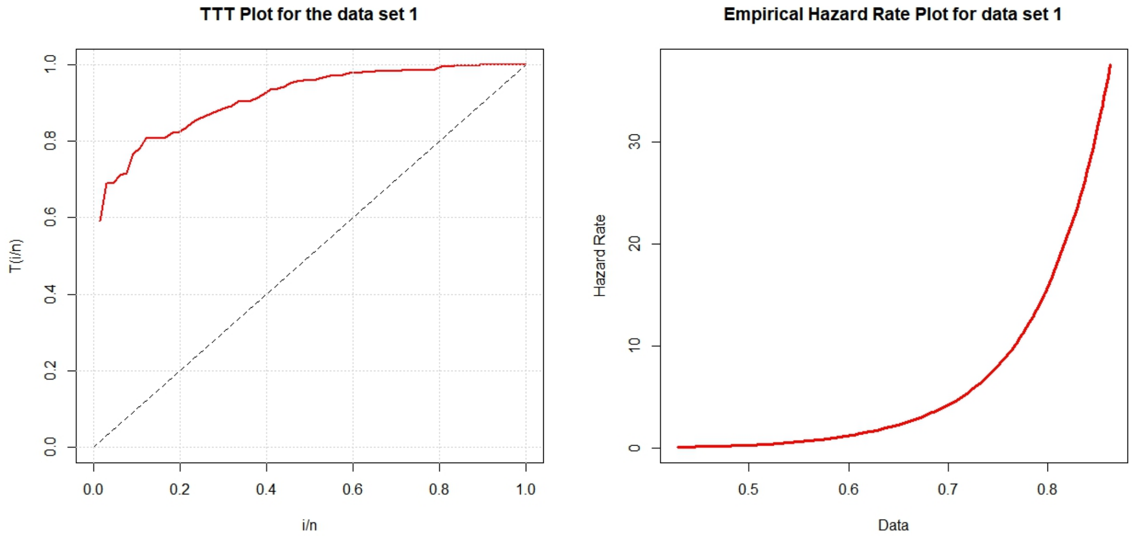

Figure 3. Preliminary diagnostic plots for Dataset I.

The MLEs of the model parameters together with their corresponding standard errors (SEs) for the UWG distribution and all competing distributions are reported in Table 6. The goodness-of-fit and model selection criteria, comprising the $-\ell$, AIC, AICC, BIC, HQIC, K–S test statistic, and its associated p-value, are summarized in Table 7. Additionally, Table 8 presents a detailed comparison between the empirical and theoretical descriptive statistics of the proposed UWG model for Dataset I.

Table 6. The MLEs (SEs) for Dataset I.

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{p}$
UWG	8.6456 (3.0976)	2.2312 (0.4318)	0.0015 (17.4251)
UW	5.1710 (1.6975)	1.8511 (0.2036)	—
PUTW	2.3185 (1.1114)	3.4566 (1.4636)	3.0427 (1.9435)
UIEW	0.9538 (0.8522)	0.6101 (0.2789)	3.0000 (2.2013)
UMW	0.2519 (0.9457)	8.3411 (0.0279)	2.2751 (1.7391)
B	6.4320 (2.2291)	2.6945 (0.8269)	—

Table 7. Goodness-of-fit for Dataset I.

Distribution	ℓ	AIC	AICC	BIC	HQIC	K–S	p-value
UWG	53.9622	−101.9243	−101.5372	−95.3553	−99.3286	0.1307	0.2098
UW	50.6363	−97.2726	−97.0821	−92.8933	−95.5421	0.1428	0.1354
PUTW	52.4577	−98.9153	−98.5283	−92.3464	−96.3196	1.1823	<0.05
UIEW	49.2909	−92.5817	−92.1946	−86.0128	−89.9860	0.2042	0.0082
UMW	52.7464	−99.4927	−99.1056	−92.9238	−96.8970	0.1345	0.1834
B	51.3017	−98.6034	−98.4129	−94.2241	−96.8729	0.1355	0.1774

An examination of Table 8 reveals that the theoretical descriptive statistics derived from the proposed UWG distribution are in close agreement with the corresponding empirical statistics computed from the observed data. Specifically, the mean, variance, kurtosis, and quartile estimates exhibit negligible discrepancies between their empirical and theoretical counterparts, thereby confirming that the model accurately characterizes the central tendency, dispersion, tail behavior, and overall distributional shape of the Dataset. A moderate departure is, however, observed in the skewness measure, where the empirical data exhibit a considerably stronger negative skewness (-0.7049) relative to the theoretical value (-0.1032) predicted by the fitted model. This discrepancy suggests that the observed data display a more pronounced left-tail asymmetry than that captured by the UWG distribution. Nevertheless, the overall coherence between the empirical and theoretical statistics confirms that the proposed model provides an adequate and statistically satisfactory characterization of Dataset I.

Table 8. Descriptive statistics for Dataset I.

No.	Statistic	Empirical	Theoretical
1	Mean	0.7239	0.7218
2	Variance	0.0118	0.0122
3	Skewness	−0.7049	−0.1032
4	Kurtosis	2.6021	2.5544
5	First quartile	0.6474	0.6448
6	Median	0.7533	0.7224
7	Third quartile	0.7975	0.8019

As shown in Table 6, it is observed that the parameter p of the UWG distribution is associated with a relatively large standard error. This suggests that, for the present Dataset, the likelihood function is relatively insensitive to variations in this parameter, resulting in greater uncertainty in its estimation. Nevertheless, the remaining parameter estimates are reasonably precise, and the overall goodness-of-fit measures indicate that the UWG distribution continues to provide an adequate and flexible fit to the observed data.

An inspection of Table 7 reveals that the UWG distribution simultaneously achieves the lowest values of $-\ell$, AIC, AICC, BIC, and HQIC, as well as the smallest K–S test statistic and the largest associated p-value among all competing models considered in this study. These results collectively provide compelling statistical evidence that the UWG distribution delivers a superior fit to Dataset I relative to all other models. Accordingly, the UWG distribution is identified as the most appropriate and parsimonious model for characterizing the probabilistic structure of Dataset I.

Figure 4 presents the fitted probability density curves of the UWG distribution and all competing models superimposed on the empirical histogram of Dataset I. It is evident that the estimated density of the UWG distribution conforms closely to the overall pattern and modal structure of the empirical histogram, outperforming the competing models in capturing the shape and concentration of the observed data. This visual agreement between the fitted density and the empirical distribution further reinforces the adequacy and flexibility of the proposed model.

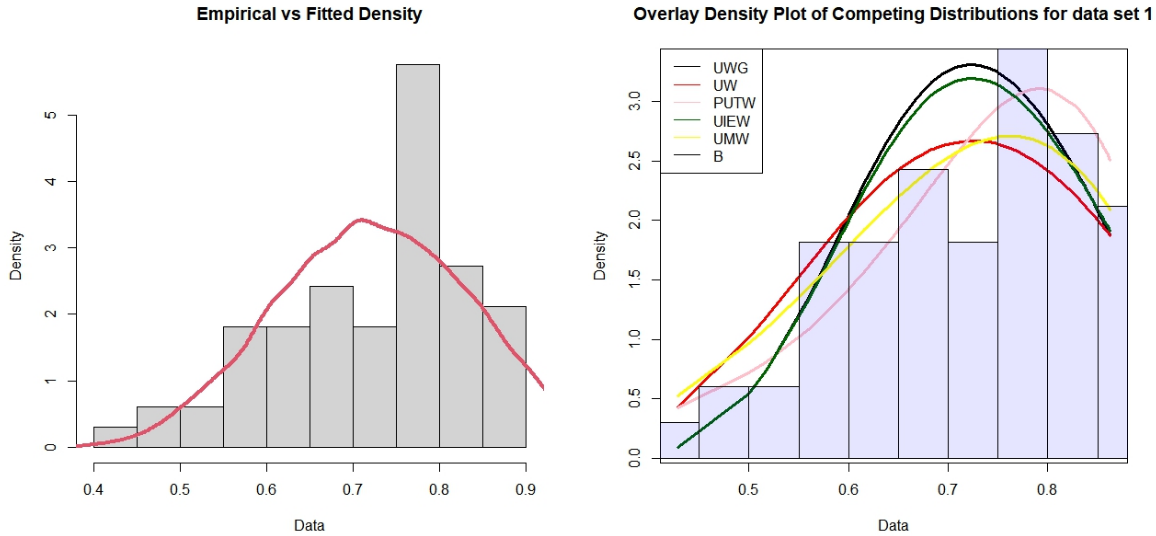

Figure 4. Empirical histogram with fitted density curves for Dataset I.

Figure 5 presents the violin plot with the embedded boxplot and the quantile–quantile (Q–Q) plot for Dataset I under the UWG distribution. The violin plot reveals that the bulk of the observations are concentrated within the interval [0.70, 0.85], with a discernible asymmetry towards the upper-middle region of the distribution. The embedded boxplot indicates a median value near the center of the interquartile range, a moderate degree of spread, and the absence of severe outliers, suggesting a well-behaved and relatively symmetric core structure in the data. The Q–Q plot demonstrates that the observed quantiles align closely with the theoretical quantiles of the UWG distribution along the reference diagonal, with only minor deviations confined to the tail regions. Such tail departures are commonly encountered in empirical Datasets and do not undermine the overall adequacy of the fit. Taken together, these graphical diagnostics confirm that the UWG distribution successfully captures the principal features and underlying stochastic behavior of Dataset I.

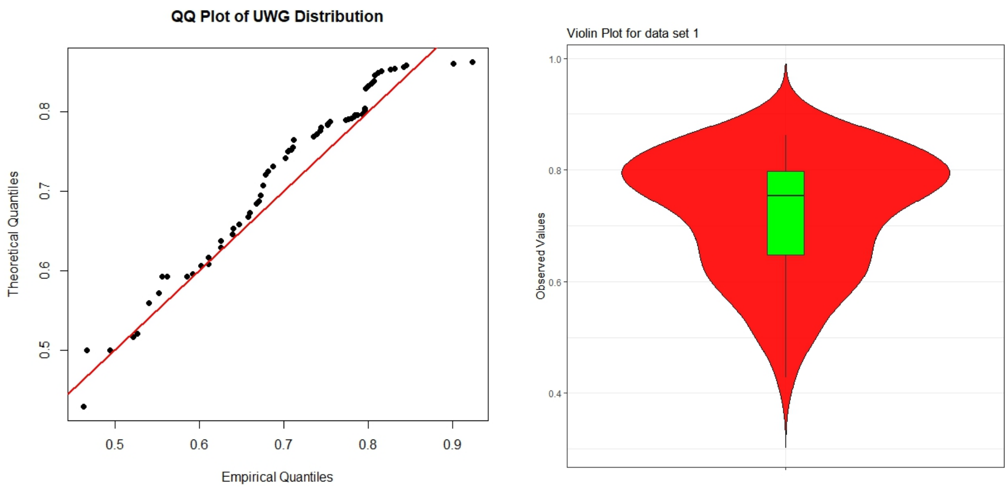

Figure 5. Distributional diagnostic plots for Dataset I.

Figure 6 displays the empirical and fitted CDFs alongside the probability–probability (P–P) plot for Dataset I. The fitted CDF of the UWG distribution tracks the empirical CDF with a high degree of precision throughout the entire range of the data, reflecting the model's ability to reproduce the observed cumulative behavior accurately. The P–P plot corroborates this finding, as the plotted points lie in close proximity to the 45° reference line with negligible systematic deviations, thereby providing further graphical confirmation of the goodness-of-fit of the proposed model.

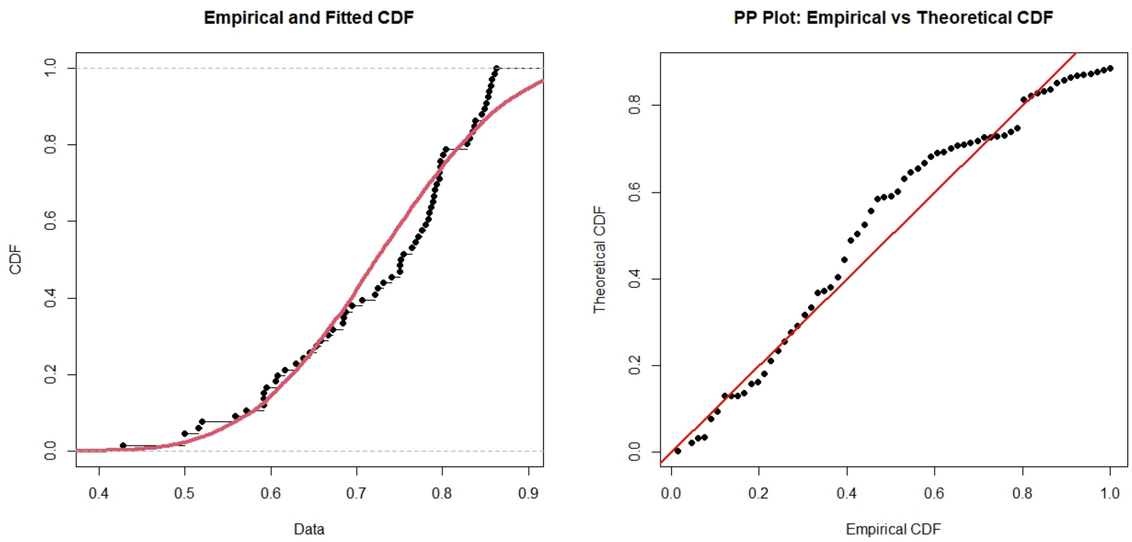

Figure 6. Cumulative distribution diagnostic plots for Dataset I under the UWG distribution.

Figure 7 presents the profile log-likelihood functions for each of the three model parameters α, β, and p of the UWG distribution fitted to Dataset I. All three profile log-likelihood curves exhibit smooth, well-behaved unimodal shapes, which is a desirable property confirming the identifiability and numerical stability of the MLEs. The profiles for α and β display distinct and sharply defined peaks at their respective MLE values, providing strong evidence for the uniqueness and precision of these estimates. For the parameter p, the profile log-likelihood reaches its maximum in the vicinity of the estimated value, which lies near the lower boundary of the admissible parameter space, and subsequently decreases monotonically with increasing values of p. This behavior indicates that the data strongly favour small values of p, while still admitting a well-defined and unique maximum likelihood estimate. The absence of multiple local maxima across all three profile log-likelihood functions, combined with their smooth and regular structure, provides compelling evidence that the model parameters are well-identified and that the optimization algorithm converged reliably to the global maximum. These results collectively affirm the reliability, stability, and robustness of the maximum likelihood estimation procedure for the proposed UWG distribution.

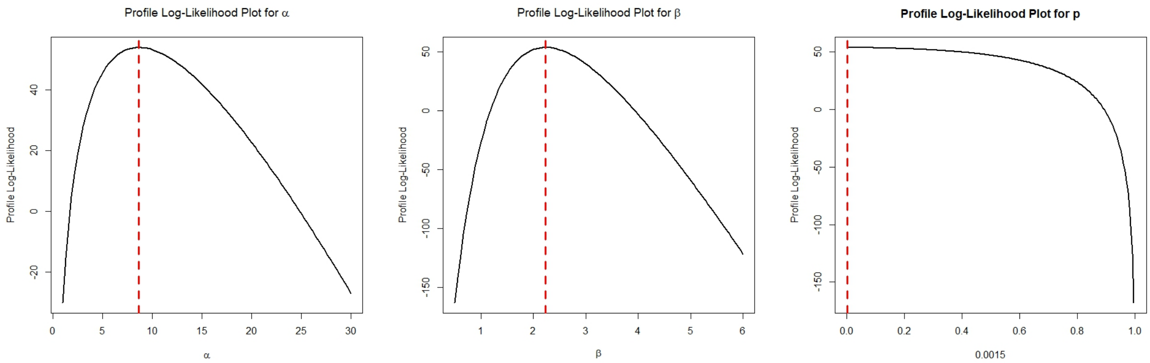

Figure 7. Profile log-likelihood functions of the UWG parameters for Dataset I.

8.2. *Application II*

The second application considers a data consisting of 48 rock samples extracted from an oil reservoir, as originally reported in [36]. The porosity measurements of these rock samples are bounded within the unit interval, rendering this data particularly well-suited for analysis using unit distributions. The complete data is presented in Table 9.

Table 9. Porosity measurements of 48 rock samples from an oil reservoir [36].

0.0903,	0.2036,	0.2043,	0.2808,	0.1976,	0.3286,	0.1486,	0.1623,	0.2627,
0.1794,	0.3266,	0.2300,	0.1833,	0.1509,	0.2000,	0.1918,	0.1541,	0.4641,
0.1170,	0.1481,	0.1448,	0.1330,	0.2760,	0.4204,	0.1224,	0.2285,	0.1138,
0.2252,	0.1769,	0.2007,	0.1670,	0.2316,	0.2910,	0.3412,	0.4387,	0.2626,
0.1896,	0.1725,	0.2400,	0.3116,	0.1635,	0.1824,	0.1641,	0.1534,	0.1618,
0.2760,	0.2538,	0.2004						

As a preliminary step prior to model fitting, diagnostic analyses were performed to investigate the underlying hazard structure of the dataset. Figure 8 presents the TTT plot and the empirical hazard rate plot for Dataset II. The concave shape of the TTT plot provides clear graphical evidence of a monotonically increasing HRF, a conclusion that is further corroborated by the upward trend exhibited in the empirical hazard rate plot. This increasing hazard structure indicates that the risk of failure grows progressively over time, a characteristic that is inherently accommodated by the UWG distribution. These diagnostic results therefore provide strong justification for the adoption of the UWG distribution as a suitable and flexible model for the analysis of Dataset II.

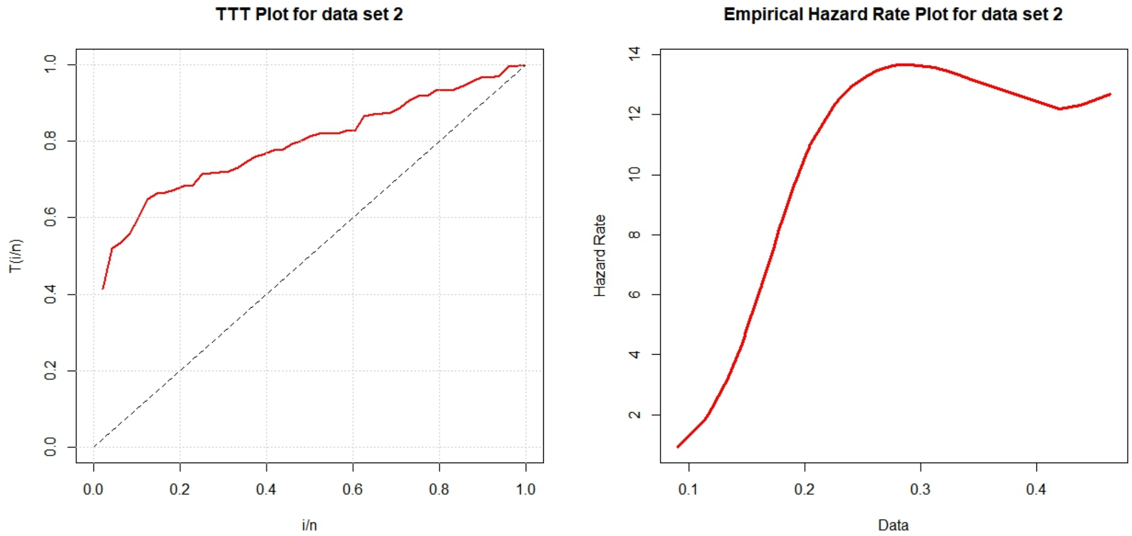

Figure 8. Preliminary diagnostic plots for Dataset II.

The MLEs of the model parameters together with their corresponding SEs for the UWG distribution and all competing distributions are reported in Table 10. The goodness-of-fit and model selection criteria, comprising the $-\ell$, AIC, AICC, BIC, HQIC, K–S test statistic, and its associated p-value, are summarized in Table 11. Furthermore, Table 12 provides a systematic comparison between the empirical and theoretical descriptive statistics of the proposed UWG model for Dataset II.

Table 10. The MLEs (SEs) for Dataset II.

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{p}$
UWG	0.1045 (0.2104)	4.6399 (1.9493)	0.3834 (1.1031)
UW	0.2049 (0.0234)	3.4813 (0.5762)	—
PUTW	7.9642 (1.5772)	0.9291 (0.3663)	6.9461 (26.7275)
UIEW	1.6440 (0.1214)	5.5002 (1.1522)	10.0930 (4.3606)
UMW	0.0010 (0.0837)	0.2180 (2.7461)	3.7310 (0.9647)
B	3.0513 (1.1813)	9.8066 (4.3465)	—

Table 11. Goodness-of-fit for Dataset II.

Distribution	ℓ	AIC	AICC	BIC	HQIC	K–S	p-value
UWG	58.4449	−110.8898	−110.3444	−105.2762	−108.7684	0.0751	0.9494
UW	52.2077	−100.4155	−100.1488	−96.6731	−99.0012	0.2455	0.0061
PUTW	55.8749	−105.7497	−105.2043	−100.1361	−103.6283	1.0925	<0.05
UIEW	54.0502	−102.1004	−101.5549	−96.4868	−99.9790	0.1527	0.2129
UMW	49.4605	−92.9211	−92.3756	−87.3075	−90.7997	0.1942	0.0535
B	50.3006	−96.6011	−96.3345	−92.8587	−95.1869	0.1585	0.1794

Table 12. Descriptive statistics for Dataset II.

No.	Statistic	Empirical	Theoretical
1	Mean	0.2181	0.2171
2	Variance	0.0061	0.0071
3	Skewness	1.1691	1.5451
4	Kurtosis	4.1091	6.6701
5	First quartile	0.1621	0.1591
6	Median	0.1981	0.1981
7	Third quartile	0.2621	0.2551

An examination of Table 12 indicates that the theoretical descriptive statistics derived from the proposed UWG distribution are generally in close agreement with the corresponding empirical statistics computed from the observed data. In particular, the mean, median, variance, and quartile estimates exhibit negligible discrepancies between their empirical and theoretical counterparts, confirming that the model accurately characterizes the central tendency, dispersion, and overall distributional behavior of Dataset II. Notably, the theoretical and empirical median values are in perfect agreement at 0.1981, which further underscores the model's ability to reproduce the central characteristics of the data with high fidelity. However, more pronounced deviations are observed in the higher-order moments, specifically in the skewness and kurtosis measures. The theoretical skewness (1.5451) exceeds the empirical value (1.1691), and the theoretical kurtosis (6.6701) is considerably larger than its empirical counterpart (4.1091). These discrepancies indicate that the UWG distribution predicts a more positively skewed distribution with heavier tails relative to what is observed in the data. Such behavior is not uncommon when fitting parametric models to small–to–moderate sample sizes, where empirical higher-order moments tend to be unstable estimators. Despite these deviations in the tail characteristics, the overall agreement between the empirical and theoretical statistics demonstrates that the proposed UWG distribution provides a statistically adequate and practically satisfactory characterization of Dataset II. An inspection of Table 11 reveals that the UWG distribution simultaneously achieves the lowest values of $-\ell$, AIC, AICC, BIC, and HQIC, as well as the smallest K–S test statistic (0.0751) and the largest associated p-value (0.9494) among all competing models. The exceptionally high p-value provides strong statistical evidence that the null hypothesis of the UWG distribution adequately fitting the data cannot be rejected. These results collectively establish that the UWG distribution delivers a markedly superior fit to Dataset II relative to all alternative models considered in this study. Accordingly, the UWG distribution is identified as the most appropriate and statistically efficient model for characterizing the probabilistic structure of Dataset II. Figure 9 presents the fitted probability density curves of the UWG distribution and all competing models superimposed on the empirical histogram of Dataset II. It is clearly evident that the estimated density of the UWG distribution conforms closely to the overall shape, modal structure, and concentration pattern of the empirical histogram, outperforming all competing models in reproducing the observed distributional characteristics. This strong visual agreement between the fitted density and the empirical distribution further reinforces the adequacy and flexibility of the proposed model for Dataset II.

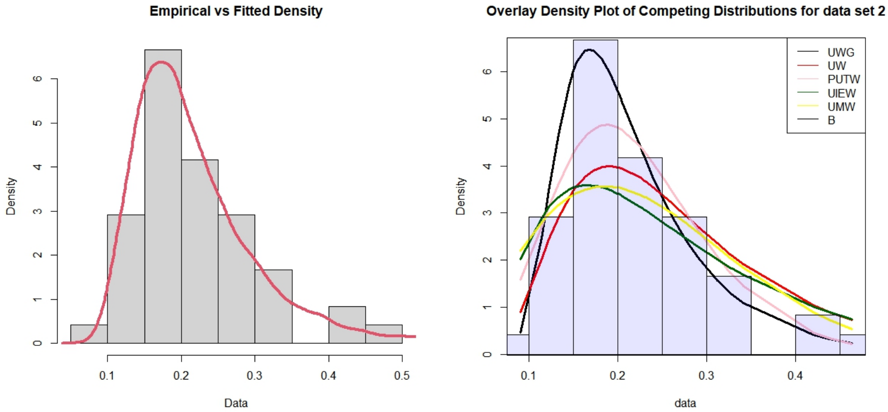

Figure 9. Empirical histogram with fitted density curves for Dataset II.

Figure 10 presents the violin plot with the embedded boxplot and the Q–Q plot for Dataset II under the UWG distribution. The violin plot reveals that the majority of observations are concentrated within the lower–to–middle range of the data, with a discernible right-skewed distributional shape and a limited number of observations extending into the upper tail, potentially indicative of a few high-porosity samples. The embedded boxplot confirms a moderate interquartile spread centered around the median, with no severe or extreme outliers disrupting the overall distributional pattern. The Q–Q plot demonstrates that the observed quantiles align closely with the theoretical quantiles of the UWG distribution along the 45° reference diagonal throughout the central range of the data, particularly within the interval [0.10, 0.35], confirming that the model accurately reproduces the core distributional behavior. Minor deviations are observed at the lower and upper tails, where a slight overestimation of the lowest quantile and moderate fluctuations in the highest quantiles are apparent. Such tail discrepancies are commonly encountered in empirical Datasets of moderate size and do not substantially compromise the overall adequacy of the fit. Collectively, these graphical diagnostics affirm that the UWG distribution successfully captures the principal stochastic features and underlying behavior of Dataset II.

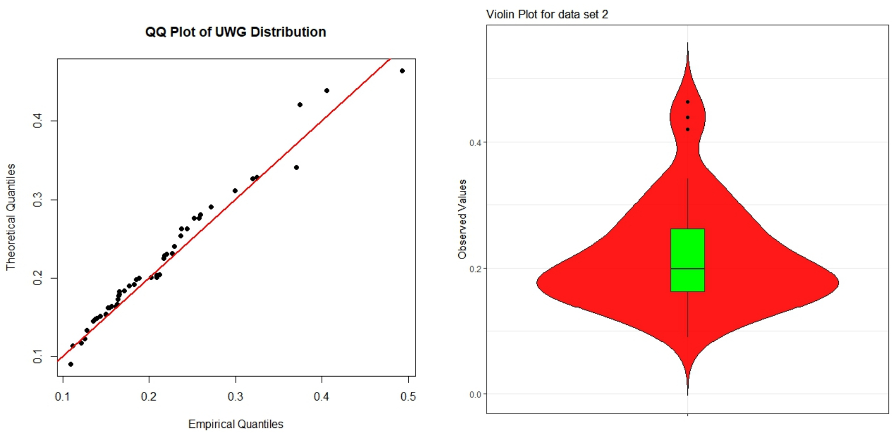

Figure 10. Distributional diagnostic plots for Dataset II.

Figure 11 presents the empirical and fitted CDFs alongside the P–P plot for Dataset II. The fitted

CDF of the UWG distribution tracks the empirical CDF with a high degree of precision across the entire support of the data, reflecting the model's capability to reproduce the observed cumulative behavior with strong fidelity. The P–P plot corroborates this finding, as the plotted points lie in close proximity to the 45° reference line with no systematic deviations, thereby providing compelling graphical confirmation of the goodness-of-fit of the proposed UWG model.

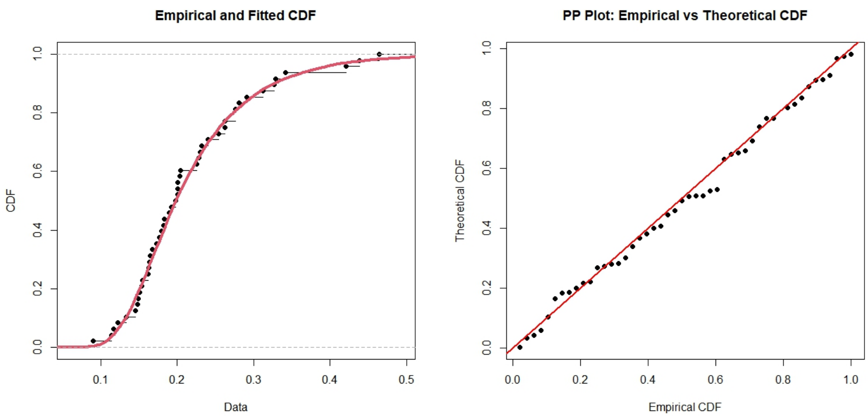

Figure 11. Cumulative distribution diagnostic plots for Dataset II.

Figure 12 presents the profile log-likelihood functions for each of the three model parameters α, β, and p of the UWG distribution fitted to Dataset II. All three profile log-likelihood curves exhibit smooth, well-behaved, and distinctly unimodal shapes, confirming the identifiability and numerical stability of the maximum likelihood estimates. The profiles for α and β display sharply defined global maxima at their respective MLE values, providing strong evidence for the uniqueness and precision of these estimates. Similarly, the profile log-likelihood for the parameter p attains a clear and well-defined peak at its estimated value, with a smooth monotonic decay on either side, further confirming the existence of a unique global maximum. The absence of multiple local maxima across all three profile curves, combined with their regular and smooth structure, provides compelling evidence that all model parameters are well-identified and that the numerical optimization converged reliably to the global maximum of the likelihood function. These results collectively affirm the reliability, stability, and robustness of the MLE procedure for the proposed UWG distribution when applied to Dataset II.

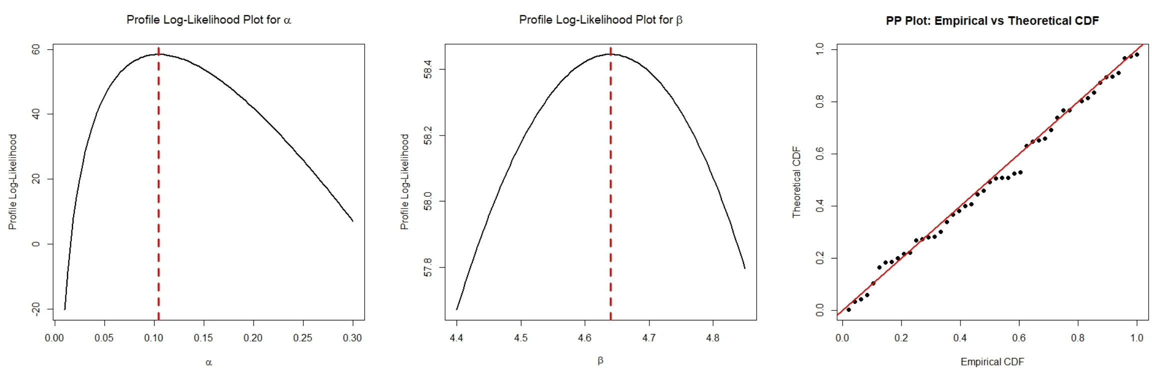

Figure 12. Profile log-likelihood functions of the UWG parameters for Dataset II.

9. Conclusions and strategic roadmap

This study introduced the UWG distribution, a novel three-parameter bounded probability distribution defined on the unit interval $(0, 1)$, constructed by compounding the unit-Weibull and geometric distributions. The physical motivation of the proposed model is grounded in reliability theory, risk assessment, and systems engineering, particularly in the context of series systems where the risk of overall system failure is governed by the minimum component lifetime. The UWG distribution offers exceptional flexibility for modeling bounded data such as proportions, rates, and fractions, while accommodating both symmetric and asymmetric shapes, ranging from positively to negatively skewed profiles, and a wide spectrum of kurtosis behaviors. A comprehensive mathematical and statistical characterization of the UWG distribution was carried out, encompassing identifiability analysis, derivation of the moment-generating structure, examination of modality, and a log-convexity and log–concavity analysis. Key reliability and information–theoretic measures were also derived, including quantile–based descriptive statistics, incomplete and conditional moments, mean residual life function, measures of inequality, and the distributional properties of order statistics. Notably, the HRF was shown to exhibit versatile shapes, including monotonically increasing, monotonically decreasing, unimodal, and bathtub-shaped profiles, rendering the model applicable to a broad range of lifetime phenomena. Model parameters were estimated via MLE, and the performance of the resulting estimators was evaluated through an extensive Monte Carlo simulation study. The simulation results, assessed in terms of bias, MSE, CP, and ACIL, confirmed the consistency, asymptotic efficiency, and reliability of the MLE procedure across various sample sizes and parameter configurations. As the sample size increased, the estimators converged toward their true values with diminishing bias and variance, validating the large-sample asymptotic properties of the proposed estimators. The practical adequacy of the UWG distribution was demonstrated through the analysis of two real-world Datasets. Preliminary diagnostics, including TTT plots and empirical hazard rate plots, supported the suitability of the proposed model for both Datasets. Further graphical assessments, comprising profile log-likelihood plots, Q–Q plots, P–P plots, violin plots, and Cullen–Frey graphs, provided a thorough evaluation of the model's fit and parameter stability. Comparative analyses based on $-\ell$, AIC, AICC, BIC, HQIC, and the K–S statistic consistently identified the UWG distribution as the best-fitting model, outperforming the UW, UMW, PUTW, UIEW, and B distributions across both Datasets. The theoretical descriptive statistics were also found to be in close agreement with the empirical characteristics of the observed data, further confirming the model's adequacy. Overall, the theoretical, inferential, and empirical results obtained in this study established the UWG distribution as a flexible and valuable addition to the family of bounded lifetime distributions, offering strong fitting performance, rich hazard rate flexibility, and sound inferential properties for analyzing unit interval data. Several avenues for future research remain open. One promising direction is the extension of the proposed UWG distribution to multivariate settings through suitable copula constructions, enabling the modeling of dependent bounded random variables while preserving the flexibility of the marginal distribution. Another important extension is the development of regression models for bounded response variables, where the parameters of the UWG distribution may be linked to explanatory variables. Such extensions are expected to introduce additional computational and inferential challenges, particularly with respect to parameter identifiability, optimization in higher-dimensional parameter spaces, and efficient estimation procedures. Addressing these challenges would further broaden the applicability of the proposed distribution in reliability, environmental, biomedical, and other fields involving bounded data. Finally, the development of new compounded and generalized families of distributions based on alternative system structures and failure mechanisms may yield more flexible models for complex real-world reliability and survival data.

Author contributions

Magrisha Namsaw, Bhanita Das, Mohamed S. Eliwa and Mohamed F. Abouelenein: Conception, analysis, and writing of the manuscript. All authors of this article have been contributed equally. 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.

Funding

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2603).

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. A. W. Marshall, I. Olkin, A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, *Biometrika*, **84** (1997), 641–652.
2. K. Adamidis, S. Loukas, A lifetime distribution with decreasing failure rate, *Stat. Probab. Lett.*, **39** (1998), 35–42. https://doi.org/10.1016/S0167-7152(98)00012-1
3. C. Kuş, A new lifetime distribution, *Comput. Stat. Data Anal.*, **51** (2007), 4497–4509. https://doi.org/10.1016/j.csda.2006.07.017
4. R. Tahmasbi, S. Rezaei, A two-parameter lifetime distribution with decreasing failure rate, *Comput. Stat. Data Anal.*, **52** (2008), 3889–3901. https://doi.org/10.1016/j.csda.2007.12.002
5. W. Barreto-Souza, A. L. Morais, G. M. Cordeiro, The Weibull-geometric distribution, *J. Stat. Comput. Simul.*, **81** (2011), 645–657. https://doi.org/10.1080/00949650903436554
6. W. Lu, D. Shi, A new compounding life distribution: The Weibull–Poisson distribution, *J. Appl. Stat.*, **39** (2012), 21–38. https://doi.org/10.1080/02664763.2011.575126
7. G. M. Cordeiro, E. M. M. Ortega, A. J. Lemonte, The exponential–Weibull lifetime distribution, *J. Stat. Comput. Simul.*, **84** (2014), 2592–2606. https://doi.org/10.1080/00949655.2013.797982
8. M. Wang, I. Elbatal, The modified Weibull geometric distribution, *Metron*, **73** (2015), 303–315.
9. I. Elbatal, M. Mansour, M. Ahsanullah, The additive Weibull-geometric distribution: Theory and applications, *J. Stat. Theory Appl.*, **15** (2016), 125–141. https://doi.org/10.2991/jsta.2016.15.2.3

10. A. A. Fattah, S. Nadarajah, A. -H. N. Ahmed, The exponentiated transmuted Weibull geometric distribution with application in survival analysis, *Commun. Stat. Simul. Comput.*, **46** (2017), 4244–4263. https://doi.org/10.1080/03610918.2015.1112909
11. Y. Chung, D. K. Dey, M. Jung, The exponentiated inverse Weibull geometric distribution, *Pak. J. Stat.*, **33** (2017), 161–178.
12. A. Asgharzadeh, S. Nadarajah, F. Sharafi, Weibull Lindley distribution, *REVSTAT-Stat. J.*, **16** (2018), 87–113. https://doi.org/10.57805/revstat.v16i1.234
13. J. B. Chakrabarty, S. Chowdhury, Compounded inverse Weibull distributions: Properties, inference and applications, *Commun. Stat. Simul. Comput.*, **48** (2019), 2012–2033. https://doi.org/10.1080/03610918.2018.1429623
14. A. Hassan, N. Akhtar, A. Rashid, A. Iqbal, A new compound lifetime distribution: Ishita power series distribution with properties and applications, *J. Stat. Appl. Pro.*, **10** (2021), 883–896. http://dx.doi.org/10.18576/jsap/100324
15. M. K. Shakhatreh, S. Dey, D. Kumar, Inverse Lindley power series distributions: A new compounding family and regression model with censored data, *J. Appl. Stat.*, **49** (2022), 3451–3476. https://doi.org/10.1080/02664763.2021.1951683
16. A. Elshahhat, E. -S. A. El-Sherpieny, A. S. Hassan, The Pareto–Poisson distribution: Characteristics, estimations and engineering applications, *Sankhya A*, **85** (2023), 1058–1099. https://doi.org/10.1007/s13171-022-00302-6
17. H. Memarbashi, J. Etminan, M. Chahkandi, On the Zhang-power series distributions with application to lifetime modeling, *J. Mahani Math. Res.*, **14** (2025), 505–525. https://doi.org/10.22103/jmmr.2024.23690.1676
18. M. Namsaw, B. Das, P. J. Hazarika, M. Alizadeh, The Alpha power modified Weibull-geometric distribution: A comprehensive mathematical framework with simulation, goodness-of-fit analysis and informed decision making using real life data, *Stat. Optim. Inf. Comput.*, **14** (2025), 1060–1087. https://doi.org/10.19139/soic-2310-5070-2607
19. M. A. Zayed, A. S. Hassan, E. M. Almetwally, A. M. Aboalkhair, A. H. Al-Nefaie, H. M. Almongy, A compound class of unit Burr XII model: Theory, estimation, fuzzy, and application, *Sci. Program.*, **2023** (2023), 4509889. https://doi.org/10.1155/2023/4509889
20. S. M. Alghamdi, M. Shrahili, A. S. Hassan, R. E. Mohamed, I. Elbatal, M. Elgarhy, Analysis of milk production and failure data: Using unit exponentiated half logistic power series class of distributions, *Symmetry*, **15** (2023), 714. https://doi.org/10.3390/sym15030714
21. M. Gabanakgosi, B. Oluyede, The Topp-Leone-Gompertz-G power series class of distributions with applications, *Pak. J. Stat. Oper. Res.*, **20** (2024), 171–195. https://doi.org/10.18187/pjsor.v20i2.4032
22. M. M. Yousef, A. S. Hassan, E. M. Almetwally, Statistical inference for the unit Gompertz power series distribution using ranked set sampling with applications, *Assiut Univ. J. Multidiscip. Sci. Res.*, **53** (2024), 154–189. https://doi.org/10.21608/aunj.2023.246431.1068
23. M. S. Eliwa, M. H. Tahir, M. A. Hussain, B. Almohaimeed, A. Al-Bossly, M. El-Morshedy, Univariate probability-G classes for scattered samples under different forms of hazard: Continuous and discrete version with their inferences tests, *Mathematics*, **11** (2023), 2929. https://doi.org/10.3390/math11132929

24. M. El-Morshedy, M. H. Tahir, M. A. Hussain, A. Al-Bossly, M. S. Eliwa, A new flexible univariate and bivariate family of distributions for unit interval (0, 1), *Symmetry*, **14** (2022), 1040. https://doi.org/10.3390/sym14051040
25. M. S. Eliwa, M. Ahsan-ul-Haq, A. Al-Bossly, M. El-Morshedy, A unit probabilistic model for proportion and asymmetric data: Properties and estimation techniques with application to model data from SC16 and P3 algorithms, *Math. Probl. Eng.*, **2022** (2022), 9289721. https://doi.org/10.1155/2022/9289721
26. E. Altun, M. El-Morshedy, M. S. Eliwa, A new regression model for bounded response variable: An alternative to the beta and unit-Lindley regression models, *PLoS ONE*, **16** (2021), e0245627. https://doi.org/10.1371/journal.pone.0245627
27. M. H. Tahir, M. A. Hussain, G. M. Cordeiro, M. El-Morshedy, M. S. Eliwa, A new Kumaraswamy generalized family of distributions with properties, applications, and bivariate extension, *Mathematics*, **8** (2020), 1989. https://doi.org/10.3390/math8111989
28. J. Mazucheli, A. F. B. Menezes, M. E. Ghitany, The unit-Weibull distribution and associated inference, *J. Appl. Probab. Stat.*, **13** (2018), 1–22.
29. G. Casella, R. Berger, *Statistical inference*, 2 Eds., New York: Chapman and Hall/CRC, 2024. https://doi.org/10.1201/9781003456285
30. D. R. Cox, D. V. Hinkley, *Theoretical statistics*, New York: Chapman and Hall, 1974. https://doi.org/10.1201/b14832
31. A. M. Sarhan, T. Manshi, M. E. Sobh, The unit-modified Weibull distribution: Theory, estimation, and real-world applications, *Stats*, **8** (2025), 81. https://doi.org/10.3390/stats8030081
32. A. M. Abd El-latif, O. A. Alqasem, J. K. Okutu, C. Tanış, L. P. Sapkota, N. A. Noori, A flexible extension of the unit upper truncated Weibull distribution: Statistical analysis with applications on geology, engineering, and radiation data, *J. Radiat. Res. Appl. Sci.*, **18** (2025), 101434. https://doi.org/10.1016/j.jrras.2025.101434
33. A. S. Hassan, R. S. Alharbi, Different estimation methods for the unit inverse exponentiated Weibull distribution, *Commun. Stat. Appl. Methods*, **30** (2023), 191–213. https://doi.org/10.29220/CSAM.2023.30.2.191
34. N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous univariate distributions*, 2 Eds., New York: John Wiley & Sons, 1995.
35. A. Z. Afify, M. Nassar, D. Kumar, G. M. Cordeiro, A new unit distribution: Properties, inference, and applications, *Electron. J. Appl. Stat. Anal.*, **15** (2022), 460–484. https://doi.org/10.1285/i20705948v15n2p460
36. G. M. Cordeiro, R. dos Santos Brito, The beta power distribution, *Braz. J. Probab. Stat.*, **26** (2012), 88–112. http://dx.doi.org/10.1214/10-BJPS124