AIMS Mathematics, 10(4): 9773–9804.
DOI: 10.3934/math.2025448
Received: 09 February 2025
Revised: 13 April 2025
Accepted: 22 April 2025
Published: 27 April 2025

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

Research article

The Weibull-generalized shifted geometric distribution: properties, estimation, and applications

Mohieddine Rahmouni[1,*] **and Dalia Ziedan**[2]

[1] Applied College, King Faisal University, Al-Ahsa, Saudi Arabia
[2] Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt

* **Correspondence:** Email: mrahmouni@kfu.edu.sa.

Abstract: This paper introduces the Weibull-generalized shifted geometric (WGSG) distribution, a novel lifetime model integrating the Weibull and shifted geometric distributions to address complex lifetime data patterns. Extending the Weibull-geometric framework, this distribution models system reliability by focusing on the k-th smallest lifetime—when k components fail—rather than the minimum. Key properties, including the probability density function, cumulative distribution function, and moments, were derived. Parameters were estimated using maximum likelihood, expectation-maximization, method of moments, and Bayesian approaches, with a simulation study comparing their performance. Applications to two real-world lifetime and reliability datasets demonstrated the distribution's superiority over classical models in handling challenging survival and reliability scenarios. This flexible model enhances the ability to capture diverse hazard behaviors, advancing lifetime data analysis.

Keywords: lifetime distributions; reliability; failure rate; order statistics; WGSG distribution; geometric distribution; entropy
Mathematics Subject Classification: 60E05, 62E15, 62F10

1. Introduction

Lifetime distributions are integral to reliability theory and survival analysis, offering critical tools for analyzing real-world data in fields such as finance, manufacturing, biological sciences, physics, and engineering [1, 2]. Traditional distributions like the exponential and Weibull are valued for their simplicity and interpretability. The exponential distribution, with its assumption of a constant failure rate, is often used to describe system reliability at the component level [3–5]. However, real-world systems frequently exhibit non-constant hazard rates, requiring more flexible distributions to capture increasing or decreasing failure behaviors. Although the Weibull distribution can accommodate both

increasing and decreasing failure rates, it often proves inadequate for complex datasets involving multiple failure causes or censored data, necessitating more advanced techniques to accurately represent intricate phenomena [6].

To address the limitations of classical lifetime distributions in capturing the diverse behaviors observed in real-world reliability and survival data, recent research has increasingly focused on constructing new families of distributions through the compounding of continuous lifetime distributions with truncated discrete distributions [7–9]. This innovative approach enhances the flexibility and applicability of standard techniques by introducing additional parameters that allow for better representation of skewness, kurtosis, and hazard rate shapes. Specifically, continuous distributions such as the exponential, Weibull, Rayleigh, and gamma have been compounded with discrete distributions like the Poisson, geometric, binomial, negative-binomial, logarithmic, and more generally, power series distributions. The foundational work by Adamidis & Loukas [10], which introduced the exponential-geometric distribution, laid the groundwork for a wide class of compound distributions. This approach was later expanded by Adamidis et al. [11] and Kuş [12], who explored further generalizations and alternative mixing distributions, leading to a surge in the development of new techniques. These compound distributions exhibit greater flexibility, enabling better fitting of datasets with complex features such as increasing, decreasing, bathtub-shaped, or upside-down bathtub hazard functions. Notable contributions in this domain include extended versions of well-known lifetime distributions that have demonstrated improved performance in applications spanning reliability engineering, biostatistics, and actuarial science [13–17]. These distributions not only offer improved statistical properties—such as the ability to accommodate over-dispersion and unobserved heterogeneity—but also maintain mathematical tractability, allowing for explicit expressions of the probability density, cumulative distribution, moments, and hazard functions. For a comprehensive overview and classification of these compounding models and their theoretical properties, interested readers may refer to the reviews by Goyal et al. [7], Maurya & Nadarajah [8], Rahmouni & Orabi [18, 19], and Tahir & Cordeiro [9].

Building on these advancements, we focus on the need to analyze intermediate failure events, as prior studies often concentrated on the minimum lifetime, $X_{(1)} = \min\{Y_i\}_{i=1}^N$, or the maximum lifetime, $X_{(N)} = \max\{Y_i\}_{i=1}^N$, of a system's components. Many real-world applications, however, demand the analysis of the k-th failure. Order statistics, arranged in ascending order as $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(N)}$, are essential in reliability analysis [20–23] and also apply to economics, where they provide distributional criteria for evaluating welfare and inequalities in income and wealth [24, 25]. To address this need, we propose the Weibull-generalized shifted geometric (WGSG) distribution, a new family of compound distributions that generalizes the Weibull-geometric (WG) distribution introduced by Barreto-Souza et al. [26]. The WG distribution describes the time to the first failure in a system with a random number of components, N, where each component's lifetime, Y_i, follows a Weibull distribution, and N follows a geometric distribution. Our approach extends this model to capture the k-th order statistic, $X_{(k)}$, enabling the analysis of any failure order, such as the second, third, or k-th lifetime. To illustrate this approach, consider a machine producing N units (e.g., light bulbs or wire fuses), each with an independent Weibull-distributed lifetime Y_i ($i = 1, 2, \ldots, N$). If N follows a shifted geometric distribution—adjusting the starting point of the geometric progression to $n \geq k$ while preserving its properties—the k-th order statistic $X_{(k)}$ provides a versatile method for analyzing system reliability. This shifted geometric distribution is particularly valuable in scenarios where counting begins at a

specific point k, such as in quality control, queueing theory, inventory management, and reliability systems with latent defects [27].

The WGSG distribution introduces additional shape parameters to capture a broad spectrum of failure behaviors, making it well-suited for analyzing the k-th lifetime in systems where failure occurs after k units fail. This model aligns with the analysis of k-out-of-n systems, where a system fails when $n - k + 1$ or more components fail [28–31]. Unlike traditional approaches that assume binary states (fully functional or failed), the WGSG distribution can accommodate degraded states in multi-state systems, which are prevalent in practical applications [32–35].

This paper investigates the properties of the WGSG distribution, explores parameter estimation techniques, and demonstrates its practical applications using real datasets. Building on the compounding methodology of Adamidis & Loukas [10] and Kuş [12], we aim to provide a versatile tool for reliability and survival analysis that addresses the limitations of traditional approaches in capturing complex failure dynamics. The paper is structured as follows: Section 2 introduces the WGSG distribution and its probability density function (PDF). Section 3 discusses its properties, including reliability, the failure rate function, the moment generating function, r-th moments, and Shannon entropy. Section 4 presents parameter estimation approaches, such as maximum likelihood estimation (MLE), the expectation-maximization (EM) algorithm, and method of moments. Section 5 presents the simulation study and Section 6 illustrates the WGSG distribution's practicality with real datasets. Finally, Section 7 concludes the paper and suggests directions for future research.

2. The WGSG distribution

Let X be a random variable following the WGSG distribution, parameterized by $\varphi = (\alpha, \theta, p)$, where $\alpha > 0$, $\theta = \frac{1}{\beta} > 0$, and $p \in (0, 1)$. Then, the PDF of the WGSG distribution is:

$$f_{X_{(k)}}(x; \alpha, \theta, p) = \frac{k(1-p)\alpha\theta^{\alpha}x^{\alpha-1}e^{-(\theta x)^{\alpha}}\left(1 - e^{-(\theta x)^{\alpha}}\right)^{k-1}}{\left(1 - pe^{-(\theta x)^{\alpha}}\right)^{k+1}}, \quad x \geq 0. \tag{2.1}$$

The cumulative distribution function (CDF) is given by:

$$F_{X_{(k)}}(x; \alpha, \theta, p) = \left(\frac{1 - e^{-(\theta x)^{\alpha}}}{1 - pe^{-(\theta x)^{\alpha}}}\right)^{k}, \quad x \geq 0. \tag{2.2}$$

This formulation characterizes the mixture between the shifted geometric distribution and the k^{th}-order statistic of Weibull-distributed random variables. A notable special case occurs when $k = 1$, where the WGSG density simplifies to the Weibull-geometric (WG) distribution, which models the time to the first failure, $X_{(1)} = \min\{Y_i\}_{i=1}^{N}$, as introduced in Barreto-Souza et al. [26]. Figures 1 and 2 display the PDF and CDF of the WGSG distribution for parameter values $\alpha \in \{0.3, 0.5, 1, 1.5, 2, 2.5\}$, $\theta = 1$, $p = 0.3$, and $k \in \{1, 2, 3, 4, 5\}$. These plots illustrate the distribution's flexibility in modeling a wide range of reliability scenarios. When $\alpha < 1$ (e.g., 0.3, 0.5), the PDF exhibits a sharp decline and a singularity at $x = 0$, capturing early failure behavior. In contrast, for $\alpha > 1$ (e.g., 1.5, 2, 2.5), the PDF becomes unimodal, with its peak shifting leftward and becoming sharper as α increases. Higher values of k yield flatter peaks and more rapid tail decay in the PDF, reflecting the compounding effect of higher-order statistics. Similarly, the CDF becomes steeper as α and k increase, indicating lighter

tails and a greater concentration of probability mass near the origin. These characteristics underline the WGSG distribution's utility in reliability engineering as well as in survival analysis involving extreme event times and in risk assessment for rare failures. The parameters k and p offer fine control over the distribution's shape, hazard rate, and tail behavior, enhancing its applicability across various practical contexts.

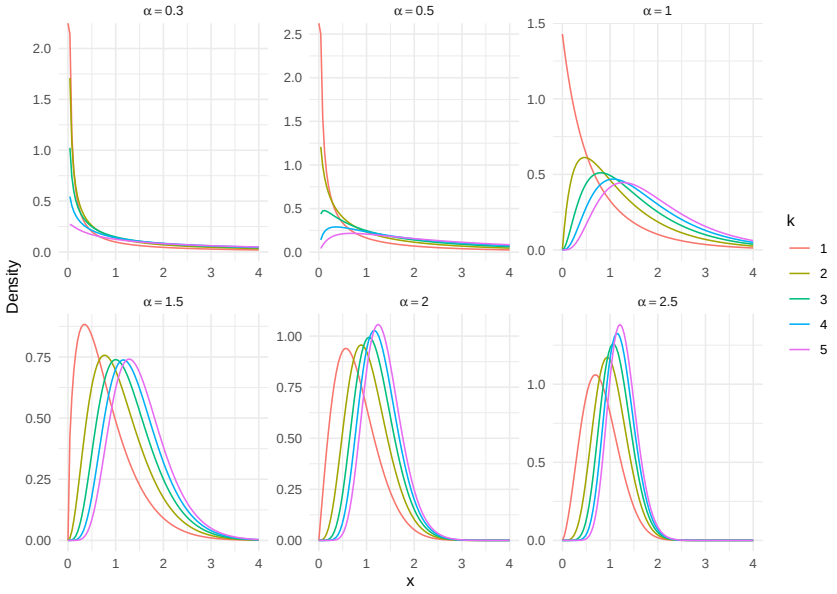

Figure 1. The PDF of the WGSG distribution.

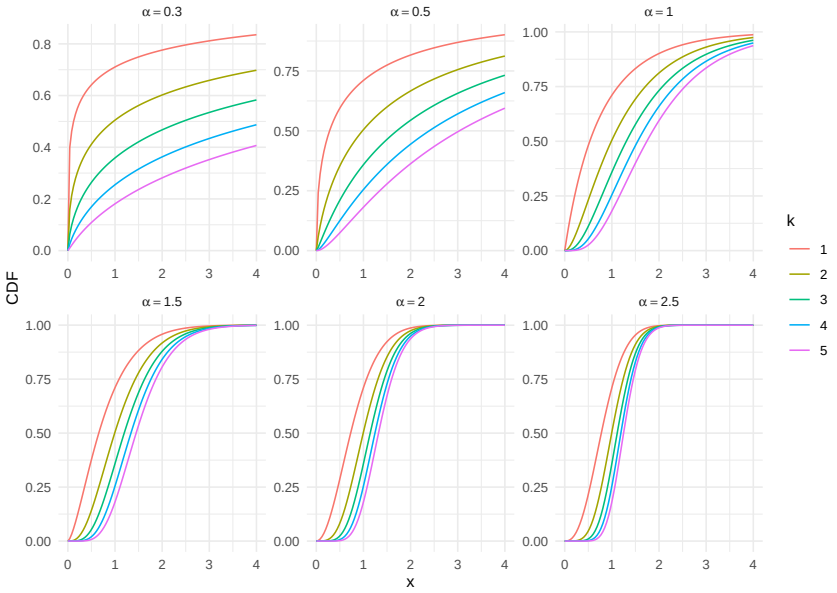

Figure 2. The CDF of the WGSG distribution.

The Weibull-like term $\alpha\theta^{\alpha}x^{\alpha-1}e^{-(\theta x)^{\alpha}}$ governs the overall shape of the PDF, typically exhibiting

unimodal behavior—initially increasing, reaching a peak, and then decreasing. The geometric component, involving the parameters p and k, adjusts the probability structure and influences the tail behavior potentially making it heavier or lighter depending on p and k. For small values of x, the Weibull-like term dominates, causing the PDF to increase as x grows, with the exponential term $e^{-(\theta x)^\alpha}$ having minimal effect. As x continues to increase, the exponential term starts to decay and the PDF reaches a maximum at some intermediate value, where the Weibull-like component and the geometric factors balance. For large x, the exponential term takes over, causing the PDF to decline. The geometric component further ensures that the distribution tail decays smoothly without forming multiple modes.

To establish that the WGSG distribution is unimodal, we examine the first derivative of the PDF with respect to x. The critical points of the PDF, where the derivative equals zero, indicate the mode. The first derivative is positive for small x, showing an increasing density, and negative for larger x, indicating a decreasing density. The presence of a single critical point where the first derivative changes sign from positive to negative suggests unimodality. Further verification via the second derivative or concavity analysis strengthens this conclusion. Therefore, the WGSG distribution is unimodal when $\alpha > 1$, characterized by a single peak followed by a steady decline, with no secondary modes present.

As $p \to 1$, the WGSG distribution becomes a Weibull distribution with shape parameter α and scale parameter θ because the shifted geometric component effectively becomes deterministic, i.e., $P(N = k) \to 1$. The mixture model collapses to a single Weibull component, and the number of components (determined by N) becomes one. The limiting behavior is given by:

$$\lim_{p\to 1} f_X(x) = \alpha\theta^\alpha x^{\alpha-1} e^{-(\theta x)^\alpha}.$$

Proof of the PDF and CDF expressions in Eqs (2.1) **and** (2.2)**:** Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from a Weibull distribution. The k-th order statistic, $X_{(k)}$, is the k-th smallest value in the sample. The PDF of $X_{(k)}$ is:

$$f_{X_{(k)}}(x;\alpha,\theta,n) = \frac{n!}{(k-1)!(n-k)!}\,[F(x;\alpha,\theta)]^{k-1}\,[1-F(x;\alpha,\theta)]^{n-k}\,f(x;\alpha,\theta),$$

where $f(x;\alpha,\theta) = \alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}$ is the PDF of a Weibull random variable and $F(x) = 1 - e^{-(\theta x)^\alpha}$ is the CDF with scale parameter $\theta > 0$ and shape parameter $\alpha > 0$. Thus, the PDF of the k-th order Weibull distribution is:

$$f_{X_{(k)}}(x;\alpha,\theta,n) = \frac{n!}{(k-1)!(n-k)!}\left(1-e^{-(\theta x)^\alpha}\right)^{k-1} e^{-(n-k+1)(\theta x)^\alpha}\alpha\theta^\alpha x^{\alpha-1}, \quad x \geq 0.$$

The probability mass function (PMF) of the distribution of N that begins counting from k is:

$$P(N = n) = (1-p)p^{n-k}, \quad n = k, k+1, k+2, \ldots,$$

where p is the probability of success on each trial. The marginal PDF is obtained by summing over all possible values of N:

$$f_{X_{(k)}}(x;\alpha,\theta,p) = \sum_{n=k}^{\infty} f_{X_{(k)}}(x;\alpha,\theta,n)P(N=n).$$

Substituting $f_{X_{(k)}}(x;\alpha,\theta,n)$ and $P(N = n)$:

$$f_{X_{(k)}}(x;\alpha,\theta,p) = \alpha\theta^{\alpha}x^{\alpha-1}\left(1-e^{-(\theta x)^{\alpha}}\right)^{k-1}e^{-(\theta x)^{\alpha}}(1-p)\sum_{n=k}^{\infty}\frac{n!}{(k-1)!(n-k)!}p^{n-k}e^{-(n-k)(\theta x)^{\alpha}}.$$

Let $n = m + k$, where $m = n - k$. The summation becomes:

$$\sum_{m=0}^{\infty}\frac{(m+k)!}{(k-1)!\,m!}p^{m}e^{-m(\theta x)^{\alpha}}.$$

Using the rising factorial (Pochhammer symbol) $(k+1)_m = \frac{\Gamma(m+k+1)}{\Gamma(k+1)}$, the summation can be expressed as:

$$k\sum_{m=0}^{\infty}\frac{(k+1)_m}{m!}\left(pe^{-(\theta x)^{\alpha}}\right)^{m}.$$

This series is the generalized hypergeometric function ${}_1F_0(-;k;z)$, where $z = pe^{-(\theta x)^{\alpha}}$. Thus:

$$\sum_{m=0}^{\infty}\frac{(k+1)_m}{m!}\left(pe^{-(\theta x)^{\alpha}}\right)^{m} = \left(1-pe^{-(\theta x)^{\alpha}}\right)^{-(k+1)}.$$

Substituting back:

$$f_{X_{(k)}}(x;\alpha,\theta,p) = \alpha\theta^{\alpha}x^{\alpha-1}\left(1-e^{-(\theta x)^{\alpha}}\right)^{k-1}\frac{(1-p)ke^{-(\theta x)^{\alpha}}}{(1-pe^{-(\theta x)^{\alpha}})^{k+1}}.$$

Then,

$$F_{X_{(k)}}(x;\alpha,\theta,p) = \int_0^x f_X(t)\,dt,$$

where:

$$f_X(t) = \alpha\theta^{\alpha}t^{\alpha-1}\left(1-e^{-(\theta t)^{\alpha}}\right)^{k-1}\frac{(1-p)ke^{-(\theta t)^{\alpha}}}{(1-pe^{-(\theta t)^{\alpha}})^{k+1}}.$$

Let $u = 1 - e^{-(\theta t)^{\alpha}}$. Then, $du = \alpha\theta^{\alpha}t^{\alpha-1}e^{-(\theta t)^{\alpha}}\,dt$.
The PDF in terms of u is:

$$f_X(u) = \frac{k(1-p)u^{k-1}}{(1-p(1-u))^{k+1}}.$$

Thus, the CDF becomes:

$$F_X(x) = \int_0^{1-e^{-(\theta x)^{\alpha}}}\frac{k(1-p)u^{k-1}}{(1-p+pu)^{k+1}}\,du = \left(\frac{1-e^{-(\theta x)^{\alpha}}}{1-pe^{-(\theta x)^{\alpha}}}\right)^{k}.$$

3. Proprieties

This section characterizes the WGSG distribution by examining its survival and hazard functions (including asymptotic tails), moment generating function, r-th moments, limiting behavior, relationships with standard distributions in reliability and survival analysis (e.g., Weibull, geometric, and exponential), and entropy.

3.1. Survival and hazard functions

The survival function $S_X(x)$ of the WGSG distribution represents the probability that a random variable X exceeds a given value x:

$$S_X(x) = 1 - F_X(x) = 1 - \left(\frac{1 - e^{-(\theta x)^\alpha}}{1 - pe^{-(\theta x)^\alpha}}\right)^k. \tag{3.1}$$

The corresponding hazard function $h_X(x)$, which quantifies the instantaneous failure rate, is given by:

$$h_X(x) = \frac{f_X(x)}{S_X(x)} = \frac{k(1-p)\alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}\left(1 - e^{-(\theta x)^\alpha}\right)^{k-1}}{\left(1 - pe^{-(\theta x)^\alpha}\right)^{k+1} - \left(1 - e^{-(\theta x)^\alpha}\right)^k\left(1 - pe^{-(\theta x)^\alpha}\right)}. \tag{3.2}$$

The tail behavior of the WGSG distribution governs the likelihood of extreme values, critical for risk assessment. As $x \to \infty$, $e^{-(\theta x)^\alpha} \to 0$. Letting $z = e^{-(\theta x)^\alpha}$, $F_X(x) = \left(\frac{1-z}{1-pz}\right)^k$, so $F_X(x) \to 1$ and $S_X(x) \to 0$, indicating a light-tailed distribution.

For large x, $1 - e^{-(\theta x)^\alpha} \to 1$ and $1 - pe^{-(\theta x)^\alpha} \to 1$, yielding $f_X(x) \sim k(1-p)\alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}$. The exponential term drives rapid decay, modulated by $x^{\alpha-1}$, with the tail exhibiting exponential decay for $\alpha = 1$, subexponential decay for $\alpha > 1$, or stretched exponential decay for $\alpha < 1$, accommodating diverse extreme event scenarios. To characterize the tail and hazard behavior, we derive asymptotic expansions for $S_X(x)$ and $h_X(x)$.

Theorem 1 (Right-tail and hazard asymptotics)**.** *For $X \sim WGSG(\alpha, \theta, p, k)$, the following hold:*

(1) As $x \to \infty$, the survival function satisfies:

$$S_X(x) \sim k(1-p)e^{-(\theta x)^\alpha}\left[1 - \left(\frac{k-1}{2} + \frac{p}{1-p}\right)e^{-(\theta x)^\alpha} + O(e^{-2(\theta x)^\alpha})\right].$$

(2) The hazard function behaves as:

$$h_X(x) \sim \begin{cases} \dfrac{k\alpha\theta^{\alpha k}}{(1-p)^k}x^{\alpha k-1}, & as\ x \to 0^+, \\ \alpha\theta^\alpha x^{\alpha-1}, & as\ x \to \infty. \end{cases}$$

Proof. (Right-tail asymptotics) Let $z = e^{-(\theta x)^\alpha}$, so $z \to 0^+$ as $x \to \infty$. The survival function is:

$$S_X(x) = 1 - \left(\frac{1-z}{1-pz}\right)^k.$$

Expand:

$$\frac{1-z}{1-pz} = (1-z)(1-pz)^{-1}.$$

Using $(1-pz)^{-1} = 1 + pz + p^2z^2 + O(z^3)$:

$$\frac{1-z}{1-pz} = (1-z)(1+pz+p^2z^2+O(z^3)) = 1-(1-p)z+p(p-1)z^2+O(z^3).$$

Compute:

$$\left(1-(1-p)z+p(p-1)z^2+O(z^3)\right)^k = 1-k(1-p)z+\left[kp(p-1)+\frac{k(k-1)}{2}(1-p)^2\right]z^2+O(z^3).$$

Thus:

$$S_X(x) = k(1-p)z-\left[kp(p-1)+\frac{k(k-1)}{2}(1-p)^2\right]z^2+O(z^3).$$

Factorize:

$$S_X(x) = k(1-p)z\left[1-\frac{kp(p-1)+\frac{k(k-1)}{2}(1-p)^2}{k(1-p)}z+O(z^2)\right].$$

Simplify the coefficient:

$$\frac{kp(p-1)+\frac{k(k-1)}{2}(1-p)^2}{k(1-p)} = \frac{p(p-1)}{1-p}+\frac{(k-1)(1-p)}{2} = \frac{k-1}{2}+\frac{p}{1-p}.$$

With $z = e^{-(\theta x)^\alpha}$:

$$S_X(x) = k(1-p)e^{-(\theta x)^\alpha}\left[1-\left(\frac{k-1}{2}+\frac{p}{1-p}\right)e^{-(\theta x)^\alpha}+O(e^{-2(\theta x)^\alpha})\right].$$

□

Proof. (Hazard rate asymptotics) As $x \to \infty$, using $f_X(x) \sim k(1-p)\alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}$ and $S_X(x) \sim k(1-p)e^{-(\theta x)^\alpha}$:

$$h_X(x) \sim \frac{k(1-p)\alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}}{k(1-p)e^{-(\theta x)^\alpha}} = \alpha\theta^\alpha x^{\alpha-1}.$$

As $x \to 0^+$, approximate $e^{-(\theta x)^\alpha} \approx 1-(\theta x)^\alpha$. Then:

$$F_X(x) \approx \left(\frac{(\theta x)^\alpha}{1-p}\right)^k = \frac{\theta^{\alpha k}x^{\alpha k}}{(1-p)^k}.$$

For the PDF:

$$f_X(x) \approx \frac{k(1-p)\alpha\theta^\alpha x^{\alpha-1}\left((\theta x)^\alpha\right)^{k-1}}{(1-p)^{k+1}} = \frac{k\alpha\theta^{\alpha k}x^{\alpha k-1}}{(1-p)^k}.$$

Since $S_X(x) \approx 1$:

$$h_X(x) \approx f_X(x) \sim \frac{k\alpha\theta^{\alpha k}}{(1-p)^k}x^{\alpha k-1}.$$

□

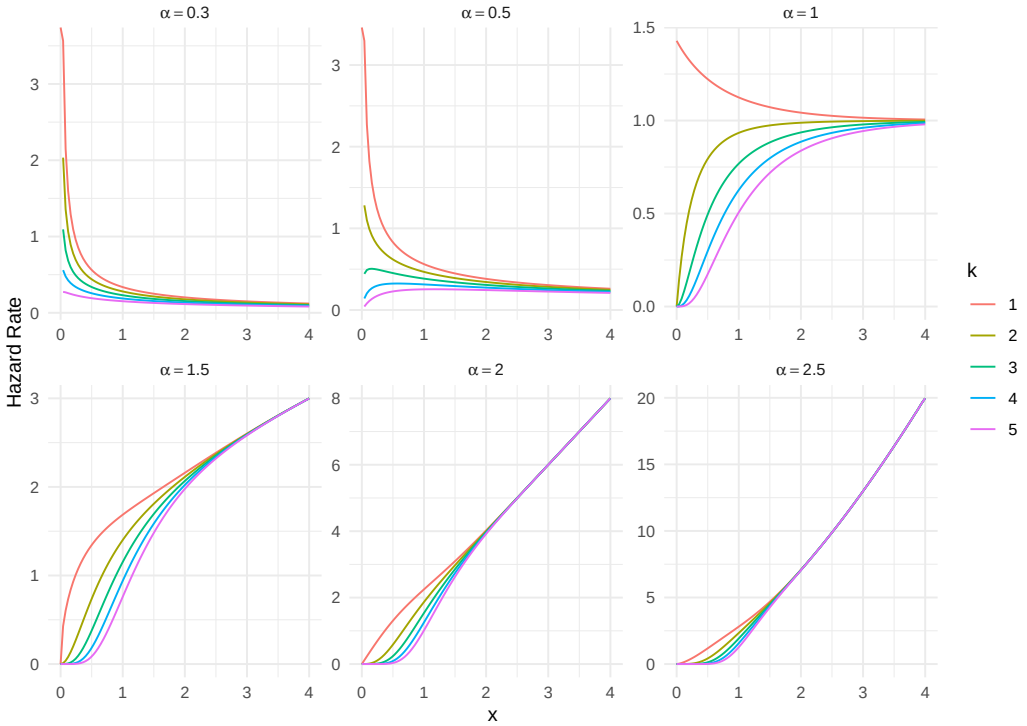

Figure 3. Hazard function for the WGSG distribution.

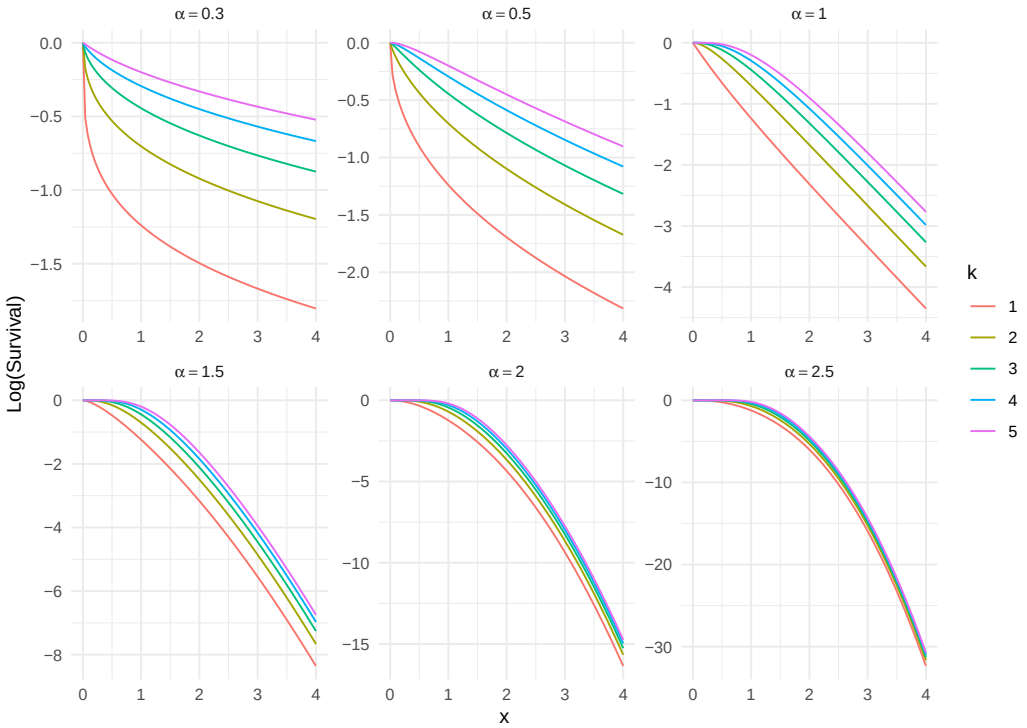

Figure 4. Log-survival function for the WGSG distribution.

The hazard rate, illustrated in Figure 3, transitions from $\frac{k\alpha\theta^{\alpha k}}{(1-p)^k}x^{\alpha k-1}$ near the origin to $\alpha\theta^{\alpha}x^{\alpha-1}$ for large x. For $\alpha < 1$, it decreases, capturing early failures; for $\alpha > 1$, it increases, reflecting wear-out phases; and for $\alpha = 1$, it remains constant.

The log-survival function, depicted in Figure 4, confirms the WGSG's light-tailed nature, with $\log(S_X(x)) \approx \log(k(1-p)) - (\theta x)^{\alpha}$ for large x. Smaller α values yield slower decay, suitable for prolonged survival, while larger α values produce steeper declines, indicating rapid failure onset.

The WGSG distribution's hazard rate supports complex patterns, such as bathtub-shaped hazard functions, common in reliability studies. The parameter p enhances adaptability to varied tail behaviors, and expansions like $(1-(1-p)z)^k$ may arise in series representations, making the WGSG distribution a robust tool for precise tail and hazard estimates in reliability engineering and survival analysis.

3.1.1. Limiting properties

The limiting behavior of the WGSG distribution reveals its flexibility and parameter sensitivity, with the survival function defined in Equation (3.1). As the shape parameter $\alpha \to 0^+$, the approximation $e^{-(\theta x)^{\alpha}} \to e^{-1}$ yields:

$$S_X(x) \to 1 - \left(\frac{e-1}{e-p}\right)^k. \tag{3.3}$$

Such a constant survival function implies a degenerate, uniform-like distribution over x, rather than a point mass at $x = 0$. Conversely, as $\alpha \to \infty$:

$$S_X(x) \to \begin{cases} 1, & x < 1/\theta, \\ 1 - \left(\frac{1-e^{-1}}{1-pe^{-1}}\right)^k, & x = 1/\theta, \\ 0, & x > 1/\theta. \end{cases} \tag{3.4}$$

The resulting step-function concentrates probability mass around $x = 1/\theta$. For the scale parameter, as $\theta \to 0^+$:

$$e^{-(\theta x)^{\alpha}} \to 1 \quad \Rightarrow \quad S_X(x) \to 1. \tag{3.5}$$

The distribution becomes improper, with probability mass dispersed toward infinity. As $\theta \to \infty$:

$$e^{-(\theta x)^{\alpha}} \to 0 \quad \text{for } x > 0 \quad \Rightarrow \quad S_X(x) \to 0. \tag{3.6}$$

The distribution collapses to a point mass at $x = 0$. As the mixing parameter $p \to 0^+$:

$$S_X(x) \to 1 - \left(1 - e^{-(\theta x)^{\alpha}}\right)^k. \tag{3.7}$$

The result is the exponentiated Weibull distribution, commonly applied in reliability modeling for independent failure modes. As $p \to 1^-$:

$$\frac{1 - e^{-(\theta x)^{\alpha}}}{1 - pe^{-(\theta x)^{\alpha}}} \to 1 \quad \text{for } x > 0 \quad \Rightarrow \quad S_X(x) \to 0. \tag{3.8}$$

All mass concentrates at $x = 0$. For large k:

$$\left(\frac{1 - e^{-(\theta x)^{\alpha}}}{1 - pe^{-(\theta x)^{\alpha}}}\right)^k \to 0 \quad \text{for } x > 0 \quad \Rightarrow \quad S_X(x) \to 0. \tag{3.9}$$

The distribution approaches a point mass at $x = 0$. As $k \to 0^+$:

$$\left(\frac{1 - e^{-(\theta x)^\alpha}}{1 - pe^{-(\theta x)^\alpha}}\right)^k \to 1 \quad \Rightarrow \quad S_X(x) \to 1. \tag{3.10}$$

Probability mass shifts toward infinity, producing an improper distribution. These limits often lead to degenerate distributions (e.g., point masses at $x = 0$ or $x = 1/\theta$, or improper distributions), highlighting parameter roles but offering limited practical utility. The exponentiated Weibull, obtained as $p \to 0^+$, is particularly valuable for reliability engineering. For practical applications, moderate parameter values, such as $\alpha \in \{0.3, 0.5, 1, 1.5, 2, 2.5\}$, $k \in \{1, 2, 3, 4, 5\}$, and $p = 0.3$, ensure flexible modeling without degeneracy, allowing precise control over tail behavior, as illustrated in Figures 3 and 4.

3.1.2. Associations with other distributions

The WGSG distribution, parameterized by shape $\alpha > 0$, scale $\theta > 0$, weight $0 \leq p < 1$, and order $k \in \mathbb{Z}^+$, reduces to classical distributions used in reliability and survival analysis under specific parameter limits.

Exponentiated Weibull distribution: As $p \to 0^+$, the WGSG becomes the exponentiated Weibull distribution:

$$S_X(x) = 1 - \left(1 - e^{-(\theta x)^\alpha}\right)^k, \quad x \geq 0.$$

For $k = 1$, this is the standard Weibull distribution with shape α and scale θ^{-1}:

$$S_X(x) = e^{-(\theta x)^\alpha}, \quad x \geq 0,$$

$$f_X(x) = \alpha\theta^\alpha x^{\alpha-1} e^{-(\theta x)^\alpha}, \quad x \geq 0.$$

For $k > 1$, the exponentiated Weibull models transformed failure times, extending the Weibull's flexibility. The parameter p enriches hazard and tail behaviors.

Exponential distribution: Setting $\alpha = 1$, $p \to 0$, and $k = 1$, the WGSG simplifies to the exponential distribution:

$$S_X(x) = e^{-\theta x}, \quad x \geq 0,$$

$$f_X(x) = \theta e^{-\theta x}, \quad x \geq 0.$$

This describes memoryless processes, such as inter-event times in a Poisson process, with constant hazard rate θ. For $k > 1$, the survival function $1 - \left(1 - e^{-\theta x}\right)^k$ is non-standard, lacking memorylessness.

Rayleigh distribution: When $\alpha = 2$, $p \to 0$, and $k = 1$, the WGSG reduces to the Rayleigh distribution:

$$S_X(x) = e^{-(\theta x)^2}, \quad x \geq 0,$$

$$f_X(x) = 2\theta^2 x e^{-(\theta x)^2}, \quad x \geq 0.$$

This models life data under accumulated stress, such as wear-related failures, aligning with the Weibull for $\alpha = 2$.

3.2. Moment-generating function (MGF)

The moment-generating function (MGF) of a random variable X is defined as:

$$M_X(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx} f_X(x)\, dx.$$

Substituting this into the definition of $M_X(t)$ gives:

$$M_X(t) = k(1-p)\alpha\theta^\alpha \int_0^\infty e^{tx} x^{\alpha-1} e^{-(\theta x)^\alpha} \frac{\left(1 - e^{-(\theta x)^\alpha}\right)^{k-1}}{\left(1 - pe^{-(\theta x)^\alpha}\right)^{k+1}}\, dx.$$

Using the binomial series expansion,

$$\frac{1}{(1 - pe^{-(\theta x)^\alpha})^{k+1}} = \sum_{i=0}^{\infty} \binom{k+i}{i} p^i e^{-i(\theta x)^\alpha}.$$

$$(1 - e^{-(\theta x)^\alpha})^{k-1} = \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^j e^{-j(\theta x)^\alpha}.$$

Substituting into $M_X(t)$:

$$M_X(t) = k(1-p) \sum_{i=0}^{\infty} \sum_{j=0}^{k-1} \binom{k+i}{i} \binom{k-1}{j} (-1)^j p^i \int_0^\infty e^{tx} x^{\alpha-1} e^{-(i+j+1)(\theta x)^\alpha}\, dx.$$

Using the known integral formula:

$$\int_0^\infty e^{tx} x^{\alpha-1} e^{-c(\theta x)^\alpha}\, dx = \frac{E_{1/\alpha,1}\left(\frac{t}{\theta c^{1/\alpha}}\right)}{c},$$

where $E_{1/\alpha,1}(z)$ is the Mittag-Leffler function defined as:

$$E_{1/\alpha,1}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma\left(\frac{k}{\alpha} + 1\right)}.$$

We set $c = i + j + 1$ and obtain:

$$M_X(t) = k(1-p) \sum_{i=0}^{\infty} \sum_{j=0}^{k-1} \binom{k+i}{i} \binom{k-1}{j} (-1)^j p^n \cdot \frac{E_{1/\alpha,1}\left(\frac{t}{\theta(i+j+1)^{1/\alpha}}\right)}{(i+j+1)}.$$

3.3. r-th Moment

The r-th moment is given by:

$$E(X^r) = \int_0^\infty x^r f_X(x)\, dx.$$

Substituting the PDF:

$$E(X^r) = k(1-p)\alpha\theta^\alpha \int_0^\infty x^{r+\alpha-1} \frac{e^{-(\theta x)^\alpha}\left(1-e^{-(\theta x)^\alpha}\right)^{k-1}}{\left[1-pe^{-(\theta x)^\alpha}\right]^{k+1}}\, dx.$$

Expanding the numerator and the denominator using the binomial series:

$$\frac{1}{(1-pe^{-(\theta x)^\alpha})^{k+1}} = \sum_{j=0}^{\infty} \binom{k+j}{j} p^j e^{-j(\theta x)^\alpha}.$$

$$(1-e^{-(\theta x)^\alpha})^{k-1} = \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^i e^{-i(\theta x)^\alpha}.$$

Thus,

$$E(X^r) = k(1-p)\alpha\theta^\alpha \sum_{j=0}^{\infty} \binom{k+j}{j} p^j \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^i \int_0^\infty x^{r+\alpha-1} e^{-[(i+j+1)(\theta x)^\alpha]}\, dx.$$

Using the substitution $y = (i+j+1)^{\frac{1}{\alpha}}\theta x$ and the gamma function identity:

$$\int_0^\infty y^{r+\alpha-1} e^{-y^\alpha}\, dy = \frac{1}{\alpha}\Gamma\left(\frac{r}{\alpha}+1\right),$$

we obtain:

$$\mu_r = E(X^r) = \frac{k(1-p)}{\theta^r}\Gamma\left(\frac{r}{\alpha}+1\right) \sum_{i=0}^{\infty}\sum_{j=0}^{k-1} \binom{i+k}{i}\binom{k-1}{j} (-1)^j (i+j+1)^{-\left(\frac{r}{\alpha}+1\right)}.$$

Mean:

The first moment, or mean, is given by $E(X) = \mu_1$:

$$E(X) = \frac{k(1-p)}{\theta}\Gamma\left(\frac{1}{\alpha}+1\right) \sum_{i=0}^{\infty}\sum_{j=0}^{k-1} \binom{i+k}{i}\binom{k-1}{j} (-1)^j (i+j+1)^{-\left(\frac{1}{\alpha}+1\right)}.$$

For $k = 1$, the formula simplifies:

$$E(X) = \frac{1-p}{\theta}\Gamma\left(\frac{1}{\alpha}+1\right) \sum_{i=0}^{\infty} (i+1)^{-\left(\frac{1}{\alpha}+1\right)}.$$

This sum involves a generalized Hurwitz zeta function:

$$E(X) = \frac{1-p}{\theta}\Gamma\left(\frac{1}{\alpha}+1\right) \zeta\left(\frac{1}{\alpha}+1\right).$$

Variance:

The variance is given by:

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 .$$

Setting $r = 2$, we get:

$$E(X^2) = \mu_2 = \frac{k(1-p)}{\theta^2}\Gamma\left(\frac{2}{\alpha}+1\right)\sum_{i=0}^{\infty}\sum_{j=0}^{k-1}\binom{i+k}{i}\binom{k-1}{j}(-1)^j(i+j+1)^{-\left(\frac{2}{\alpha}+1\right)}.$$

For $k = 1$, this simplifies to:

$$E(X^2) = \frac{1-p}{\theta^2}\Gamma\left(\frac{2}{\alpha}+1\right)\sum_{i=0}^{\infty}(i+1)^{-\left(\frac{2}{\alpha}+1\right)}.$$

This sum involves a generalized Hurwitz zeta function:

$$E(X^2) = \frac{1-p}{\theta^2}\Gamma\left(\frac{2}{\alpha}+1\right)\zeta\left(\frac{2}{\alpha}+1\right).$$

Thus, the variance $\mathrm{Var}(X)$ is:

$$\mathrm{Var}(X) = \frac{1-p}{\theta^2}\Gamma\left(\frac{2}{\alpha}+1\right)\zeta\left(\frac{2}{\alpha}+1\right) - \left(\frac{1-p}{\theta}\Gamma\left(\frac{1}{\alpha}+1\right)\zeta\left(\frac{1}{\alpha}+1\right)\right)^2 .$$

3.4. Entropy

Entropy is a fundamental concept in information and communication theory, introduced by [36]. It quantifies the average missing information or uncertainty associated with a random source. Shannon defined the entropy of a continuous random variable X as:

$$H(\theta,\alpha,p,k) = \mathbb{E}\left[\log\left(\frac{1}{f_X(x)}\right)\right] = -\int_0^{\infty} f_X(x)\log f_X(x)\,dx,$$

where $\log\left(\frac{1}{f_X(x)}\right)$ is referred to as the uncertainty associated with the outcome x. The probability distribution $f_X(x)$ fully describes the probabilistic characteristics of a random variable. However, when comparing two or more probability distributions, entropy provides a quantitative measure to compare the randomness of different distributions.

For the new family of distributions, the entropy is given by:

$$H(\theta,\alpha,p,k) = -\int_0^{\infty}\frac{k(1-p)\alpha\theta^{\alpha}x^{\alpha-1}e^{-(\theta x)^{\alpha}}\left(1-e^{-(\theta x)^{\alpha}}\right)^{k-1}}{\left(1-pe^{-(\theta x)^{\alpha}}\right)^{k+1}} \times \log\left(\frac{k(1-p)\alpha\theta^{\alpha}x^{\alpha-1}e^{-(\theta x)^{\alpha}}\left(1-e^{-(\theta x)^{\alpha}}\right)^{k-1}}{\left(1-pe^{-(\theta x)^{\alpha}}\right)^{k+1}}\right)dx.$$

Expanding the logarithm term using its properties:

$$\log f_X(x) = \log k + \log(1-p) + \log\alpha + \alpha\log\theta + (\alpha-1)\log x - (\theta x)^\alpha + (k-1)\log\left(1-e^{-(\theta x)^\alpha}\right) - (k+1)\log\left(1-pe^{-(\theta x)^\alpha}\right).$$

Taking expectations, the entropy simplifies to:

$$H(\alpha,\theta,p,k) = -\log k - \log(1-p) - \log\alpha - \alpha\log\theta - (\alpha-1)\mathbb{E}[\log X] + \mathbb{E}[(\theta X)^\alpha] - (k-1)\mathbb{E}[\log(1-e^{-(\theta X)^\alpha})] + (k+1)\mathbb{E}[\log(1-pe^{-(\theta X)^\alpha})],$$

where $\mathbb{E}[\log X]$ captures the spread of the distribution and $\mathbb{E}[(\theta X)^\alpha]$ relates to the moment of X. The last two expectation terms capture the influence of the geometric component and tail behavior.

The partial derivative with respect to the parameters are given by:

$$\frac{\partial H(\alpha,\theta,p,k)}{\partial\theta} = -\frac{\alpha}{\theta} + \alpha\theta^{\alpha-1}\mathbb{E}[X^\alpha] - (k-1)\alpha\theta^{\alpha-1}\mathbb{E}\left[\frac{X^\alpha e^{-(\theta X)^\alpha}}{1-e^{-(\theta X)^\alpha}}\right] + (k+1)p\alpha\theta^{\alpha-1}\mathbb{E}\left[\frac{X^\alpha e^{-(\theta X)^\alpha}}{1-pe^{-(\theta X)^\alpha}}\right],$$

$$\frac{\partial H(\alpha,\theta,p,k)}{\partial\alpha} = -\frac{1}{\alpha} - \log\theta - \mathbb{E}[\log X] + \theta^\alpha\mathbb{E}[X^\alpha\log(\theta X)] - (k-1)\theta^\alpha\mathbb{E}\left[\frac{X^\alpha\log(\theta X)e^{-(\theta X)^\alpha}}{1-e^{-(\theta X)^\alpha}}\right] + (k+1)p\theta^\alpha\mathbb{E}\left[\frac{X^\alpha\log(\theta X)e^{-(\theta X)^\alpha}}{1-pe^{-(\theta X)^\alpha}}\right],$$

$$\frac{\partial H(\alpha,\theta,p,k)}{\partial k} = -\frac{1}{k} - \mathbb{E}\left[\log\left(\frac{1-e^{-(\theta X)^\alpha}}{1-pe^{-(\theta X)^\alpha}}\right)\right],$$

$$\frac{\partial H(\alpha,\theta,p,k)}{\partial p} = -\frac{1}{1-p} - (k+1)\mathbb{E}\left[\frac{e^{-(\theta X)^\alpha}}{1-pe^{-(\theta X)^\alpha}}\right].$$

The residual entropy, also known as conditional entropy in this context, quantifies the uncertainty in a random variable $X_{(k)}$ given that it exceeds a threshold $t \geq 0$. For the WGSG distribution, this corresponds to the differential entropy of the residual lifetime distribution $X_{(k)} - t \mid X_{(k)} > t$. The residual entropy is defined as:

$$H(X_{(k)} \mid X_{(k)} > t) = -\int_t^\infty \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)}\log\left(\frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)}\right)dx, \tag{3.11}$$

where

$$f_{X_{(k)}}(x;\alpha,\theta,p) = \frac{k(1-p)\alpha\theta^{\alpha}x^{\alpha-1}e^{-(\theta x)^{\alpha}}\left(1-e^{-(\theta x)^{\alpha}}\right)^{k-1}}{\left(1-pe^{-(\theta x)^{\alpha}}\right)^{k+1}}, \quad x \geq 0, \tag{3.12}$$

$$S_{X_{(k)}}(x) = 1 - \left(\frac{1-e^{-(\theta x)^{\alpha}}}{1-pe^{-(\theta x)^{\alpha}}}\right)^{k}. \tag{3.13}$$

Applying the logarithmic property $\log\left(\frac{a}{b}\right) = \log a - \log b$, the expression simplifies as follows:

$$\begin{aligned} H(X_{(k)} \mid X_{(k)} > t) = &-\int_t^{\infty} \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)} \log f_{X_{(k)}}(x)\, dx \\ &+ \log S_{X_{(k)}}(t) \int_t^{\infty} \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)}\, dx. \end{aligned} \tag{3.14}$$

Since $\log S_{X_{(k)}}(t)$ is constant with respect to x, the second integral becomes:

$$\int_t^{\infty} \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)} \log S_{X_{(k)}}(t)\, dx = \log S_{X_{(k)}}(t) \int_t^{\infty} \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)} dx. \tag{3.15}$$

Noting that $\int_t^{\infty} f_{X_{(k)}}(x)\, dx = S_{X_{(k)}}(t)$, we have:

$$\int_t^{\infty} \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)} dx = 1, \tag{3.16}$$

$$\int_t^{\infty} \frac{f_{X_{(k)}}(x)}{S_{X_{(k)}}(t)} \log S_{X_{(k)}}(t)\, dx = \log S_{X_{(k)}}(t).$$

Thus:

$$H(X_{(k)} \mid X_{(k)} > t) = \frac{1}{S_{X_{(k)}}(t)} \left(-\int_t^{\infty} f_{X_{(k)}}(x) \log f_{X_{(k)}}(x)\, dx\right) + \log S_{X_{(k)}}(t), \tag{3.17}$$

$$H(X_{(k)} \mid X_{(k)} > t) = \frac{1}{S_{X_{(k)}}(t)} \left(-\int_t^{\infty} f_{X_{(k)}}(x) \log f_{X_{(k)}}(x)\, dx + S_{X_{(k)}}(t) \log S_{X_{(k)}}(t)\right). \tag{3.18}$$

This form highlights two components: the entropy integral over the tail of the distribution (from t to ∞) and a correction term involving the survival probability at time t.

For $t = 0$, the survival function is:

$$S_{X_{(k)}}(0) = 1 - \left(\frac{1-e^{0}}{1-pe^{0}}\right)^{k} = 1,$$

$$S_{X_{(k)}}(0) \log S_{X_{(k)}}(0) = 0.$$

Thus:

$$H(X_{(k)} \mid X_{(k)} > 0) = \frac{1}{1}\left(-\int_0^{\infty} f_{X_{(k)}}(x) \log f_{X_{(k)}}(x)\, dx + 0\right) = -\int_0^{\infty} f_{X_{(k)}}(x) \log f_{X_{(k)}}(x)\, dx,$$

which is the differential entropy of $X_{(k)}$, often referred to as the Shannon entropy in this context:

$$H(X_{(k)}) = -\int_0^\infty f_{X_{(k)}}(x) \log f_{X_{(k)}}(x)\, dx. \tag{3.19}$$

This measure is valuable in reliability engineering and information theory, as it quantifies the uncertainty in a system's remaining lifetime after surviving up to time t, aiding in reliability analysis and decision-making under uncertainty.

4. Parameter estimation

4.1. Maximum likelihood estimation

Let $x = (x_1, x_2, \ldots, x_n)$ be a random sample drawn from the WGSG distribution with the unknown parameter vector $\varphi = (\alpha, \theta, p)$. Using the PDF of the WGSG distribution, the log-likelihood function is given by:

$$\log L(\varphi; x) = \sum_{i=1}^{n} \log f_X(x_i; \varphi).$$

Substituting $f_X(x_i; \varphi)$ into the log-likelihood function, we obtain:

$$\log L(\varphi; x) = \sum_{i=1}^{m} \log \left[\frac{k(1-p)\alpha\theta^\alpha x_i^{\alpha-1} e^{-(\theta x_i)^\alpha}}{(1 - pe^{-(\theta x_i)^\alpha})^{k+1}} \cdot \left(1 - e^{-(\theta x_i)^\alpha}\right)^{k-1} \right].$$

Thus,

$$\begin{aligned} \log L(\varphi; x) = \sum_{i=1}^{n} \Bigg[& \log(k) + \log(1-p) + \log(\alpha) + \alpha \log(\theta) + (\alpha - 1)\log(x_i) - (\theta x_i)^\alpha \\ & - (k+1)\log\left(1 - pe^{-(\theta x_i)^\alpha}\right) + (k-1)\log\left(1 - e^{-(\theta x_i)^\alpha}\right) \Bigg]. \end{aligned}$$

The partial derivatives with respect to α, θ, and p are:

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^{n} \left[\frac{1}{\alpha} + \log(\theta) + \log(x_i) - (\theta x_i)^\alpha \log(\theta x_i) \right],$$

$$\frac{\partial \log L}{\partial \theta} = \sum_{i=1}^{n} \left[\frac{\alpha}{\theta} - \alpha(\theta x_i)^{\alpha-1} x_i \right],$$

$$\frac{\partial \log L}{\partial p} = \sum_{i=1}^{n} \left[-\frac{1}{1-p} + (k+1)\frac{e^{-(\theta x_i)^\alpha}}{1 - pe^{-(\theta x_i)^\alpha}} \right].$$

These nonlinear equations are typically solved numerically using iterative methods such as the Newton-Raphson method, the expectation-maximization (EM) algorithm, and gradient-based optimization (e.g., BFGS).

4.2. EM algorithm

Let the observed data be $\{x_i\}_{i=1}^n$ and the latent variable N, which follows a truncated geometric distribution. The joint PDF is:

$$f_{X,N}(x, n; \alpha, \theta, p, k) = P(N = n) \cdot f_{(k)}(x; \alpha, \theta, n).$$

Thus, the complete data log-likelihood function is:

$$\ell_c(\alpha, \theta, p, k) = \sum_{i=1}^{n} \log\left(\sum_{n=k}^{\infty} P(N = n) f_{(k)}(x_i; \alpha, \theta, n)\right).$$

E-step (Expectation step): In the E-step, we calculate the expected value of the latent variable N given the observed data x_i. The posterior distribution of N given x_i is:

$$P(N = n \mid x_i) = \frac{P(N = n) f_{(k)}(x_i; \alpha, \theta, n)}{\sum_{n=k}^{\infty} P(N = n) f_{(k)}(x_i; \alpha, \theta, n)}.$$

Thus, the expected value of N given x_i is:

$$E(N \mid x_i) = \sum_{n=k}^{\infty} n P(N = n \mid x_i).$$

Substituting for the posterior:

$$E(N \mid x_i) = \sum_{n=k}^{\infty} n \frac{P(N = n) f_{(k)}(x_i; \alpha, \theta, n)}{\sum_{n=k}^{\infty} P(N = n) f_{(k)}(x_i; \alpha, \theta, n)}.$$

Simplifying, we get the following expression:

$$E(N \mid x_i) = k + \frac{(k+1) p e^{-(\theta x_i)^\alpha}}{1 - p e^{-(\theta x_i)^\alpha}}.$$

M-step (Maximization step): In the M-step, we maximize the expected complete-data log-likelihood with respect to the parameters α, θ, and p. We calculate the updates for each parameter.

$$p^{(r+1)} = \frac{\sum O_i^{(r)}}{n + \sum O_i^{(r)}},$$

$$\theta^{(r+1)} = \left[\frac{1}{n} \sum_{i=1}^{n} x_i^{\alpha^{(r+1)}} \left(O_i^{(r)} + 1 - \frac{(k-1) a_i^{(r+1)}}{1 - a_i^{(r+1)}}\right)\right]^{-\frac{1}{\alpha^{(r+1)}}},$$

$$\alpha^{(r+1)} = n \left[\sum_{i=1}^{n} \ln b_i^{(r+1)} \left((O_i^{(r)} + 1)(b_i^{(r+1)})^{\alpha^{(r+1)}} - \frac{(k-1)(b_i^{(r+1)})^{\alpha^{(r+1)}} a_i^{(r+1)}}{1 - a_i^{(r+1)}} - 1\right)\right]^{-1},$$

where

$$a_i^{(r+1)} = e^{-(\theta^{(r+1)} x_i)^{\alpha^{(r+1)}}},$$

$$b_i^{(r+1)} = \theta^{(r+1)} x_i,$$

and

$$O_i^{(r)} = \frac{p^{(r)} e^{-(\theta^{(r)} x_i)^{\alpha^{(r)}}}}{1 - p^{(r)} e^{-(\theta^{(r)} x_i)^{\alpha^{(r)}}}}.$$

4.3. Method of moments estimation

The method of moments (MM) provides consistent estimators for the parameters $\alpha > 0$, $\theta > 0$, and $0 \leq p < 1$ of the WGSG distribution by equating sample moments to their theoretical counterparts. For a fixed-order $k \in \mathbb{Z}^+$, the estimation leverages the first three moments to solve a nonlinear system.

The r-th raw moment of the WGSG distribution is:

$$\mu_r = \frac{k(1-p)}{\theta^r} \Gamma\left(\frac{r}{\alpha} + 1\right) S_r(\alpha, p, k), \tag{4.1}$$

where:

$$S_r(\alpha, p, k) = \sum_{i=0}^{\infty} \sum_{j=0}^{k-1} \binom{i+k}{i} \binom{k-1}{j} (-1)^j (i+j+1)^{-\left(\frac{r}{\alpha}+1\right)},$$

which equals $\mathbb{E}\left[(I+J+1)^{-\left(\frac{r}{\alpha}+1\right)}\right]$, with $I \sim \text{NB}(k, 1-p)$ (negative binomial, failures before k successes, success probability $1-p$) and $J \sim \text{Bin}(k-1, 1/2)$ independent. The series converges absolutely due to the rapid decay of $(i+j+1)^{-\left(\frac{r}{\alpha}+1\right)}$.

For a sample of size n, the r-th sample moment is $\hat{\mu}_r = n^{-1} \sum_{i=1}^{n} x_i^r$. Matching moments for $r = 1, 2, 3$ yields:

$$\hat{\mu}_1 = \frac{k(1-p)}{\theta} M_1(\alpha, p, k), \tag{4.2}$$

$$\hat{\mu}_2 = \frac{k(1-p)}{\theta^2} M_2(\alpha, p, k), \tag{4.3}$$

$$\hat{\mu}_3 = \frac{k(1-p)}{\theta^3} M_3(\alpha, p, k), \tag{4.4}$$

where $M_r(\alpha, p, k) = \Gamma\left(\frac{r}{\alpha} + 1\right) S_r(\alpha, p, k)$. We solve for θ:

$$\theta = \frac{\hat{\mu}_1 M_2(\alpha, p, k)}{\hat{\mu}_2 M_1(\alpha, p, k)}. \tag{4.5}$$

Substituting into Eq (4.2):

$$p = 1 - \frac{\hat{\mu}_1^2 M_2(\alpha, p, k)}{k \hat{\mu}_2 M_1(\alpha, p, k)^2}. \tag{4.6}$$

From Eqs (4.4) and (4.3):

$$\frac{\hat{\mu}_3}{\hat{\mu}_2} = \frac{M_3(\alpha, p, k)}{\theta M_2(\alpha, p, k)}, \tag{4.7}$$

and equating with (4.5):

$$\frac{\hat{\mu}_1 M_2(\alpha, p, k)}{\hat{\mu}_2 M_1(\alpha, p, k)} = \frac{\hat{\mu}_2 M_3(\alpha, p, k)}{\hat{\mu}_3 M_2(\alpha, p, k)}, \tag{4.8}$$

$$\frac{M_2(\alpha, p, k)^2}{M_1(\alpha, p, k) M_3(\alpha, p, k)} = \frac{\hat{\mu}_2^2}{\hat{\mu}_1 \hat{\mu}_3}. \tag{4.9}$$

This nonlinear equation can be solved numerically using a root-finding algorithm, such as the Newton-Raphson method or a bracketing method (e.g., bisection), implemented in software like `R` or `Python`. For instance, in `R`, the `nleqslv` package can be used by defining the following system of equations:

$$h_1(\alpha, p) = p - 1 + \frac{\hat{\mu}_1^2 M_2(\alpha, p, k)}{k \hat{\mu}_2 M_1(\alpha, p, k)^2}, \tag{4.10}$$

$$h_2(\alpha, p) = \frac{M_2(\alpha, p, k)^2}{M_1(\alpha, p, k) M_3(\alpha, p, k)} - \frac{\hat{\mu}_2^2}{\hat{\mu}_1 \hat{\mu}_3}, \tag{4.11}$$

and iteratively adjusting α and p to minimize the functions h_1 and h_2. Appropriate initial guesses (e.g., $\alpha = 1$, $p = 0.5$) and parameter bounds ($\alpha > 0$, $0 \leq p < 1$) help ensure convergence. The function $S_r(\alpha, p, k)$ is approximated by truncating the infinite series at a sufficiently large upper index $i_{\max} \approx 1000$, where additional terms become negligible. Once α is estimated, the remaining parameters θ and p are computed from Eqs (4.5) and (4.6), respectively.

4.3.1. Bayesian estimation

Bayesian estimation provides a robust framework for estimating the parameters $\alpha > 0$, $\theta > 0$, and $0 \leq p < 1$ of the WGSG distribution, leveraging prior knowledge to yield full posterior distributions, while treating the fixed parameter $k \in \mathbb{N}$ as a known value that defines a distinct distribution for each integer. For a random sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ from the WGSG distribution with the PDF given in Eq (2.1), the likelihood function is:

$$L(\alpha, \theta, p \mid \mathbf{x}) = \prod_{i=1}^{n} f_{X_{(k)}}(x_i \mid \alpha, \theta, p).$$

Independent priors are assigned to reflect prior beliefs: $\alpha \sim \text{Gamma}(a_\alpha, b_\alpha)$, $\theta \sim \text{Gamma}(a_\theta, b_\theta)$, and $p \sim \text{Beta}(a_p, b_p)$, chosen for their domain compatibility and computational tractability, with hyperparameters (a_α, b_α), (a_θ, b_θ), and (a_p, b_p) set to weakly informative values (e.g., $a_\alpha = b_\alpha = 0.001$) to approximate noninformative priors when prior knowledge is limited. The posterior distribution is defined as:

$$\pi(\alpha, \theta, p \mid \mathbf{x}) = \frac{L(\alpha, \theta, p \mid \mathbf{x}) \pi(\alpha, \theta, p)}{\int_0^\infty \int_0^\infty \int_0^1 L(\alpha, \theta, p \mid \mathbf{x}) \pi(\alpha, \theta, p) \, d\alpha \, d\theta \, dp}.$$

The posterior expected loss is:

$$\int L((\alpha, \theta, p), a) \pi(\alpha, \theta, p \mid \mathbf{x}) \, d\alpha \, d\theta \, dp.$$

The joint posterior density, proportional to the likelihood times the priors, is:

$$\pi(\alpha, \theta, p \mid \mathbf{x}) \propto L(\alpha, \theta, p \mid \mathbf{x}) \cdot \pi(\alpha) \cdot \pi(\theta) \cdot \pi(p),$$

but its complexity, driven by the WGSG's nested exponential and polynomial terms, prevents a closed-form solution, necessitating Markov Chain Monte Carlo (MCMC) methods like the Metropolis-Hastings algorithm or Hamiltonian Monte Carlo, implemented via tools such as `Stan` or `JAGS`, to approximate posterior summaries. Under mean square error (MSE) loss, the Bayes estimator for a parameter $\varphi \in \{\alpha, \theta, p\}$, denoted $\hat{\varphi}_{\text{Bayes}}$, is the posterior mean:

$$\hat{\varphi}_{\text{Bayes}} = \mathbb{E}[\varphi \mid \mathbf{x}] = \int_0^\infty \int_0^\infty \int_0^1 \varphi \pi(\alpha, \theta, p \mid \mathbf{x})\, d\alpha\, d\theta\, dp,$$

which minimizes the expected posterior loss and is approximated using MCMC samples, with analogous expressions for $\hat{\alpha}_{\text{Bayes}}$, $\hat{\theta}_{\text{Bayes}}$, and $\hat{p}_{\text{Bayes}}$.

5. Simulation study

To evaluate the performance of the maximum likelihood (ML), method of moments (MM), and Bayesian estimation methods for the WGSG distribution, we conducted a simulation study. We generated 1000 random samples of sizes $n = 20$, 50, and 100 from the WGSG distribution, considering three parameter settings for $\varphi = (\alpha, \theta, p)$: $(0.5, 1, 0.5)$, $(1, 1, 0.5)$, and $(2, 1, 0.5)$, combined with $k \in \{1, 2, 3\}$. The simulations were implemented using the R programming language.

For each sample, we estimated the parameters $\widehat{\varphi} = (\widehat{\alpha}, \widehat{\theta}, \widehat{p})$ and computed the bias and mean squared error (MSE) of the estimators. The bias and MSE are defined as:

$$\text{Bias}(\widehat{\varphi}) = E(\widehat{\varphi}) - \varphi,$$

$$\text{MSE}(\widehat{\varphi}) = E[(\widehat{\varphi} - \varphi)^2] = [\text{Bias}(\widehat{\varphi})]^2 + \text{Var}(\widehat{\varphi}),$$

where $\text{Var}(\widehat{\varphi})$ is the variance of the estimator. An estimator $\widehat{\varphi}_1$ is considered more efficient than $\widehat{\varphi}_2$ if $\text{MSE}(\widehat{\varphi}_1) < \text{MSE}(\widehat{\varphi}_2)$.

The results are summarized in Table 1, which reports the bias and mean squared error (MSE) of the parameter estimates for the ML, MM, and Bayesian methods across all scenarios. The findings reveal that for each order statistic $k \in \{1, 2, 3\}$, the MSEs decrease as the sample size n increases, indicating improved estimation precision with larger samples. Moreover, the Bayesian method consistently exhibits lower MSEs compared to the ML and MM approaches, suggesting greater efficiency under these simulation settings. As the sample size increases to $n = 100$, the bias of the ML estimators tends to decrease, aligning with their known property of asymptotic unbiasedness. However, the MSE results suggest that ML estimation may struggle in small samples when applied to the WGSG distribution, primarily due to the sensitivity of numerical optimization routines to initial values and the complexity of the likelihood surface. In contrast, the Bayesian approach employs Metropolis-Hastings MCMC sampling to generate posterior samples, with parameters estimated using posterior summaries such as the mean or median. While the Bayesian estimators may be influenced by the choice of prior and exhibit some bias, they often achieve lower MSE by effectively regularizing parameter estimates, especially in small-sample scenarios where likelihood-based estimates may be unstable. The Bayesian

MSE for $\widehat{p}$ is notably low (e.g., 0.004 at $n = 20, k = 2, \alpha = 2$) compared to ML (0.142) and MM (0.038), underlining its superior efficiency. Similarly, for $\widehat{\alpha}$, the Bayesian bias is substantially lower (e.g., -0.592 at $n = 20, k = 3, \alpha = 2$) than the corresponding ML bias (1.957), demonstrating the stabilizing effect of the prior. Furthermore, Bayesian estimation shows greater robustness to changes in k and α: for example, the MSE for $\widehat{\alpha}$ increases only modestly from 0.021 at $k = 1, \alpha = 0.5$ to 0.414 at $k = 3, \alpha = 2$, whereas ML and MM estimators exhibit more pronounced increases.

The MM approach estimates parameters by equating the theoretical moments of the WGSG distribution to their sample counterparts and solving for α, θ, and p. This method is computationally simpler than ML and Bayesian approaches, as it avoids complex likelihood evaluations and MCMC sampling. However, it tends to be less reliable for distributions with intricate moment structures such as the WGSG, which are sensitive to sample variability and outliers. The simulation results reflect this limitation: the MM approach consistently yields the highest bias and MSE across all parameters. For example, at $n = 20, k = 3, \alpha = 0.5$, the MM bias for $\widehat{\alpha}$ is 1.490, and the MSE is 2.863—substantially worse than ML (bias: 0.487, MSE: 0.323) and Bayesian (bias: -0.067, MSE: 0.014) results in the same setting. Even at larger sample sizes ($n = 100$), the MM estimator remains biased (e.g., 1.264 at $k = 3, \alpha = 0.5$), suggesting limited improvement with increased data. This underperformance likely stems from the difficulty of accurately estimating higher-order moments in the WGSG distribution. Although ML estimation offers favorable large-sample properties, its finite-sample performance can be inferior to the Bayesian approach. The Bayesian method provides more stable and accurate estimates across various parameter settings and sample sizes. The MM approach, while computationally attractive, proves to be the least reliable among the three, particularly in small-sample contexts and in the presence of complex moment behavior.

Table 1. Comparison of ML, Bayesian, and MM estimation for the WGSG distribution ($\alpha \in \{0.5, 1, 2\}$; $\theta = 1$; $p = 0.5$).

			ML						Bayesian						MM					
			Bias			MSE			Bias			MSE			Bias			MSE		
n	k	α	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{p}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{p}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{p}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{p}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{p}$	$\widehat{\alpha}$	$\widehat{\theta}$	$\widehat{p}$
20	1	0.5	0.018	0.678	-0.087	0.014	4.052	0.144	0.099	-0.307	-0.078	0.021	0.144	0.010	0.886	-0.889	0.099	1.408	0.798	0.049
		1	0.090	0.092	-0.068	0.078	0.402	0.145	0.157	-0.236	-0.058	0.067	0.081	0.006	0.623	-0.712	0.158	1.260	0.565	0.078
		2	0.199	-0.003	-0.061	0.335	0.087	0.144	0.169	-0.152	-0.067	0.157	0.032	0.007	-0.580	-0.172	0.066	0.812	0.081	0.029
	2	0.5	0.264	-0.696	0.072	0.111	0.558	0.146	-0.028	0.656	0.220	0.009	0.600	0.055	1.311	-0.920	0.242	2.462	0.849	0.120
		1	0.580	-0.493	0.090	0.514	0.302	0.145	-0.115	0.805	0.096	0.041	0.741	0.015	1.098	-0.706	0.298	2.119	0.543	0.145
		2	1.292	-0.287	0.087	2.463	0.112	0.142	-0.394	0.541	-0.022	0.231	0.324	0.004	-0.441	-0.206	0.087	0.902	0.085	0.038
	3	0.5	0.487	-0.787	0.110	0.323	0.645	0.145	-0.067	0.692	0.375	0.014	0.679	0.142	1.490	-0.930	0.313	2.863	0.867	0.155
		1	0.974	-0.568	0.108	1.308	0.355	0.150	-0.184	1.269	0.266	0.062	1.842	0.076	1.213	-0.671	0.321	2.404	0.530	0.160
		2	1.957	-0.355	0.106	5.240	0.145	0.150	-0.592	1.113	0.048	0.414	1.292	0.007	-0.304	-0.312	0.120	1.194	0.134	0.051
50	1	0.5	-0.010	0.521	-0.108	0.006	1.951	0.133	0.073	-0.360	-0.103	0.010	0.158	0.015	0.941	-0.935	0.130	1.456	0.876	0.065
		1	0.028	0.051	-0.074	0.035	0.254	0.129	0.141	-0.253	-0.078	0.038	0.076	0.010	0.697	-0.682	0.214	1.293	0.516	0.100
		2	0.059	-0.004	-0.077	0.140	0.061	0.129	0.215	-0.143	-0.086	0.109	0.025	0.012	-0.414	-0.125	0.075	0.626	0.053	0.037
	2	0.5	0.240	-0.727	0.131	0.078	0.580	0.125	-0.040	0.935	0.230	0.006	1.119	0.057	1.302	-0.946	0.286	2.219	0.896	0.142
		1	0.511	-0.527	0.150	0.352	0.322	0.133	-0.148	1.069	0.056	0.033	1.237	0.009	1.088	-0.695	0.330	1.908	0.542	0.162
		2	1.076	-0.318	0.149	1.535	0.122	0.128	-0.411	0.611	-0.073	0.206	0.393	0.009	-0.368	-0.183	0.096	0.749	0.065	0.042
	3	0.5	0.435	-0.790	0.132	0.234	0.649	0.137	-0.075	1.018	0.382	0.009	1.398	0.147	1.382	-0.940	0.355	2.299	0.892	0.176
		1	0.852	-0.568	0.115	0.928	0.354	0.148	-0.223	1.770	0.241	0.060	3.404	0.063	1.174	-0.683	0.347	2.055	0.538	0.174
		2	1.751	-0.362	0.133	3.868	0.148	0.147	-0.639	1.301	-0.008	0.437	1.746	0.005	-0.027	-0.313	0.169	1.235	0.125	0.070
100	1	0.5	-0.014	0.326	-0.071	0.004	0.970	0.106	0.060	-0.368	-0.126	0.006	0.153	0.021	1.026	-0.949	0.176	1.591	0.902	0.088
		1	0.012	0.021	-0.054	0.021	0.162	0.105	0.124	-0.242	-0.100	0.025	0.068	0.016	0.652	-0.652	0.217	1.139	0.485	0.100
		2	0.030	-0.012	-0.052	0.086	0.041	0.104	0.203	-0.135	-0.114	0.076	0.021	0.019	-0.331	-0.103	0.070	0.514	0.040	0.038
	2	0.5	0.227	-0.745	0.167	0.063	0.588	0.097	-0.050	1.313	0.208	0.004	2.071	0.047	1.256	-0.943	0.322	1.974	0.898	0.159
		1	0.498	-0.547	0.193	0.307	0.332	0.118	-0.179	1.337	-0.013	0.037	1.894	0.008	1.053	-0.695	0.347	1.732	0.529	0.167
		2	1.021	-0.333	0.191	1.278	0.127	0.115	-0.453	0.670	-0.120	0.225	0.464	0.019	-0.357	-0.153	0.088	0.652	0.052	0.040
	3	0.5	0.434	-0.813	0.189	0.217	0.678	0.119	-0.080	1.449	0.375	0.008	2.671	0.141	1.264	-0.928	0.389	1.898	0.878	0.191
		1	0.827	-0.578	0.149	0.832	0.361	0.139	-0.247	2.388	0.190	0.066	6.091	0.042	1.073	-0.649	0.329	1.786	0.515	0.169
		2	1.681	-0.366	0.160	3.416	0.148	0.138	-0.683	1.450	-0.073	0.480	2.148	0.010	0.072	-0.301	0.186	1.226	0.120	0.076

6. Applications

In this section, we compare the adequacy of the WGSG distribution for 5 special cases ($k = 1, 2, 3, 4, 5$) compared with the Weibull, Weibull-geometric (WG), and Weibull-Poisson (WP) distributions using two real datasets. The first data, sourced from [26], represent the fatigue life (rounded to the nearest thousand cycles) for 67 specimens of Alloy T7987 that failed before accumulating 300 thousand cycles. The dataset is as follows:

(94, 118, 139, 159, 171, 189, 227, 96, 121, 140, 159, 172, 190, 256, 99, 121, 141, 159, 173, 196, 257, 99, 123, 141, 159, 176, 197, 269, 104, 129, 143, 162, 177, 203, 271, 108, 131, 144, 168, 180, 205, 274, 112, 133, 149, 168, 180, 211, 291, 114, 135, 149, 169, 184, 213, 117, 136, 152, 170, 187, 224, 117, 139, 153, 170, 188, 226).

The second dataset represents the time intervals (in hours) between successive failures of the air conditioning systems in a fleet of seven Boeing 720 airplanes, specifically aircraft numbers 7910, 7911, 7912, 7913, 7914, 7915, and 7916. This subset, consisting of 125 observations, was analyzed by [37]. It is part of a larger dataset of 213 observations, originally studied by [38] and further explored by [10, 12, 39, 40]. The 125 observations for the specified aircraft are as follows:

(74, 57, 48, 29, 502, 12, 70, 21, 29, 386, 59, 27, 153, 26, 326, 55, 320, 56, 104, 220, 239, 47, 246, 176, 182, 33, 15, 104, 35, 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95, 97, 51, 11, 4, 141, 18, 142, 68, 77, 80, 1, 16, 106, 206, 82, 54, 31, 216, 46, 111, 39, 63, 18, 191, 18, 163, 24, 50, 44, 102, 72, 22, 39, 3, 15, 197, 188, 79, 88, 46, 5, 5, 36, 22, 139, 210, 97, 30, 23, 13, 14, 359, 9, 12, 270, 603, 3, 104, 2, 438, 50, 254, 5, 283, 35, 12).

Weibull distribution:

$$f_W(x;\alpha,\theta) = \alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}, \quad x \geq 0.$$

$$F_W(x;\alpha,\theta) = 1 - e^{-(\theta x)^\alpha}, \quad x \geq 0.$$

Weibull-geometric (WG) distribution:

$$f_{WG}(x;\alpha,\theta,p) = (1-p)\alpha\theta^\alpha x^{\alpha-1}e^{-(\theta x)^\alpha}\left(1 - pe^{-(\theta x)^\alpha}\right)^{-2}, \quad x \geq 0.$$

$$F_{WG}(x;\alpha,\theta,p) = \frac{1 - e^{-(\theta x)^\alpha}}{1 - pe^{-(\theta x)^\alpha}}, \quad x \geq 0.$$

Weibull-Poisson (WP) distribution:

$$f_{WP}(x;\alpha,\theta,\lambda) = \frac{\alpha\theta^\alpha\lambda x^{\alpha-1}}{1-e^{-\lambda}}e^{-(\theta x)^\alpha}e^{-\lambda\left(1-e^{-(\theta x)^\alpha}\right)}, \quad x \geq 0.$$

$$F_{WP}(x;\alpha,\theta,\lambda) = \frac{e^{(\lambda\exp(-(x\theta)^\alpha)} - e^\lambda}{1-e^\lambda}, \quad x \geq 0.$$

Table 2. Estimated parameters for fatigue life data and air conditioning system failures.

Model	Fatigue Life Data	Air Conditioning System Failures
Weibull	$(\alpha = 3.726, \theta = 0.005)$	$(\alpha = 0.863, \theta = 0.012)$
WG	$(\alpha = 6.036, \theta = 0.004, p = 0.952)$	$(\alpha = 1.990, \theta = 0.006, p = 0.714)$
WP	$(\alpha = 4.552, \theta = 0.004, \lambda = 3.067)$	$(\alpha = 0.977, \theta = 0.007, \lambda = 1.547)$
WGSG ($k = 1$)	$(\alpha = 6.036, \theta = 0.004, p = 0.952)$	$(\alpha = 1.990, \theta = 0.006, p = 0.714)$
WGSG ($k = 2$)	$(\alpha = 4.737, \theta = 0.004, p = 0.940)$	$(\alpha = 0.752, \theta = 0.012, p = 0.607)$
WGSG ($k = 3$)	$(\alpha = 2.740, \theta = 0.006, p = 0.478)$	$(\alpha = 0.607, \theta = 0.022, p = 0.528)$
WGSG ($k = 4$)	$(\alpha = 3.955, \theta = 0.005, p = 0.944)$	$(\alpha = 0.539, \theta = 0.032, p = 0.527)$
WGSG ($k = 5$)	$(\alpha = 2.285, \theta = 0.007, p = 0.488)$	$(\alpha = 0.502, \theta = 0.041, p = 0.555)$

Table 3. Model comparison results for fatigue life data and air conditioning system failures.

Model	Log-Lik.	AIC	BIC	Entropy	KS Stat.	KS p-value
			Fatigue Life Data			
Weibull	-353.292	710.584	714.993	5.273	0.097	0.519
WG	-348.552	703.104	709.718	5.202	0.053	0.988
WP	-350.631	707.262	713.876	5.233	0.059	0.961
WGSG ($k = 1$)	-348.552	703.104	709.718	5.202	0.053	0.988
WGSG ($k = 2$)	-347.671	701.341	707.955	5.189	0.061	0.952
WGSG ($k = 3$)	-349.124	704.248	710.862	5.211	0.066	0.916
WGSG ($k = 4$)	-347.367	700.735	707.349	5.185	0.063	0.941
WGSG ($k = 5$)	-348.550	703.100	709.714	5.202	0.057	0.974
			Air Conditioning System Failures			
Weibull	-687.632	1379.265	1384.921	5.501	0.036	0.846
WG	-686.117	1378.233	1386.718	5.489	0.046	0.946
WP	-686.693	1379.387	1387.872	5.494	0.046	0.946
WGSG ($k = 1$)	-686.117	1378.233	1386.718	5.489	0.046	0.946
WGSG ($k = 2$)	-685.279	1376.557	1385.042	5.482	0.044	0.956
WGSG ($k = 3$)	-686.261	1378.522	1387.007	5.490	0.046	0.956
WGSG ($k = 4$)	-686.565	1379.131	1387.615	5.493	0.042	0.976
WGSG ($k = 5$)	-686.839	1379.678	1388.163	5.495	0.045	0.947

Table 2 presents the estimated parameters for various lifetime models fitted to the fatigue life data and air conditioning system failure data. The WGSG distribution, evaluated across different order statistics k, demonstrates flexibility in capturing varying data characteristics through changes in its parameters. For the numerical computation of maximum likelihood estimation, we use the L-BFGS-B method, as proposed by [41]. This method is a limited-memory variation of the BFGS quasi-Newton approach, applied when solving a system of simultaneous equations. The maximum likelihood estimates (MLEs) of the parameters, along with the p-values of the K-S statistics for these models, are summarized in Table 3. The results also include the AIC, BIC, and Shannon's entropy

as model selection criteria. It is important to note that the Kolmogorov-Smirnov (K-S) test compares an empirical distribution with a known (non-estimated) distribution. This test helps determine if a sample originates from a population that follows a specific distribution (H_0: the data follow a specified distribution).

The analysis of the fatigue life data reveals that the WGSG distribution with $k = 4$ provides the best fit among all tested models, outperforming the Weibull, Weibull-geometric (WG), and Weibull-Poisson (WP) distributions across all key metrics. Specifically, the WGSG ($k = 4$) model achieves the highest log-likelihood (−347.367), indicating the best overall fit to the data. It also records the lowest Akaike Information Criterion (AIC = 700.735) and Bayesian Information Criterion (BIC = 707.349), which balance goodness-of-fit with model complexity, reinforcing its superiority. Furthermore, it exhibits the lowest entropy (5.185), suggesting it offers the most precise representation of the data's underlying distribution. The Kolmogorov-Smirnov (K-S) test supports the adequacy of all models, with p-values exceeding 0.05 (WGSG $k = 4$: 0.941), indicating that none can be rejected based on fit to the empirical distribution.

Notably, the WG and WGSG ($k = 1$) models, which are mathematically equivalent, achieve the highest K-S p-value (0.988) and perform exceptionally well, with a log-likelihood of −348.552, AIC of 703.104, and BIC of 709.718. This excellent fit reflects the WGSG's reduction to the WG distribution when $k = 1$. However, the WGSG ($k = 4$) edges out these models by further optimizing the AIC, BIC, and entropy, suggesting it captures additional nuances in the data. In contrast, the standard Weibull model performs the worst, with the lowest log-likelihood (−353.292), highest AIC (710.584), highest BIC (714.993), and highest entropy (5.273). This indicates that the Weibull distribution is too simplistic to fully model the complexity of fatigue life data.

For the air conditioning system failures dataset, the WGSG distribution with $k = 2$ emerges as the best-fitting model, surpassing the Weibull, WG, and WP distributions across most metrics. It achieves the highest log-likelihood (−685.279), lowest AIC (1376.557), and lowest BIC (1385.042), indicating an optimal balance of fit and complexity. Additionally, it has the lowest entropy (5.482), suggesting a precise representation of the failure data. The K-S test yields a high p-value (0.956), confirming the model's adequacy and inability to be rejected.

The WG and WGSG ($k = 1$) models, again identical, perform strongly with a log-likelihood of −686.117, AIC of 1378.233, BIC of 1386.718, and a K-S p-value of 0.946, closely trailing the WGSG ($k = 2$). Meanwhile, the WGSG ($k = 4$) achieves the highest K-S p-value (0.976), indicating an excellent fit to the empirical distribution, though its AIC (1379.131) and BIC (1387.615) are slightly higher than those of WGSG ($k = 2$). The standard Weibull model again performs the worst, with a log-likelihood of −687.632, AIC of 1379.265, BIC of 1384.921, and entropy of 5.501, underscoring its inadequacy for capturing the complexity of air conditioning failure times. These results suggest that the WGSG model is particularly well-suited for modeling air conditioning system failures, offering practical value for maintenance scheduling and system reliability predictions. As with the fatigue life data, the increased complexity of the WGSG distribution requires careful validation to ensure generalizability.

To explore the empirical behaviors that the failure rate function may exhibit, we use the graphical method based on the total time on test (TTT) plot, introduced by [42]. In its empirical form, the TTT

plot is constructed using the values $r|n$ and $G(r|n)$, where:

$$G(r|n) = \frac{\sum_{i=1}^{r} X_{(i:n)} + (n-r)X_{(r:n)}}{\sum_{i=1}^{n} X_{(i:n)}}.$$

Here, $r = 1, \ldots, n$, and $X_{(i:n)}$ denotes the order statistics of the sample. The TTT plot can take different forms, resembling the Gini index, and serves as a rough indicator of the shape of the failure rate function. Specifically, when the curve approaches a straight diagonal, a constant failure rate is appropriate, suggesting the data follow an exponential distribution. A convex curve indicates the data are likely from a decreasing failure rate (DFR) distribution.

The empirical and theoretical distribution fits are visually evaluated using Figures 5 and 6. Figure 5 compares the empirical PDFs and CDFs with their theoretical counterparts for both the fatigue life and air conditioning system failures datasets. For the fatigue life data, the WGSG distribution with $k = 4$ closely matches the empirical PDF and CDF, reinforcing the numerical evidence of its superior performance. Similarly, for the air conditioning system failures data, the WGSG ($k = 2$) provides a precise fit to the empirical distributions, particularly in the PDF, where it captures the failure time distribution's shape more accurately than the Weibull, WG, or WP distributions.

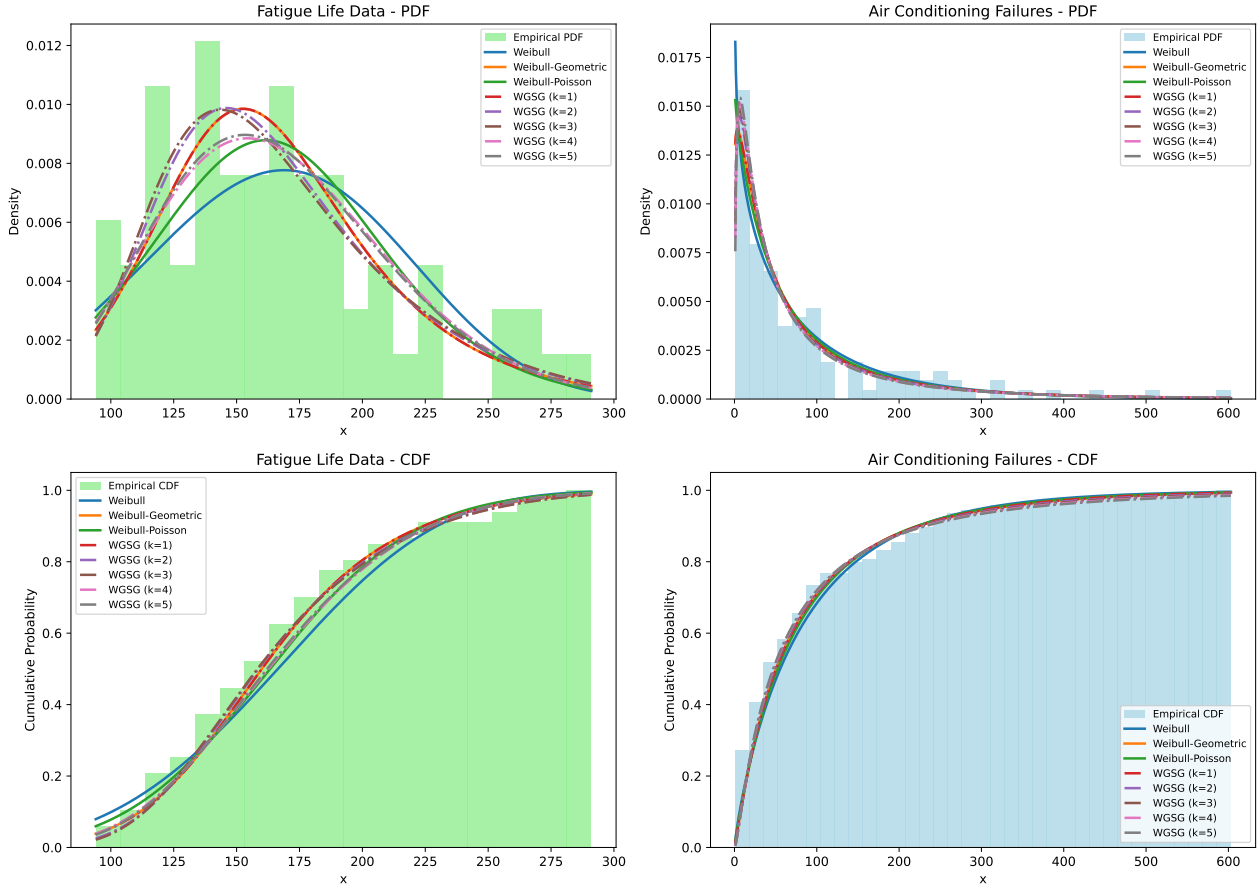

Figure 5. Comparison of empirical and theoretical distributions.

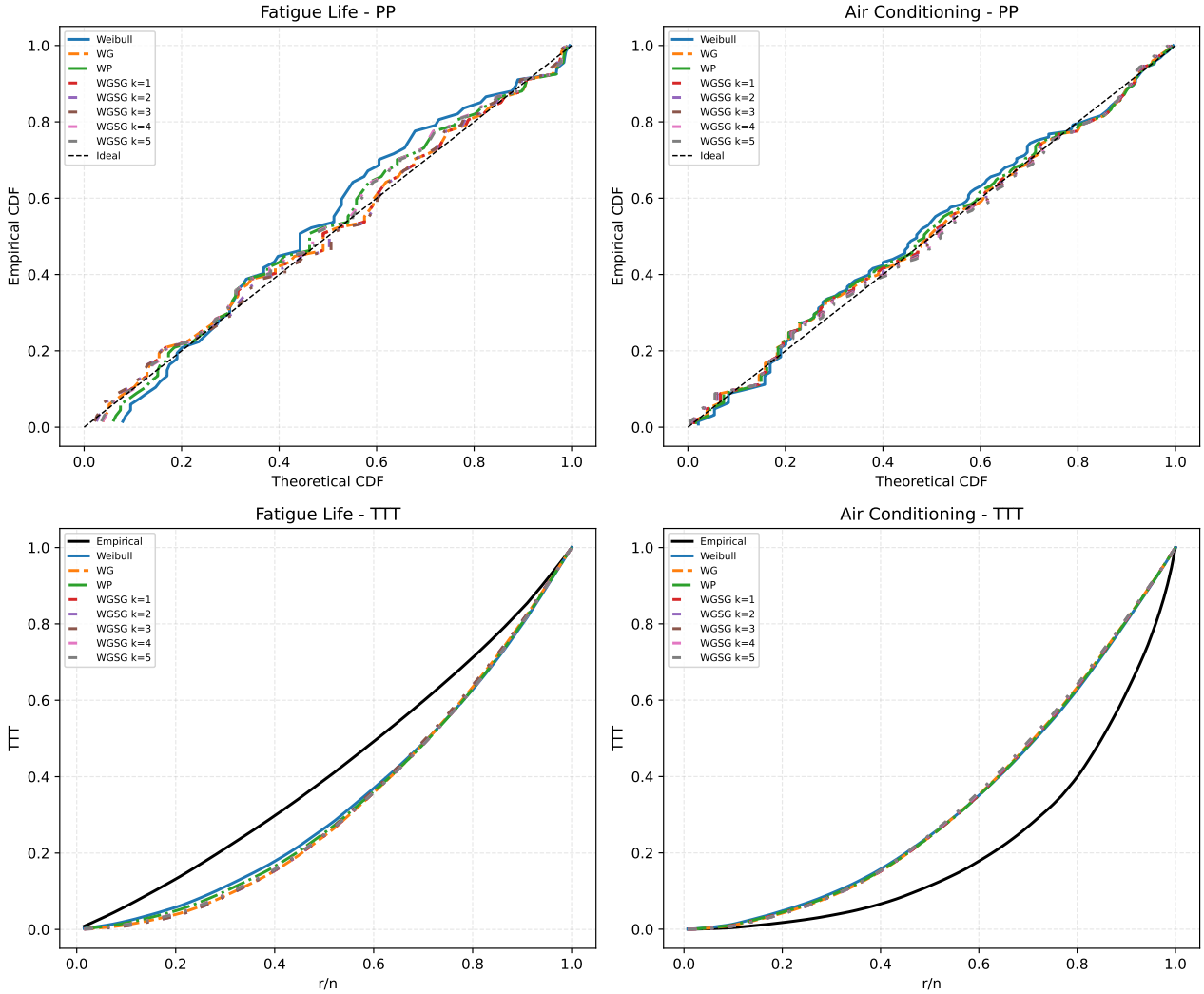

Figure 6. Probability-probability (PP) and TTT plots for all distributions.

Figure 6 presents probability-probability (PP) and TTT plots for all distributions. The PP plots for both datasets show that the WGSG distribution lies closest to the diagonal line, indicating strong agreement between the theoretical and empirical distributions. The TTT plots offer additional insight into the failure rate behavior, which is explored further in the next section. For the fatigue life data, the TTT plot exhibits a convex shape, suggesting a decreasing failure rate (DFR), which the WGSG ($k = 4$) effectively captures, as evidenced by its superior numerical and visual fit. Likewise, the TTT plot for the air conditioning system failures data displays a convex shape, indicating a DFR, which the WGSG ($k = 2$) accurately reflects through its strong fit metrics and close alignment in the PP plot. These findings underscore the WGSG distribution's suitability for both datasets, particularly its ability to accommodate decreasing failure rate behaviors.

Across both datasets, the WGSG distribution consistently outperforms simpler alternatives like the Weibull, WG, and WP distributions, with the optimal k value varying by dataset ($k = 4$ for fatigue life data and $k = 2$ for air conditioning system failures). This adaptability highlights the WGSG's

capacity to balance fit and complexity, making it more effective than traditional distributions for diverse data structures. While the Weibull distribution remains statistically adequate (K-S p-values > 0.05), it consistently underperforms, suggesting it may be too rigid for reliability and survival data with heterogeneous failure patterns. The WGSG's superior performance, supported by both numerical metrics and graphical analysis, has significant implications for lifetime and reliability engineering, where precise failure time analysis can enhance design, maintenance, and safety practices.

7. Conclusions

The WGSG distribution offers a versatile framework for lifetime data analysis, integrating Weibull and shifted geometric properties to model diverse failure patterns with precision. Its unique focus on the k-th order statistic enhances flexibility beyond models like the Weibull-geometric, capturing varying hazard rates and tail behaviors, as detailed in theoretical analyses of entropy and asymptotic properties. Simulations validate robust parameter estimation via maximum likelihood estimation, Expectation-Maximization, Bayesian methods, and the method of moments, while applications to real datasets demonstrate superior fit compared to Weibull, Weibull-Poisson, and Weibull-geometric distributions. This adaptability positions the WGSG as a valuable tool for reliability and risk assessment. Future work will extend it to censored data, multivariate structures, and alternative methods like minimum distance estimation and generalized method of moments, broadening its statistical impact.

Author contributions

M. Rahmouni: Conceptualization, visualization, methodology, writing—original draft, writing—review and editing, formal analysis, validation, investigation, software; D. Ziedan: visualization, data curation, writing—review, investigation. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

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

Funding

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU251549].

Conflict of interest

The authors declare no conflicts of interest in this paper.

References

1. G. M. Cordeiro, R. B. Silva, A. D. C. Nascimento, *Recent advances in lifetime and reliability models*, Bentham Science Publishers, 2020. https://doi.org/10.2174/9781681083452120010l
2. M. Imran, N. Alsadat, M. H. Tahir, F. Jamal, M. Elgarhy, H. Ahmad, et al., The development of an extended Weibull model with applications to medicine, industry and actuarial sciences, *Sci. Rep.*, **14** (2024), 1–24. https://doi.org/10.1038/s41598-024-61308-8
3. K. Balakrishnan, *The exponential distribution: theory, methods and applications*, 1 Ed., Routledge, 1995. https://doi.org/10.1201/9780203756348
4. R. E. Barlow, F. Proschan, *Statistical theory of reliability and life testing: probability models*, Holt, Rinehart and Winston, New York, 1975.
5. S. K. Sinha, B. K. Kale, *Life testing and reliability estimation*, New Delhi: Wiley Eastern ltd., 1980.
6. M. Goldoust, S. Rezaei, Y. Si, S. Nadarajah, Lifetime distributions motivated by series and parallel structures, *Commun. Stat.-Simul. Comput.*, **48** (2019), 556–579. https://doi.org/10.1080/03610918.2017.1390122
7. T. Goyal, S. K. Maurya, S. Nadarajah, Geometric generated family of distributions: a review, *Braz. J. Probab. Stat.*, **35** (2021), 442–465. https://doi.org/10.1214/20-bjps485
8. S. K. Maurya, S. Nadarajah, Poisson generated family of distributions: a review, *Sankhya B*, **83** (2021), 484–540. https://doi.org/10.1007/s13571-020-00237-8
9. M. H. Tahir, G. M. Cordeiro, Compounding of distributions: a survey and new generalized classes, *J. Stat. Distrib. Appl.*, **3** (2016), 13. https://doi.org/10.1186/s40488-016-0052-1
10. K. Adamidis, S. Loukas, A lifetime distribution with decreasing failure rate, *Stat. Probabil. Lett.*, **39** (1998), 35–42. https://doi.org/10.1016/s0167-7152(98)00012-1
11. K. Adamidis, T. Dimitrakopoulou, S. Loukas, On an extension of the exponential-geometric distribution, *Stat. Probabil. Lett.*, **73** (2005), 259–269. https://doi.org/10.1016/j.spl.2005.03.013
12. C. Kuş, A new lifetime distribution, *Comput. Stat. Data Anal.*, **51** (2007), 4497–4509. https://doi.org/10.1016/j.csda.2006.07.017
13. M. Bourguignon, R. B. Silva, G. M. Cordeiro, A new class of fatigue life distributions, *J. Stat. Comput. Simul.*, **84** (2014), 2619–2635. https://doi.org/10.1080/00949655.2013.799164
14. V. G. Cancho, F. Louzada-Neto, G. D. C. Barriga, The poisson-exponential lifetime distribution, *Comput. Stat. Data Anal.*, **55** (2011), 677–686. https://doi.org/10.1016/j.csda.2010.05.033
15. V. Pappas, K. Adamidis, S. Loukas, A generalization of the exponential-logarithmic distribution, *J. Stat. Theory Pract.*, **9** (2015), 122–133. https://doi.org/10.1080/15598608.2014.898604
16. G. C. Rodrigues, F. Louzada, P. L. Ramos, Poisson-exponential distribution: different methods of estimation, *J. Appl. Stat.*, **45** (2018), 128–144. https://doi.org/10.1080/02664763.2016.1268571
17. L. Tafakori, A. Pourkhanali, S. Nadarajah, A new lifetime model with different types of failure rate, *Commun. Stat.-Theory Methods*, **47** (2018), 4006–4020. https://doi.org/10.1080/03610926.2017.1367811

18. M. Rahmouni, A. Orabi, The exponential-generalized truncated geometric (EGTG) distribution: a new lifetime distribution, *Int. J. Stat. Probabil.*, **7** (2018), 1–20. https://doi.org/10.5539/ijsp.v7n1p1

19. M. Rahmouni, A. Orabi, A generalization of the exponential-logarithmic distribution for reliability and life data analysis, *Life Cycle Reliab. Saf. Eng.*, **7** (2018), 159–171. https://doi.org/10.1007/s41872-018-0049-5

20. A. E. Sarhan, B. G. Greenberg, *Contributions to order statistics*, Wiley New York, 1962.

21. P. Pledger, F. Proschan, Comparisons of order statistics and of spacings from heterogeneous distributions, In: J. S. Rustagi, *Optimizing methods in statistics*, New York: Academic Press, 1971, 89–113. https://doi.org/10.1016/B978-0-12-604550-5.50011-0

22. H. A. David, *Order statistics*, Wiley Online Library, 1981.

23. L. Bain, *Statistical analysis of reliability and life-testing models: theory and methods*, 2 Eds., New York: Routledge, 1991. https://doi.org/10.1201/9780203738733

24. J. A. Weymark, Generalized gini inequality indices, *Math. Soc. Sci.*, **1** (1981), 409–430. https://doi.org/10.1016/0165-4896(81)90018-4

25. E. M. Parrado-Gallardo, E. Bárcena-Martín, L. J. Imedio-Olmedo, Inequality, welfare and order statistics, In: J. A. Bishop, J. G. Rodríguez, *Economic well-being and inequality: papers from the fifth ECINEQ meeting*, Emerald Group Publishing Limited, **22** (2014), 383–399.

26. W. Barreto-Souza, A. L. de Morais, G. M. Cordeiro, The weibull-geometric distribution, *J. Stat. Comput. Simul.*, **81** (2011), 645–657. https://doi.org/10.1080/00949650903436554

27. A. P. Basu, J. P. Klein, *Some recent results in competing risks theory*, In: J. Crowley, R. A. Johnson, *Survival analysis*, IMS Lecture Notes–Monograph Series 2, Hayward, CA: IMS, 1982, 216–229.

28. L. Xie, Z. Wang, Reliability degradation of mechanical components and systems. In: K. B. Misra, *Handbook of performability engineering*, Springer, 2008, 413–429. https://doi.org/10.1007/978-1-84800-131-2_27

29. L. Y. Xie, J. Zhou, Y. Wang, X. Wang, Load-strength order statistics interference models for system reliability evaluation, *Int. J. Performabil. Eng.*, **1** (2005), 23–36.

30. F. Proschan, J. Sethuraman, Stochastic comparisons of order statistics from heterogeneous populations, with applications in reliability, *J. Multivariate Anal.*, **6** (1976), 608–616. https://doi.org/10.1016/0047-259X(76)90008-7

31. J. S. Kim, F. Proschan, J. Sethuraman, Stochastic comparisons of order statistics, with applications in relibility, *Commun. Stat.-Theory Methods*, **17** (1988), 2151–2172. https://doi.org/10.1080/03610928808829739

32. J. E. Ramirez-Marquez, D. W. Coit, Composite importance measures for multi-state systems with multi-state components, *IEEE Trans. Reliabil.*, **54** (2005), 517–529. https://doi.org/10.1109/TR.2005.853444

33. S. Eryilmaz, Lifetime of multistate k-out-of-n systems, *Qual. Reliab. Eng. Int.*, **30** (2014), 1015–1022. https://doi.org/10.1002/qre.1529

34. M. J. Anzanello, A simplified approach for reliability evaluation and component allocation in three-state series and parallel systems composed of non-identical components, *Gest. Prod.*, **16** (2009), 54–62. https://doi.org/10.1590/S0104-530X2009000100006

35. G. Yingkui, L. Jing, Multi-state system reliability: a new and systematic review, *Procedia Eng.*, **29** (2012), 531–536. https://doi.org/10.1016/j.proeng.2011.12.756

36. C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, **27** (1948), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

37. R. C. Gupta, M. Waleed, Analysis of survival data by a Weibull-Bessel distribution, *Commun. Stat.-Theory Methods*, **47** (2018), 980–995. https://doi.org/10.1080/03610926.2017.1316402

38. F. Proschan, Theoretical explanation of observed decreasing failurerate, *Technometrics*, **5** (1963), 375–383. https://doi.org/10.1080/00401706.1963.10490105

39. 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

40. M. Chahkandi, M. Ganjali, On some lifetime distributions with decreasing failure rate, *Comput. Stat. Data Anal.*, **53** (2009), 4433–4440. https://doi.org/10.1016/j.csda.2009.06.016

41. R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Comput.*, **16** (1995), 1190–1208. https://doi.org/10.1137/0916069

42. M. V. Aarset, How to identify a bathtub hazard rate, *IEEE Trans. Reliabil.*, **R-36** (1987), 106–108. https://doi.org/10.1109/tr.1987.5222310

AIMS Press