AIMS Mathematics, 10(12): 29927–29954.
DOI: 10.3934/math.20251315
Received: 07 September 2025
Revised: 22 October 2025
Accepted: 11 November 2025
Published: 19 December 2025

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

Research article

From exponential-geometric to Lomax: A unified survival model via gamma frailty

Mohieddine Rahmouni*

Applied College, King Faisal University, Al-Ahsa, Saudi Arabia

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

Abstract: This paper introduces a new flexible lifetime distribution that unifies the exponential-geometric and Lomax models through a random system size and a shared gamma frailty. It models the first failure time in a system with a geometrically distributed number of conditionally exponential components with gamma frailty, producing a geometric mixture of Lomax distributions. The model generalizes classical distributions, reducing to exponential-geometric, Lomax, or exponential under appropriate limits. We derived analytical expressions for the cumulative distribution, survival, and hazard functions, discussed maximum likelihood estimation, and illustrated its competitive and versatile fits on real datasets.

Keywords: survival analysis; gamma frailty; geometric distribution; Lomax distribution; mixture model; reliability; first failure time
Mathematics Subject Classification: 62E15, 60E05, 62F10

1. Introduction

Modeling lifetime (survival) data is central in reliability engineering [1–3], biomedical research [4, 5], economics [6, 7], social sciences [8], and actuarial science [9–11], informing analyses of system failures, disease progression, employment duration, and financial risk. Such data often exhibit decreasing failure rates (DFR), heavy tails, and unobserved heterogeneity [12]. Recent advances in reliability include fatigue-failure models incorporating stress spectra for mechanical parts [13] and Weibull-based reliability assessments enhanced by the Bayesian bootstrap [14], underscoring the demand for flexible, data-driven lifetime distributions. These developments motivate unified models that accommodate both structural redundancy and frailty-driven heterogeneity, such as the framework proposed here [15].

A seminal contribution in this area is the exponential-geometric (EG) distribution introduced by [16]. This model represents the time to the first failure in a system consisting of a random number

of components, where the system size follows a geometric distribution. The EG distribution captures decreasing hazard rates naturally and has been widely applied in reliability studies and survival analysis. Subsequent extensions, such as the exponential-generalized truncated geometric (EGTG) distribution [17], generalized the EG framework to account for higher-order failures, further enhancing its applicability in modeling complex system behaviors.

Another well-established approach to modeling heterogeneity is the use of frailty models [18–20]. Frailty introduces unobserved random effects into the hazard function, effectively capturing latent variability among units. Among the various frailty distributions, the gamma distribution has been particularly influential due to its mathematical tractability and ability to induce heavy-tailed behavior [21]. For instance, applying a gamma frailty to exponential lifetimes leads to the Lomax (Pareto Type II) distribution, which accommodates long-tailed failure times frequently observed in reliability data [22].

Positioning and novelty. Prior research falls into three strands. *(i) EG-family redundancy models* assume a random number of parallel components with exponential lifetimes; a geometric count injects extra variability in first-failure times and yields simple likelihoods, but tail/hazard shaping is limited without further structure [16, 23]. *(ii) Shared-frailty models* multiply hazards by a latent random effect (often gamma), capturing heavy tails and over-dispersion with clear inferential tools, yet they do not encode randomness in system size; with exponential baselines, gamma frailty yields Lomax marginals [19, 21]. *(iii) Compound/mixture approaches* (for example, flexible baselines and finite mixtures) fit diverse shapes but can sacrifice mechanistic first-failure interpretation and parameter transparency [24].

Our contribution integrates (i) and (ii) in the first-failure setting. We model $Y = \min\{X_1, \ldots, X_N\}$ with $N \sim \text{Geom}(1-\eta)$ on $\{1, 2, \ldots\}$ and a shared gamma frailty $Z \sim \Gamma(\lambda, \lambda)$ (mean 1), where, given Z, $X_1, \ldots, X_N \overset{\text{iid}}{\sim} \text{Exp}(\theta Z)$. Marginalizing N and Z yields a tractable geometric–Lomax mixture for Y with three interpretable parameters $\boldsymbol{\vartheta} = (\theta, \lambda, \eta)$: θ sets the time scale, λ controls the frailty concentration (tail index), and η tunes the redundancy/mixing intensity. The construction bridges EG ($\lambda \to \infty$), Lomax ($\eta \to 0$), and exponential ($\eta \to 0,\ \lambda \to \infty$) limits, preserves a decreasing-failure-rate hazard suited to infant-mortality/first-failure regimes, and offers greater flexibility than EG while retaining mechanistic meaning and estimation tractability (via maximum likelihood estimation, MLE; the expectation–maximization algorithm, EM; and its Monte Carlo variant, MCEM).

To assess practical relevance, we analyze reliability datasets and find that the unified model competes with—and often improves on—classical alternatives by capturing both early failures and long tails.

The remainder of the paper is organized as follows. Section 2 presents the hierarchical construction and the marginal law of Y. Section 3 develops distributional properties (cdf, survival, hazard, limits). Section 4 details maximum likelihood estimation and EM/MCEM procedures. Section 5 presents a Monte Carlo simulation study conducted to assess the finite-sample performance of the MLE for the proposed unified distribution. Section 6 reports real-data applications, and Section 7 outlines implications and future directions.

2. Distribution derivation

Consider a system with a random number of components $N \sim \text{Geometric}(p = 1-\eta)$, where $\Pr(N = n) = (1-\eta)\eta^{n-1}$ for $n \geq 1$. The component lifetimes are subject to a shared frailty $Z \sim \Gamma(\lambda, \lambda)$ with

density $f_Z(z) = \frac{\lambda^\lambda}{\Gamma(\lambda)} z^{\lambda-1} e^{-\lambda z}, z > 0$. Conditional on $Z = z$, let $X_i \mid Z = z \sim \text{Exp}(\theta z)$ independently for $i = 1, \ldots, N$, and define the system failure time as $Y = \min\{X_1, \ldots, X_N\}$. This hierarchical construction induces a marginal distribution for Y that unifies the EG and Lomax distributions. The following theorem establishes the closed-form expression for the PDF of Y.

Theorem 1 (Marginal PDF of the first failure time)**.** *Under the hierarchical model above, the marginal PDF of $Y = X_{(1)}$ is*

$$f_Y(y) = \theta\lambda^{\lambda+1}(1-\eta)\sum_{n=1}^{\infty} \frac{n\eta^{n-1}}{(n\theta y + \lambda)^{\lambda+1}}, \quad y > 0. \tag{2.1}$$

Proof. We derive the marginal PDF in two steps: first by integrating over the frailty Z and then by summing over the random system size N.

Conditional PDF given $N = n$ (step 1): Given $N = n$ and $Z = z$, the component lifetimes are independent and identically distributed (i.i.d.) $\text{Exp}(\theta z)$, so the minimum $Y = \min\{X_1, \ldots, X_n\}$ has the PDF:

$$f_{Y|N=n,Z=z}(y) = n\theta z e^{-n\theta z y}, \quad y > 0.$$

The marginalization over $Z \sim \Gamma(\lambda, \lambda)$ gives:

$$f_{Y|N=n}(y) = \int_0^\infty n\theta z e^{-n\theta z y} \cdot \frac{\lambda^\lambda}{\Gamma(\lambda)} z^{\lambda-1} e^{-\lambda z} dz = \frac{n\theta\lambda^\lambda}{\Gamma(\lambda)} \int_0^\infty z^\lambda e^{-(\lambda+n\theta y)z} dz.$$

Using the standard gamma integral

$$\int_0^\infty z^a e^{-bz} dz = \frac{\Gamma(a+1)}{b^{a+1}}, \quad a > -1,\ b > 0,$$

with $a = \lambda$, $b = \lambda + n\theta y$, we obtain:

$$f_{Y|N=n}(y) = \frac{n\theta\lambda^\lambda}{\Gamma(\lambda)} \cdot \frac{\Gamma(\lambda+1)}{(\lambda+n\theta y)^{\lambda+1}}.$$

Since $\Gamma(\lambda+1) = \lambda\Gamma(\lambda)$, this simplifies to:

$$f_{Y|N=n}(y) = \frac{n\theta\lambda^{\lambda+1}}{(\lambda+n\theta y)^{\lambda+1}}.$$

This is the PDF of a Lomax (Pareto Type II) distribution with shape parameter λ and scale parameter $\lambda/(n\theta)$.

Marginalization over N (step 2): Summing over $N \sim \text{Geometric}(1-\eta)$ gives:

$$f_Y(y) = \sum_{n=1}^\infty \Pr(N = n) f_{Y|N=n}(y) = \sum_{n=1}^\infty (1-\eta)\eta^{n-1} \cdot \frac{n\theta\lambda^{\lambda+1}}{(\lambda+n\theta y)^{\lambda+1}}.$$

Factoring out constants yields:

$$f_Y(y) = \theta\lambda^{\lambda+1}(1-\eta)\sum_{n=1}^{\infty} \frac{n\eta^{n-1}}{(\lambda+n\theta y)^{\lambda+1}}, \quad y > 0.$$

□

Corollary 2 (Geometric mixture representation)**.** *The first failure time Y follows a geometric mixture of Lomax distributions:*

$$Y \sim \sum_{n=1}^{\infty} w_n \cdot \text{Lomax}\left(\lambda, \frac{\lambda}{n\theta}\right),$$

where the mixing weights are $w_n = (1-\eta)\eta^{n-1}$*,* $n = 1, 2, \ldots$*, and each component has the PDF*

$$f_n(y) = \frac{n\theta\lambda^{\lambda+1}}{(\lambda + n\theta y)^{\lambda+1}}, \quad y > 0.$$

Proof. This follows directly from Theorem 1 and the fact that $f_{Y|N=n}(y)$ is the PDF of a Lomax$(\lambda, \lambda/(n\theta))$ distribution. The marginal density is a weighted sum of these component densities with weights $\Pr(N = n) = (1-\eta)\eta^{n-1}$, which confirms the mixture representation. □

3. Distributional properties

3.1. Shape of the PDF

Let $Y = X_{(1)}$ denote the first-failure time under the unified frailty model with parameters $\theta > 0$, $\lambda > 0$, and $\eta \in [0, 1)$, whose PDF is given in Equation (2.1).

Theorem 3 (Mode of the unified frailty distribution)**.** *The PDF* $f_Y(y)$ *is strictly decreasing on* $(0, \infty)$*, so its global maximum occurs at the boundary*

$$y_{\text{mode}} = 0, \qquad f_Y(0) = \frac{\theta}{1-\eta}.$$

Hence, the unified frailty distribution is unimodal with a peak at the origin.

Proof. Each summand of the series representation of $f_Y(y)$ is strictly decreasing. Since $f_Y(y)$ is a positive linear combination of these decreasing terms, it is strictly decreasing as well. Evaluating the series at $y = 0$ gives the stated maximum.

$$f_Y(0) \;=\; \theta(1-\eta)\sum_{n=1}^{\infty} n\eta^{n-1} = \theta(1-\eta)\cdot\frac{1}{(1-\eta)^2} = \frac{\theta}{1-\eta}.$$

When $\eta = 0$, this reduces to the Lomax case: $f_Y(0) = \theta$. As $\eta \to 1^-$, $f_Y(0) \to \infty$, meaning the density at 0 explodes. □

Corollary 4 (Behavior at the origin)**.** *At* $y = 0$*, the PDF has negative slope and positive curvature:*

$$f_Y'(0) = -\frac{\lambda+1}{\lambda}\,\theta^2\,\frac{1+\eta}{(1-\eta)^2} < 0, \qquad f_Y''(0) = \frac{(\lambda+1)(\lambda+2)}{\lambda^2}\,\theta^3\,\frac{1+4\eta+\eta^2}{(1-\eta)^3} > 0.$$

Consequently, the mode at the origin is sharp and convex.

Proof. This directly follows by differentiating the series termwise and evaluating at $y = 0$ (see the previous derivation for details). □

3.2. CDF and survival function

From the mixture representation in Corollary 2, the CDF of Y is obtained by averaging the Lomax CDF over the geometric mixing distribution:

$$F_Y(y) = (1-\eta)\sum_{n=1}^{\infty}\eta^{n-1}\left[1-\left(\frac{\lambda}{n\theta y+\lambda}\right)^{\lambda}\right], \qquad y>0.$$

Equivalently, the survival function ($S_Y(y)$) can be expressed as

$$S_Y(y) = 1-F_Y(y) = (1-\eta)\sum_{n=1}^{\infty}\eta^{n-1}\left(\frac{\lambda}{n\theta y+\lambda}\right)^{\lambda}, \qquad y>0.$$

These series expressions highlight the geometric-Lomax mixture structure: for fixed n, the survival probability is inherited from a Lomax$(\lambda, \lambda/(n\theta))$ component, and the weights $(1-\eta)\eta^{n-1}$ describe the probability of the system having n components.

3.3. Hazard function

The hazard rate of Y is

$$h_Y(y) = \frac{f_Y(y)}{S_Y(y)}, \qquad y>0,$$

where, from Theorem 1 and Section 3,

$$S_Y(y) = (1-\eta)\,\lambda^{\lambda}\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{(n\theta y+\lambda)^{\lambda}}, \qquad f_Y(y) = \theta\,\lambda^{\lambda+1}(1-\eta)\sum_{n=1}^{\infty}\frac{n\,\eta^{n-1}}{(n\theta y+\lambda)^{\lambda+1}}.$$

Hence the factors $(1-\eta)$ and λ^{λ} cancel, yielding the explicit ratio

$$h_Y(y) = \theta\,\lambda\,\frac{\displaystyle\sum_{n=1}^{\infty}\frac{n\,\eta^{n-1}}{(n\theta y+\lambda)^{\lambda+1}}}{\displaystyle\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{(n\theta y+\lambda)^{\lambda}}}, \qquad y>0.$$

Boundary behavior. Using the above series,

$$\lim_{y\downarrow 0} S_Y(y) = 1, \qquad \lim_{y\downarrow 0} f_Y(y) = \theta(1-\eta)\sum_{n=1}^{\infty} n\,\eta^{n-1} = \frac{\theta}{1-\eta},$$

so

$$\lim_{y\downarrow 0} h_Y(y) = \frac{\theta}{1-\eta}.$$

For large y, $(n\theta y+\lambda)\sim n\theta y$, and

$$S_Y(y)\sim(1-\eta)\frac{\lambda^{\lambda}}{(\theta y)^{\lambda}}\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^{\lambda}}, \qquad f_Y(y)\sim(1-\eta)\frac{\theta\,\lambda^{\lambda+1}}{(\theta y)^{\lambda+1}}\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^{\lambda}},$$

so the common series cancels in the ratio, giving the universal tail law

$$\lim_{y\to\infty} y\,h_Y(y) = \lambda \quad\Longrightarrow\quad h_Y(y)\sim\frac{\lambda}{y} \text{ as } y\to\infty.$$

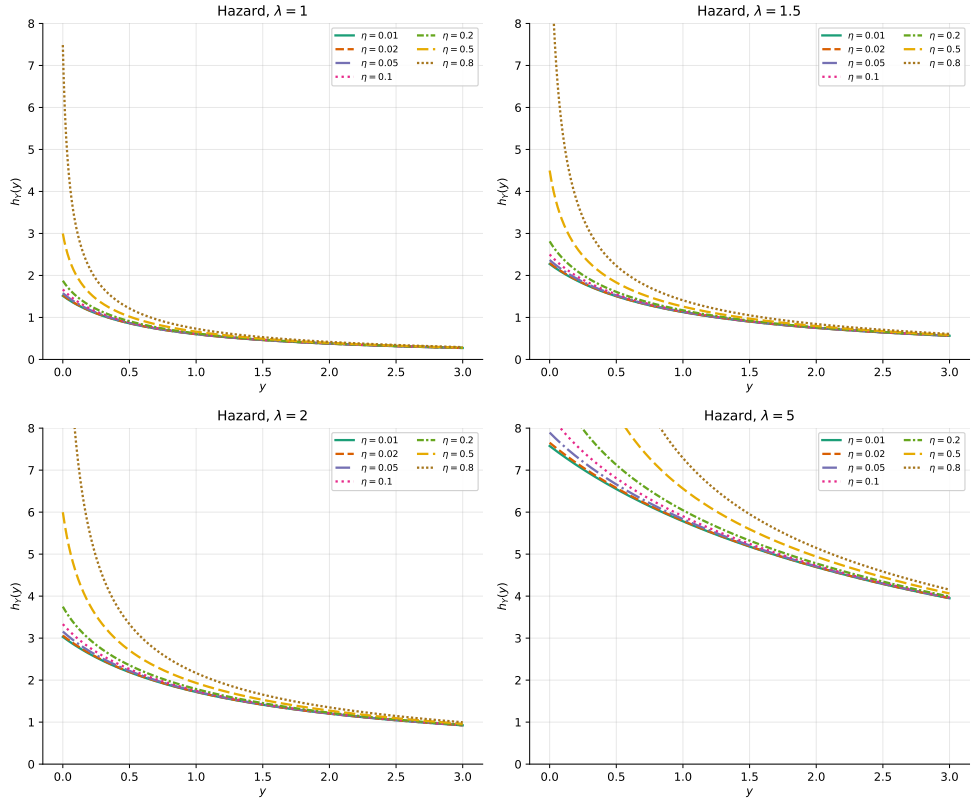

Figure 1. Hazard functions $h_Y(y)$ for the unified model with $\theta = 1.5$. Panels show $\lambda \in \{1, 1.5, 2, 5\}$; within each panel, curves vary over $\eta \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 0.8\}$.

Shape. Using the series forms for f_Y and S_Y above,

$$h_Y(0^+) = \frac{\theta}{1-\eta}, \qquad h_Y(y) \sim \frac{\lambda}{y} \text{ as } y \to \infty,$$

so the hazard starts finite and eventually decays to 0. Conditional on $N = n$, the component is Lomax$(\lambda, \lambda/(n\theta))$ with hazard $h_n(y) = \frac{n\theta\lambda}{\lambda+n\theta y}$, which is decreasing in y. For this model, the hazard can be shown to have a DFR for all $\theta > 0, \lambda > 0, \eta \in [0, 1)$; in particular, $h_Y(0^+) = \theta/(1-\eta)$ and $h_Y(y) \approx \lambda/y$ as $y \to \infty$.

3.4. Limiting and special cases

The proposed hierarchical frailty model for the time to first failure Y provides a unifying framework that encompasses several classical lifetime distributions as special or limiting cases, highlighting its flexibility in reliability and survival analysis. Specifically, as the frailty parameter $\lambda \to \infty$, the gamma-distributed frailty $Z \sim \Gamma(\lambda, \lambda)$ converges in probability to 1 (since $\mathbb{E}[Z] = 1$ and $\text{Var}(Z) = 1/\lambda \to 0$), effectively removing the shared heterogeneity; conditionally, $Y \mid N = n$ becomes Exp$(n\theta)$, and marginalizing over $N \sim$ Geometric$(1-\eta)$ yields the EG distribution of Adamidis & Loukas [16], with PDF $f_Y(y) = \theta(1-\eta)e^{-\theta y}[1-\eta e^{-\theta y}]^{-2}$ for $y > 0$, CDF $F_Y(y) = 1-(1-\eta)[1-\eta e^{-\theta y}]^{-1}e^{-\theta y}$, and a monotonically decreasing hazard rate $h_Y(y) = \theta[1-\eta e^{-\theta y}]^{-1}$. Conversely, as $\eta \to 0$, $\Pr(N = 1) = 1-\eta \to 1$, reducing the model to a single component with lifetime modulated by Z, so Y follows the Lomax (or

Pareto Type II) distribution with shape parameter λ and scale λ/θ, having PDF $f_Y(y) = \theta\lambda^\lambda(\lambda+\theta y)^{-\lambda-1}$ for $y > 0$, survival function $S_Y(y) = [\lambda/(\lambda+\theta y)]^\lambda$, and hazard rate $h_Y(y) = \theta\lambda/(\lambda+\theta y)$ that decreases to 0 as $y \to \infty$. Finally, taking both limits simultaneously ($\eta \to 0$ and $\lambda \to \infty$), the Lomax degenerates to the exponential distribution Exp(θ) with constant hazard θ, PDF $f_Y(y) = \theta e^{-\theta y}$ for $y > 0$, and memoryless property $\Pr(Y > y + t \mid Y > t) = \Pr(Y > y)$. These reductions are confirmed by direct computation of the limiting forms of the general PDF from Theorem 1, underscoring the model's role as a generalization bridging light- and heavy-tailed behaviors. These limiting cases are summarized in Figure 2.

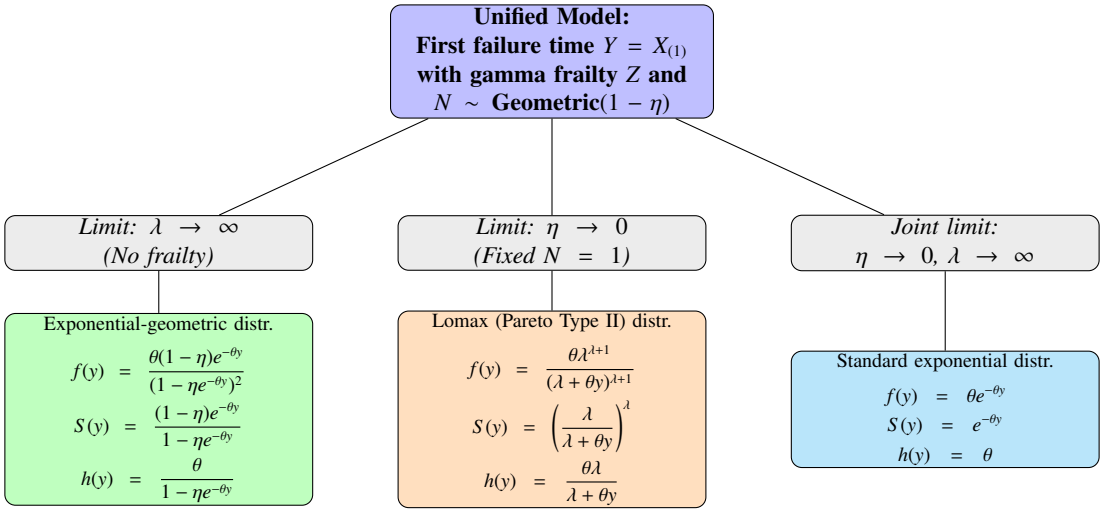

Figure 2. Special and limiting cases of the unified model for the first failure time.

Figures 3 and 4 show how η and λ shape the unified density. Larger η shifts mass toward the origin (more weight on larger n in the geometric mixture), while larger λ concentrates frailty and moves the model toward the EG limit. In the joint limit $\eta \to 0,\ \lambda \to \infty$, the model becomes exponential.

The Lomax case corresponds to $\eta \to 0$ with fixed λ. For any fixed λ,

$$f_Y(0^+) = \theta(1-\eta)\sum_{n\geq 1} n\eta^{n-1} = \frac{\theta}{1-\eta}, \qquad S_Y(y) \sim C(\eta,\lambda)\, y^{-\lambda} \quad (y \to \infty),$$

so, relative to the Lomax ($f_L(0^+) = \theta$), the unified density has a higher initial peak by the factor $1/(1-\eta)$ while preserving the same tail index λ. In the figure this appears as unified curves ($\eta > 0$) lying above the Lomax near $y = 0$, but running parallel in the tail (both hazards satisfy $h_Y(y) \sim \lambda/y$). Smaller λ yields heavier tails in both models; larger λ lightens the tail and moves the unified model toward the EG limit near the origin. Overall, (θ, λ, η) provide an interpretable bridge between the classical cases: θ sets the time scale, λ controls tail heaviness via the frailty concentration, and η tunes the early-failure intensity via geometric mixing.

3.5. Moment generating function (MGF)

We now derive the moment generating function (MGF) of the unified frailty distribution. Recall that if $Y = X_{(1)}$ denotes the first-failure time with the PDF given in Eq (2.1), then the MGF is defined

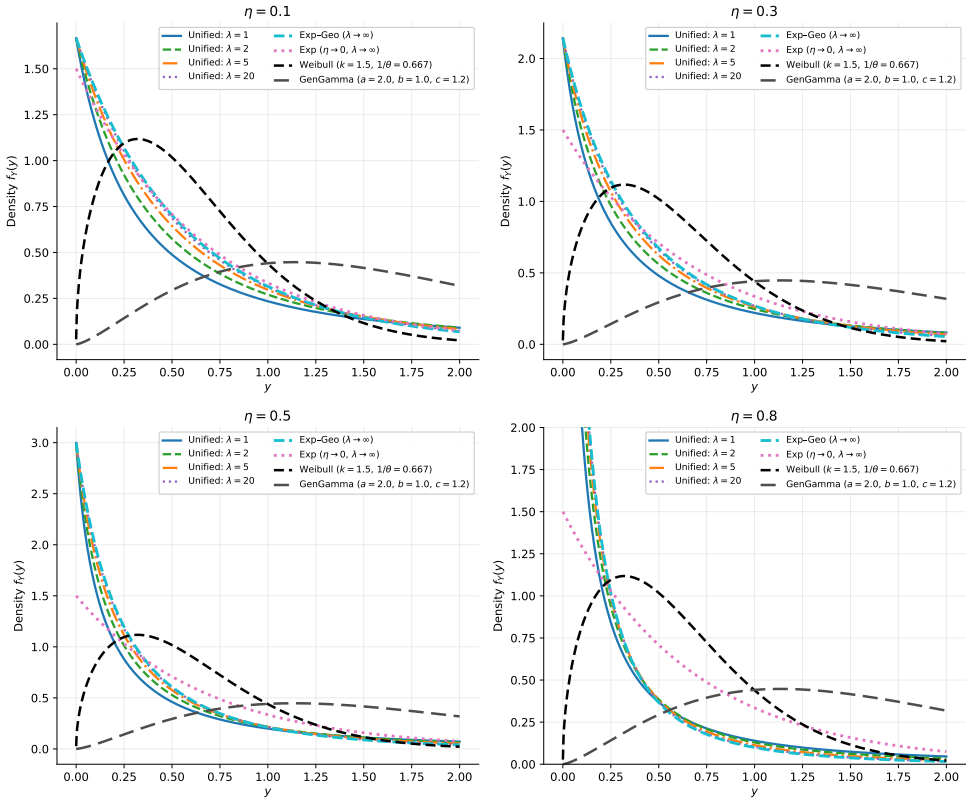

Figure 3. Unified model densities for $\theta = 1.5$ across $\eta \in \{0.1, 0.3, 0.5, 0.8\}$. Each panel overlays $\lambda \in \{1, 2, 5, 20\}$ and comparator curves: EG ($\lambda \to \infty$), Exponential ($\eta \to 0,\ \lambda \to \infty$), Weibull ($k = 1.5$, scale $= 1/\theta$), and Generalized Gamma ($a = 2.0, b = 1.0, c = 1.2$). Increasing η shifts mass toward the origin; larger λ approaches the EG limit.

as

$$M_Y(t) = \mathbb{E}[e^{tY}] = \int_0^\infty e^{ty} f_Y(y)\, dy, \qquad t < 0.$$

Theorem 5 (MGF of the unified frailty distribution)**.** *For $t < 0$, the MGF of Y is*

$$M_Y(t) = \theta\lambda^{\lambda+1}(1-\eta) \sum_{n=1}^{\infty} n\eta^{n-1} \int_0^\infty \frac{e^{ty}}{(n\theta y + \lambda)^{\lambda+1}}\, dy.$$

After the change of variables $u = n\theta y + \lambda$, this reduces to

$$M_Y(t) = (1-\eta) \sum_{n=1}^{\infty} \eta^{n-1} \left(\frac{\lambda}{n\theta}\right)^{\lambda} e^{-\frac{t\lambda}{n\theta}} \int_\lambda^\infty u^{-(\lambda+1)} e^{\frac{t}{n\theta}u}\, du.$$

Proof. Starting from the definition of the MGF and substituting the series representation of the PDF yields the first expression. The substitution $u = n\theta y + \lambda$ (with $dy = du/(n\theta)$) transforms the denominator into a power of u, and the exponential term into $e^{\frac{t}{n\theta}(u-\lambda)}$, giving the second expression. □

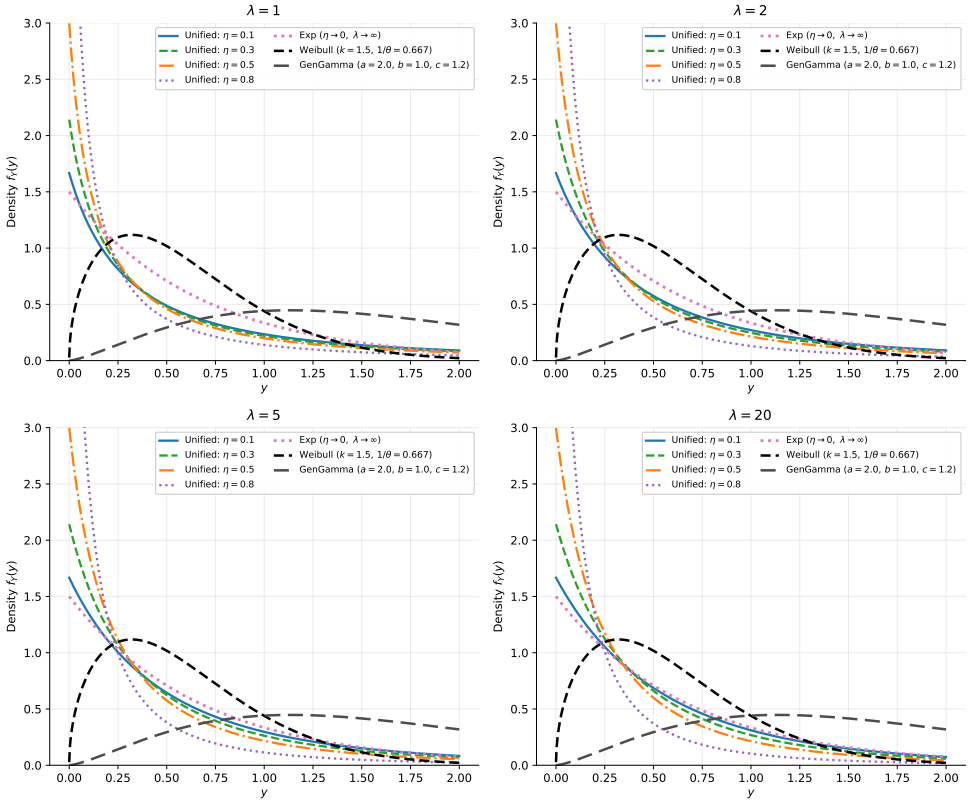

Figure 4. Unified model densities for $\theta = 1.5$ across $\lambda \in \{1, 2, 5, 20\}$. Each panel overlays $\eta \in \{0.1, 0.3, 0.5, 0.8\}$ with the same comparators (EG, Exp., Weibull, GenGamma). Larger λ concentrates frailty and lightens the tail; smaller λ yields heavier tails.

Corollary 6 (Expression via the incomplete gamma function)**.** *The MGF admits the representation*

$$M_Y(t) = (1-\eta)\sum_{n=1}^{\infty} \eta^{n-1} \left(\frac{\lambda}{n\theta}\right)^{\lambda} e^{-\frac{t\lambda}{n\theta}} \cdot \frac{1}{n\theta}\,\Gamma\Big(-\lambda, -\tfrac{t}{n\theta}\lambda\Big), \qquad t < 0,$$

where $\Gamma(s, x)$ *denotes the upper incomplete gamma function.*

Proof. The integral in Theorem 5 has the form of an incomplete gamma function once the substitution $z = -\frac{t}{n\theta}u$ is applied. Simplification then yields the stated result. □

3.6. Moments

The moments of Y are obtained from the mixture representation

$$Y \mid (N = n) \sim \text{Lomax}(\alpha = \lambda,\ b_n = \tfrac{\lambda}{n\theta}), \qquad \Pr(N = n) = (1-\eta)\eta^{n-1},\ n \geq 1.$$

Recall that for a Lomax(α, b) distribution the mean exists for $\alpha > 1$ and the variance exists for $\alpha > 2$, with

$$\mathbb{E}[Y] = \frac{b}{\alpha - 1}, \qquad \text{Var}(Y) = \frac{b^2\alpha}{(\alpha-1)^2(\alpha-2)}.$$

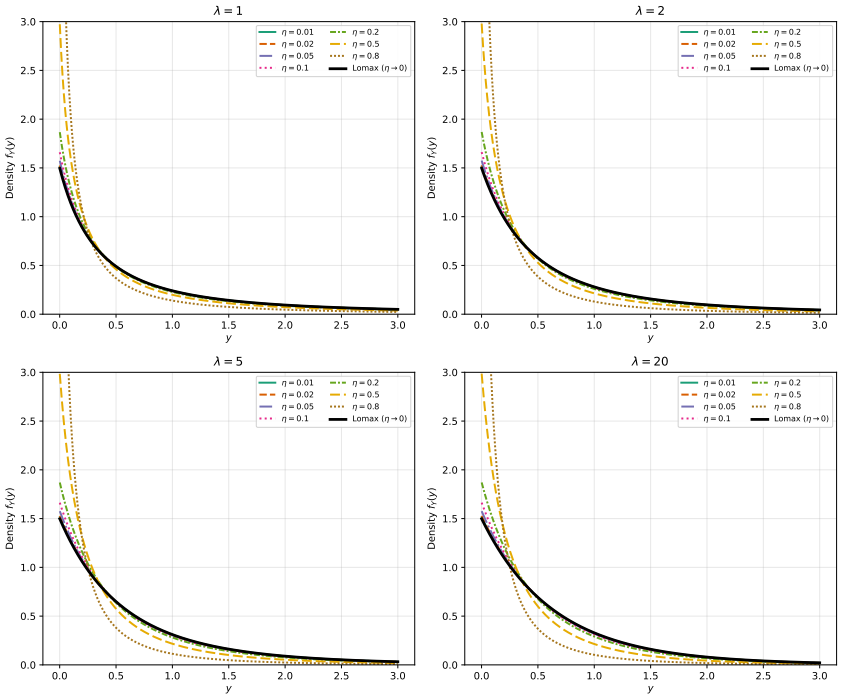

Figure 5. Unified vs. Lomax ($\eta \to 0$) densities with $\theta = 1.5$. Panels show $\lambda \in \{1, 2, 5, 20\}$; within each panel, unified curves vary over $\eta \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 0.8\}$ and the Lomax limit is overlaid.

Theorem 7 (Mean and variance)**.** *Assume $\theta > 0$, $\lambda > 0$, and $\eta \in [0, 1)$. Define the sums*

$$S_1(\eta) := (1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n} = \frac{1-\eta}{\eta}[-\ln(1-\eta)], \qquad S_2(\eta) := (1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^2} = \frac{1-\eta}{\eta}\,\mathrm{Li}_2(\eta),$$

where Li_2 is the dilogarithm (polylogarithm of order 2).

(i) If $\lambda > 1$ then the mean of Y exists and is

$$\mathbb{E}[Y] = \frac{\lambda}{(\lambda-1)\,\theta}\,S_1(\eta) = \frac{\lambda}{(\lambda-1)\,\theta}\,\frac{1-\eta}{\eta}[-\ln(1-\eta)].$$

(ii) If $\lambda > 2$ then the variance of Y exists and is

$$\mathrm{Var}(Y) = \underbrace{\frac{(1-\eta)\,\lambda^3}{\theta^2(\lambda-1)^2(\lambda-2)}\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^2}}_{\mathbb{E}[\mathrm{Var}(Y|N)]} + \underbrace{\left(\frac{\lambda}{(\lambda-1)\theta}\right)^2\left\{(1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^2} - \left[(1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n}\right]^2\right\}}_{\mathrm{Var}(\mathbb{E}[Y|N])}$$

$$= \frac{\lambda^3}{\theta^2(\lambda-1)^2(\lambda-2)}\,S_2(\eta) + \left(\frac{\lambda}{(\lambda-1)\theta}\right)^2\{S_2(\eta) - [S_1(\eta)]^2\}.$$

Proof. Conditional moments follow from the Lomax form. For $N = n$,

$$\mathbb{E}[Y \mid N = n] = \frac{b_n}{\lambda - 1} = \frac{\lambda}{(\lambda-1)n\theta}, \qquad \mathrm{Var}(Y \mid N = n) = \frac{b_n^2\lambda}{(\lambda-1)^2(\lambda-2)} = \frac{\lambda^3}{(\lambda-1)^2(\lambda-2)}\frac{1}{n^2\theta^2},$$

provided $\lambda > 2$.

Use the law of total expectation and total variance. The mean is

$$\mathbb{E}[Y] = (1-\eta)\sum_{n=1}^{\infty} \eta^{n-1}\,\frac{\lambda}{(\lambda-1)n\theta} = \frac{\lambda}{(\lambda-1)\theta}\,(1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n},$$

which gives the formula via the identity $\sum_{n=1}^{\infty} \eta^n/n = -\ln(1-\eta)$.

Proposition 8 (Second moment)**.** *If $\lambda > 2$, then*

$$\mathbb{E}[Y^2] \;=\; \frac{2\,\lambda^2}{(\lambda-1)(\lambda-2)\,\theta^2}\,S_2(\eta), \qquad S_2(\eta) := (1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^2} = \frac{1-\eta}{\eta}\,\mathrm{Li}_2(\eta).$$

In the Lomax limit $\eta \to 0$, $S_2(\eta) \to 1$ and thus $\mathbb{E}[Y^2] \to \dfrac{2\,\lambda^2}{(\lambda-1)(\lambda-2)\,\theta^2}$.

For the variance,

$$\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y \mid N)] + \mathrm{Var}(\mathbb{E}[Y \mid N]),$$

and substituting the conditional expressions yields the first boxed representation. Recognizing the sums $(1-\eta)\sum_{n\geq 1}\eta^{n-1}n^{-2} = \frac{1-\eta}{\eta}\,\mathrm{Li}_2(\eta)$ and $(1-\eta)\sum_{n\geq 1}\eta^{n-1}n^{-1} = S_1(\eta)$ gives the compact second form. □

Remarks. The mean exists *if and only if* $\lambda > 1$ and the variance exists *if and only if* $\lambda > 2$, mirroring the classical Lomax moment conditions inherited by the mixture components. In the Lomax limit $\eta \to 0$, the unified model reduces to $\mathrm{Lomax}(\lambda, \lambda/\theta)$, yielding $\mathbb{E}[Y] \to \lambda/((\lambda-1)\theta)$ and $\mathrm{Var}(Y) \to \lambda^3/\{\theta^2(\lambda-1)^2(\lambda-2)\}$, as expected. For weak mixing (η small), corrections to these moments admit compact series in η via polylogarithms $\mathrm{Li}_s(\eta)$, providing accurate approximations to $\mathbb{E}[Y]$ and $\mathrm{Var}(Y)$ near the Lomax regime.

Proposition 9 (General moments for the unified model)**.** *Let $Y = X_{(1)}$ follow the unified frailty model with parameters $\theta > 0$, $\lambda > 0$, and $\eta \in [0, 1)$. For $r \geq 0$ define*

$$S_r(\eta) := (1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^r} \;=\; \frac{1-\eta}{\eta}\,\mathrm{Li}_r(\eta), \qquad S_0(\eta) = 1,$$

where Li_r is the polylogarithm of order r. If $\lambda > r$, the r-th raw moment exists and

$$\mathbb{E}[Y^r] \;=\; \Gamma(r+1)\,\frac{\Gamma(\lambda-r)}{\Gamma(\lambda)}\left(\frac{\lambda}{\theta}\right)^r S_r(\eta).$$

In particular, for $\lambda > 1$ and $\lambda > 2$,

$$\mathbb{E}[Y] \;=\; \frac{\lambda}{(\lambda-1)\,\theta}\,S_1(\eta), \qquad \mathbb{E}[Y^2] \;=\; \frac{2\,\lambda^2}{(\lambda-1)(\lambda-2)\,\theta^2}\,S_2(\eta).$$

Proof. Condition on the random system size $N = n$. Given $N = n$, Y is Lomax with shape λ and scale $b_n = \lambda/(n\theta)$, so for $\lambda > r$,

$$\mathbb{E}[Y^r \mid N = n] = b_n^r\,\frac{\Gamma(r+1)\Gamma(\lambda-r)}{\Gamma(\lambda)} = \Gamma(r+1)\,\frac{\Gamma(\lambda-r)}{\Gamma(\lambda)}\left(\frac{\lambda}{n\theta}\right)^r.$$

Averaging over $N \sim \text{Geom}(1-\eta)$ on $\{1, 2, \ldots\}$ yields

$$\mathbb{E}[Y^r] = \Gamma(r+1)\,\frac{\Gamma(\lambda-r)}{\Gamma(\lambda)}\left(\frac{\lambda}{\theta}\right)^r (1-\eta)\sum_{n=1}^{\infty}\frac{\eta^{n-1}}{n^r} = \Gamma(r+1)\,\frac{\Gamma(\lambda-r)}{\Gamma(\lambda)}\left(\frac{\lambda}{\theta}\right)^r S_r(\eta).$$

□

Remark (EG limit). As $\lambda \to \infty$, the Lomax component $f_{Y|N=n}(y) = n\theta\,\lambda^{\lambda+1}(\lambda + n\theta y)^{-(\lambda+1)}$ satisfies $(1 + n\theta y/\lambda)^{-(\lambda+1)} \to e^{-n\theta y}$, so $Y \mid N = n \Rightarrow \text{Exp}(n\theta)$ and the model reduces to

$$Y \mid N = n \sim \text{Exp}(n\theta), \qquad \Pr(N = n) = (1-\eta)\eta^{n-1},\ n \geq 1.$$

Consequently, the EG limit retains the same geometric weighting and the same $S_r(\eta) = (1-\eta)\sum_{n\geq 1}\eta^{n-1}/n^r$ in all moment expressions.

Corollary 10 (Moments in the EG limit)**.** *Let Y follow the EG distribution with parameters $\theta > 0$ and $\eta \in [0, 1)$, i.e., $Y \mid N = n \sim \text{Exp}(n\theta)$ and $\Pr(N = n) = (1-\eta)\eta^{n-1}$ for $n \geq 1$. Then, for every $r \geq 0$,*

$$\mathbb{E}[Y^r] \;=\; \Gamma(r+1)\,\theta^{-r}\,S_r(\eta), \qquad S_r(\eta) = \frac{1-\eta}{\eta}\,\text{Li}_r(\eta),\ \ S_0(\eta) = 1.$$

In particular,

$$\mathbb{E}[Y] \;=\; \frac{S_1(\eta)}{\theta}, \qquad \mathbb{E}[Y^2] \;=\; \frac{2\,S_2(\eta)}{\theta^2}, \qquad \text{Var}(Y) \;=\; \frac{1}{\theta^2}\Big(2S_2(\eta) - [S_1(\eta)]^2\Big).$$

Moreover, $S_r(\eta) < \infty$ for all $r > 0$ and $S_0(\eta) = 1$; as $\eta \to 0$, $S_r(\eta) \to 1$; and as $\eta \to 1^-$, $S_r(\eta) \to 0$ for any fixed $r > 0$.

3.7. Quantiles

Let $p \in (0, 1)$ and define the p-quantile $q_p = \inf\{y > 0 : F_Y(y) \geq p\}$. From the series forms,

$$F_Y(y) = (1-\eta)\sum_{n\geq 1}\eta^{n-1}\left[1 - \left(\frac{\lambda}{\lambda + n\theta y}\right)^{\lambda}\right], \quad S_Y(y) = (1-\eta)\sum_{n\geq 1}\eta^{n-1}\left(\frac{\lambda}{\lambda + n\theta y}\right)^{\lambda},$$

so q_p is the (unique) solution of $F_Y(q_p) = p$. A closed form is generally unavailable, but accurate limits and asymptotics are as follows.

Lower tail ($p \downarrow 0$). For $y \downarrow 0$,

$$\left(\tfrac{\lambda}{\lambda+n\theta y}\right)^{\lambda} = (1 + \tfrac{n\theta y}{\lambda})^{-\lambda} = 1 - n\theta y + o(y)$$

termwise, giving

$$F_Y(y) = \frac{\theta}{1-\eta}\,y + o(y), \qquad y \downarrow 0.$$

Hence the small-p quantile is

$$q_p \;=\; \frac{1-\eta}{\theta}\,p(1 + o(1)), \qquad p \downarrow 0,$$

i.e., $q_p \approx p/h_Y(0^+)$ since $h_Y(0^+) = \theta/(1-\eta)$.

Upper tail ($p \uparrow 1$). As $y \to \infty$, $n\theta y + \lambda \sim n\theta y$ so

$$S_Y(y) \sim (1-\eta)\,\frac{\lambda^\lambda}{(\theta y)^\lambda} \sum_{n\geq 1} \frac{\eta^{n-1}}{n^\lambda} = (1-\eta)\,\frac{\lambda^\lambda}{(\theta y)^\lambda}\, C(\eta,\lambda),$$

where $C(\eta,\lambda) := \sum_{n\geq 1} \eta^{n-1} n^{-\lambda} = \eta^{-1}\,\mathrm{Li}_\lambda(\eta)$. Solving $S_Y(q_p) \approx 1-p$ yields

$$q_p \;=\; \frac{\lambda}{\theta}\left(\frac{(1-\eta)\,C(\eta,\lambda)}{1-p}\right)^{1/\lambda}(1+o(1)), \qquad p \uparrow 1.$$

Limiting closed forms.

- *EG limit ($\lambda \to \infty$).* With $u = e^{-\theta y}$, $S_Y(y) = \frac{(1-\eta)u}{1-\eta u}$. Setting $S_Y(q_p) = 1-p$ gives

$$q_p \;=\; -\frac{1}{\theta}\ln\!\left(\frac{1-p}{1-\eta p}\right) \qquad (\lambda \to \infty),$$

  which reduces to $q_p = -\theta^{-1}\ln(1-p)$ as $\eta \to 0$.
- *Lomax limit ($\eta \to 0$).* For $\eta = 0$, $F_Y(y) = 1 - (\frac{\lambda}{\lambda+\theta y})^\lambda$, so

$$q_p \;=\; \frac{\lambda}{\theta}\Big[(1-p)^{-1/\lambda} - 1\Big] \qquad (\eta = 0).$$

Practical note. For general (θ,λ,η), compute q_p by one–dimensional root finding on $F_Y(y) - p = 0$. Use the small-p and large-p approximations above for reliable initialization, and the closed-form EG ($\lambda \to \infty$) and Lomax ($\eta = 0$) expressions when those limits apply.

4. Estimation

4.1. Maximum likelihood estimation

Let $Y_1, \ldots, Y_n$ be i.i.d. from the unified model with density

$$f_Y(y \mid \theta,\lambda,\eta) = (1-\eta)\sum_{m=1}^{\infty} \eta^{m-1}\,\frac{m\theta\,\lambda^{\lambda+1}}{(m\theta y + \lambda)^{\lambda+1}}, \qquad \theta > 0,\ \lambda > 0,\ \eta \in (0,1).$$

For each observation $y_i > 0$, define

$$A_{m,i} := \frac{m\theta\,\lambda^{\lambda+1}}{(m\theta y_i + \lambda)^{\lambda+1}}, \quad S_i := \sum_{m=1}^{\infty} \eta^{m-1} A_{m,i}, \quad f_i = (1-\eta)S_i,$$

so that the log-likelihood is

$$\ell(\theta,\lambda,\eta) = \sum_{i=1}^{n} \log f_i = \sum_{i=1}^{n} [\log(1-\eta) + \log S_i].$$

Termwise differentiation is valid on interior compacts (Appendix A), and the existence and local uniqueness of the MLE on the interior (with boundary faces EG/Lomax included by continuity) follow from the conditions in Appendix B. The score functions are

$$\frac{\partial \ell}{\partial \eta} = \sum_{i=1}^{n} \left(-\frac{1}{1-\eta} + \frac{\sum_{m\geq 1}(m-1)\eta^{m-2}A_{m,i}}{S_i} \right),$$

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{n} \frac{\sum_{m\geq 1} \eta^{m-1}\, \partial_\theta A_{m,i}}{S_i}, \qquad \frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} \frac{\sum_{m\geq 1} \eta^{m-1}\, \partial_\lambda A_{m,i}}{S_i},$$

with numerically stable multiplicative forms

$$\partial_\theta A_{m,i} = A_{m,i}\left[\frac{1}{\theta} - \frac{(\lambda+1)\, m y_i}{m\theta y_i + \lambda} \right],$$

$$\partial_\lambda A_{m,i} = A_{m,i}\left[\ln \lambda + 1 - \ln(m\theta y_i + \lambda) - \frac{\lambda+1}{m\theta y_i + \lambda} \right].$$

We maximize ℓ using a quasi-Newton method (e.g., L–BFGS–B) under the reparameterization $(\log\theta, \log\lambda, \operatorname{logit}\eta)$ to enforce parameter bounds. Infinite series are evaluated via adaptive truncation using the tail bounds in Appendix A and accumulated with a log–sum–exp routine to prevent underflow or overflow. Standard errors are computed from the observed information matrix $J_n(\widehat{\boldsymbol{\vartheta}}) = -\partial^2 \ell / \partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\vartheta}^\top$.

In the Lomax limit ($\eta \to 0$), only the $m = 1$ term remains:

$$f_Y(y \mid \theta, \lambda) = \frac{\theta\, \lambda^{\lambda+1}}{(\lambda+\theta y)^{\lambda+1}}, \quad \ell(\theta,\lambda) = n\ln\theta + n(\lambda+1)\ln\lambda - (\lambda+1)\sum_{i=1}^{n} \ln(\lambda+\theta y_i).$$

The corresponding score equations are

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} - (\lambda+1)\sum_{i=1}^{n} \frac{y_i}{\lambda+\theta y_i}, \qquad \frac{\partial \ell}{\partial \lambda} = n\left(\ln\lambda + \frac{\lambda+1}{\lambda} \right) - \sum_{i=1}^{n}\left[\ln(\lambda+\theta y_i) + \frac{\lambda+1}{\lambda+\theta y_i} \right],$$

which coincide with the standard MLE equations for the Lomax distribution, confirming the coherence of the unified model's limiting behavior. Existence of a maximizer (interior or on the EG/Lomax faces) and local uniqueness near the true value are guaranteed under the regularity conditions in Appendix B.

4.2. EM algorithm

We treat the component counts $N_1, \ldots, N_n$ as latent. Throughout, we use the same notation as in Section 4.1: the observed data are $y_1, \ldots, y_n > 0$, the parameter is $\boldsymbol{\vartheta} = (\theta, \lambda, \eta)$ with $\theta > 0$, $\lambda > 0$, and $\eta \in (0,1)$, and $N \sim \text{Geom}(1-\eta)$ on $\{1, 2, \ldots\}$ with $\Pr(N = m) = (1-\eta)\eta^{m-1}$.

Given $N_i = m$, the conditional density is

$$f_{Y|N}(y_i \mid m; \theta, \lambda) = \frac{m\theta\, \lambda^{\lambda+1}}{(m\theta y_i + \lambda)^{\lambda+1}}.$$

Hence the complete-data log-likelihood is

$$\ell_c(\theta,\lambda,\eta) = \sum_{i=1}^{n}\sum_{m=1}^{\infty} \mathbf{1}\{N_i = m\}\Big\{ \log(1-\eta) + (m-1)\log\eta + \log[m\theta\, \lambda^{\lambda+1}] - (\lambda+1)\log(m\theta y_i + \lambda) \Big\}.$$

E-step. At iteration t, define the posterior ("responsibility") of component count m for observation i as

$$p_{i,m}^{(t)} := \Pr(N_i = m \mid Y_i = y_i; \theta^{(t)}, \lambda^{(t)}, \eta^{(t)}) = \frac{(1-\eta^{(t)})\,\eta^{(t)\,m-1}\, f_{Y|N}(y_i \mid m; \theta^{(t)}, \lambda^{(t)})}{\sum_{r=1}^{\infty}(1-\eta^{(t)})\,\eta^{(t)\,r-1}\, f_{Y|N}(y_i \mid r; \theta^{(t)}, \lambda^{(t)})}.$$

Let the posterior mean count be $\widehat{N}_i^{(t)} := \sum_{m\geq 1} m\, p_{i,m}^{(t)}$. The EM Q-function is

$$Q(\theta, \lambda, \eta \mid \theta^{(t)}, \lambda^{(t)}, \eta^{(t)}) = \sum_{i=1}^{n}\sum_{m=1}^{\infty} p_{i,m}^{(t)}\Big\{\log(1-\eta) + (m-1)\log\eta + \log[m\theta\,\lambda^{\lambda+1}] - (\lambda+1)\log(m\theta y_i + \lambda)\Big\}.$$

M-step. *Update for η.* Maximizing Q w.r.t. η yields the closed form

$$\eta^{(t+1)} = \frac{\sum_{i=1}^{n}(\widehat{N}_i^{(t)} - 1)}{\sum_{i=1}^{n}\widehat{N}_i^{(t)}}.$$

Updates for θ and λ. Differentiating Q gives the (expected) score equations

$$\frac{\partial Q}{\partial \theta} = \sum_{i=1}^{n}\sum_{m=1}^{\infty} p_{i,m}^{(t)}\left[\frac{1}{\theta} - \frac{(\lambda+1)\,m\,y_i}{\lambda + m\theta y_i}\right] = 0,$$

$$\frac{\partial Q}{\partial \lambda} = \sum_{i=1}^{n}\sum_{m=1}^{\infty} p_{i,m}^{(t)}\left[\frac{1}{\lambda} + \log\lambda - \log(\lambda + m\theta y_i) - \frac{\lambda+1}{\lambda + m\theta y_i}\right] = 0.$$

We solve these two equations numerically (e.g., Newton or quasi-Newton) to obtain $(\theta^{(t+1)}, \lambda^{(t+1)})$.

To ensure numerical stability and consistent bounds across the algorithm, we work with the reparameterization $(\log\theta, \log\lambda, \operatorname{logit}\eta)$ and evaluate the infinite sums in $p_{i,m}^{(t)}$ using the adaptive truncation and log-sum-exp accumulation described in Appendix A.

4.3. Estimation via importance-sampling MCEM

When the geometric series in (2.1) converges slowly (e.g., $\eta \to 1$ or small λ), direct MLE or a deterministic EM can be costly. We therefore adopt a Monte Carlo EM (MCEM) scheme where the E-step is approximated by importance sampling (IS) over the latent system size N_i, while the frailty Z_i is integrated out analytically (Rao–Blackwellization). This focuses Monte Carlo effort on a one-dimensional latent variable and reduces variance.

Complete-data structure and Q-function. Let $\boldsymbol{\vartheta} = (\theta, \lambda, \eta)$ and $y_i > 0$ for $i = 1, \ldots, n$. Conditioning on $N_i = m$ and integrating out Z_i yields the Lomax contribution

$$f_{Y|N}(y_i \mid m; \boldsymbol{\vartheta}) = \frac{m\theta\,\lambda^{\lambda+1}}{(\lambda + m\theta y_i)^{\lambda+1}}.$$

The complete-data log-likelihood contribution is

$$\ell_{c,i}(\boldsymbol{\vartheta}; N_i) = \log f_{Y|N}(y_i \mid N_i; \boldsymbol{\vartheta}) + \log(1-\eta) + (N_i - 1)\log\eta,$$

and the EM target is

$$Q(\boldsymbol{\vartheta} \mid \boldsymbol{\vartheta}^{(t)}) = \sum_{i=1}^{n} \mathbb{E}_{N_i|y_i;\boldsymbol{\vartheta}^{(t)}}[\ell_{c,i}(\boldsymbol{\vartheta}; N_i)].$$

E-step via IS over N. For each i, approximate the expectation over N_i using a truncated-geometric proposal

$$q_{N,i}(m) = \frac{(1-\eta^{(t)})(\eta^{(t)})^{m-1}}{1-(\eta^{(t)})^{M_i}}, \qquad m = 1, \ldots, M_i,$$

where M_i is chosen adaptively so the remainder of the monotone positive series for $p(N_i = m \mid y_i; \boldsymbol{\vartheta}^{(t)})$ is below a tolerance (e.g., 10^{-8}); the bound depends on y_i via $(\lambda^{(t)} + m\theta^{(t)} y_i)^{-(\lambda^{(t)}+1)}$. Draw K samples $m_{ik} \sim q_{N,i}$. The (unnormalized) IS weight is

$$w_{ik} \propto \frac{(1-\eta^{(t)})(\eta^{(t)})^{m_{ik}-1} f_{Y|N}(y_i \mid m_{ik}; \boldsymbol{\vartheta}^{(t)})}{q_{N,i}(m_{ik})}.$$

Compute weights in the log domain and normalize via a log–sum–exp routine to obtain $\tilde{w}_{ik}$. For any function g, $\mathbb{E}[g(N_i) \mid y_i; \boldsymbol{\vartheta}^{(t)}] \approx \sum_{k=1}^{K} \tilde{w}_{ik}\, g(m_{ik})$. Monitor $\mathrm{ESS}_i = 1/\sum_{k=1}^{K} \tilde{w}_{ik}^2$ and increase K or M_i until ESS_i exceeds a threshold (e.g., 200).

M-step updates. Insert the Monte Carlo expectations into $Q(\boldsymbol{\vartheta} \mid \boldsymbol{\vartheta}^{(t)})$ and maximize over (η, θ, λ).

Update for η. From $\partial Q/\partial \eta = 0$ with $\eta \in (0,1)$,

$$-\frac{n}{1-\eta} + \frac{\sum_{i=1}^{n} \mathbb{E}[N_i - 1]}{\eta} = 0 \quad \Rightarrow \quad \eta^{(t+1)} = 1 - \frac{n}{\sum_{i=1}^{n} \mathbb{E}[N_i]} = \frac{\sum_{i=1}^{n} (\mathbb{E}[N_i] - 1)}{\sum_{i=1}^{n} \mathbb{E}[N_i]}.$$

Updates for θ and λ. For $y > 0$ and $N = m$,

$$\frac{\partial}{\partial \theta} \log f_{Y|N}(y \mid m; \boldsymbol{\vartheta}) = \frac{1}{\theta} - (\lambda+1)\frac{my}{\lambda + m\theta y},$$

$$\frac{\partial}{\partial \lambda} \log f_{Y|N}(y \mid m; \boldsymbol{\vartheta}) = \log \lambda + \frac{\lambda+1}{\lambda} - \log(\lambda + m\theta y) - \frac{\lambda+1}{\lambda + m\theta y}.$$

Hence the MCEM M-step solves the *expected* score equations

$$\sum_{i=1}^{n} \mathbb{E}\left[\frac{1}{\theta} - (\lambda+1)\frac{N_i y_i}{\lambda + N_i \theta y_i}\right] = 0, \qquad \sum_{i=1}^{n} \mathbb{E}\left[\log \lambda + \frac{\lambda+1}{\lambda} - \log(\lambda + N_i \theta y_i) - \frac{\lambda+1}{\lambda + N_i \theta y_i}\right] = 0,$$

with expectations taken under $N_i \mid y_i; \boldsymbol{\vartheta}^{(t)}$ via the IS weights. We solve these with Newton–Raphson and a backtracking line search, enforcing $\theta > 0$ and $\lambda > 0$. This method efficiently handles challenging parameter regimes, as validated by simulations in Section 5.

5. Simulation study

This section presents a Monte Carlo simulation study conducted to assess the finite-sample performance of the MLE for the proposed unified distribution. The study includes 200 replicates for each of two sample sizes ($n = 100$ and $n = 500$) across four distinct parameter scenarios. Key performance metrics include bias, root mean squared error (RMSE), 95% coverage probability based on the empirical sampling distribution, and convergence rate.

Samples were generated from the unified model with parameters (θ, λ, η) under the following scenarios:

- Baseline ($\lambda = 5, \eta = 0.3$): A balanced setting with $\theta = 1.0$, $\lambda = 5.0$, and $\eta = 0.3$.
- Lomax-like ($\eta = 0.01$): A scenario mimicking Lomax-like behavior with $\theta = 1.0$, $\lambda = 5.0$, and $\eta = 0.01$.
- High Variance ($\lambda = 1.1, \eta = 0.5$): A high-variance case with $\theta = 1.0$, $\lambda = 1.1$, and $\eta = 0.5$.
- EG-like ($\lambda = 20, \eta = 0.4$): A setting resembling the exponential-geometric subfamily with $\theta = 1.0$, $\lambda = 20.0$, and $\eta = 0.4$.

For each scenario and sample size, MLEs were computed using the L-BFGS-B optimization algorithm with a numerically stable log-likelihood evaluation. Coverage probabilities were estimated using the heuristic interval $\hat{\psi} \pm 1.96 \cdot \widehat{\mathrm{sd}}(\hat{\psi})$, where $\widehat{\mathrm{sd}}(\hat{\psi})$ is the empirical standard deviation of the replicate estimates. This targets the estimator's sampling distribution rather than a single-sample Wald confidence interval.

Table 1 provides a detailed summary of the simulation outcomes, including true parameter values, bias, RMSE, and coverage probabilities. Convergence rates, derived from the summary data, are reported alongside the findings for each scenario.

Table 1. Simulation study: Bias, RMSE, and coverage of MLEs (200 replicates).

Scenario	n	Param	True	Bias	RMSE	Coverage	Converged (%)
Baseline ($\lambda = 5, \eta = 0.3$)	100	θ	1.00	0.146	0.303	0.927	75.0
	100	λ	5.00	1.852	6.656	0.933	75.0
	100	η	0.30	-0.169	0.267	0.973	75.0
	500	θ	1.00	0.090	0.208	0.895	85.5
	500	λ	5.00	-0.265	0.943	0.918	85.5
	500	η	0.30	-0.078	0.180	1.000	85.5
Lomax-like ($\eta = 0.01$)	100	θ	1.00	-0.061	0.188	0.932	88.5
	100	λ	5.00	5.569	10.544	0.842	88.5
	100	η	0.01	0.054	0.168	0.910	88.5
	500	θ	1.00	-0.027	0.085	0.940	92.0
	500	λ	5.00	0.652	1.779	0.946	92.0
	500	η	0.01	0.042	0.096	0.891	92.0
High Variance ($\lambda = 1.1$)	100	θ	1.00	-0.087	0.701	0.985	98.0
	100	λ	1.10	0.592	1.158	0.883	98.0
	100	η	0.50	-0.002	0.379	1.000	98.0
	500	θ	1.00	0.027	0.563	0.995	98.5
	500	λ	1.10	0.168	0.582	0.949	98.5
	500	η	0.50	-0.062	0.310	1.000	98.5
EG-like ($\lambda = 20, \eta = 0.4$)	100	θ	1.00	0.375	0.446	0.707	78.5
	100	λ	20.00	-7.555	12.480	1.000	78.5
	100	η	0.40	-0.320	0.354	0.236	78.5
	500	θ	1.00	0.405	0.421	0.055	73.0
	500	λ	20.00	-13.146	13.486	0.021	73.0
	500	η	0.40	-0.300	0.321	0.247	73.0

- Baseline ($\lambda = 5, \eta = 0.3$): At $n = 100$, the MLE shows a modest positive bias for θ (approximately 0.15) and λ (approximately 1.85), with a small negative bias for η (approximately -0.17). Coverage is near the nominal 0.95 (0.927–0.973), improving to 0.895–1.000 at $n = 500$, where biases decrease (for λ bias to -0.265). Convergence rates are robust at 75.0% ($n = 100$) and 85.5% ($n = 500$).
- Lomax-like ($\eta = 0.01$): With η near the boundary, θ is nearly unbiased, while η is slightly overestimated (bias 0.04–0.05), reflecting boundary estimation challenges. At $n = 100$, λ exhibits a significant positive bias (5.57) with a high RMSE (10.54) and slightly low coverage (0.842), improving to a bias of 0.652 and coverage of 0.946 at $n = 500$. Convergence rates are strong at 88.5% ($n = 100$) and 92.0% ($n = 500$).
- High Variance ($\lambda = 1.1, \eta = 0.5$): Despite heavy tails, biases are moderate, with λ bias at 0.592 (RMSE 1.16) and θ with small negative bias at $n = 100$. Coverage is near or above nominal (0.883–1.000), improving at $n = 500$ (λ bias to 0.168, coverage 0.949). Convergence is excellent at 98.0% ($n = 100$) and 98.5% ($n = 500$).
- EG-like ($\lambda = 20, \eta = 0.4$): As the frailty weakens, θ is overestimated (bias 0.375–0.405), and η is underestimated (bias -0.32 to -0.30), indicating partial confounding. λ is severely downward biased, with estimates drifting toward smaller values (e.g., $\hat{\lambda} \approx 6.9$ when $\lambda = 20$, reflecting poor identifiability in the EG-like regime), with a severe negative bias (-7.555 at $n = 100$, -13.146 at $n = 500$) and under-coverage dropping to 0.021 at $n = 500$. The coverage of 1.000 at $n = 100$ arises from extreme estimator variability (RMSE = 12.48), yielding overly wide intervals; as n increases, bias dominates and coverage collapses. Convergence rates are lower at 78.5% ($n = 100$) and 73.0% ($n = 500$), likely due to a flat likelihood surface.

The simulation results underscore the unified survival model's effectiveness for moderate parameter values, where MLEs exhibit near-zero bias for $n \geq 500$, achieving coverage above 90%. Near the Lomax limit ($\eta = 0.01$), the η estimator displays a slight positive bias yet accurately identifies the boundary, with performance enhancing as sample size increases. Conversely, in the EG-like regime ($\lambda = 20$), identifiability weakens, with η coverage dropping below 25% and λ significantly underestimated, indicative of a flattened likelihood surface. This highlights the model's robustness for typical scenarios but suggests that its performance near extreme parameter boundaries could be improved with advanced techniques, such as informative priors, model reparameterization, or profile-likelihood confidence intervals in applied contexts.

Based on these insights, the MLE emerges as the preferred method for standard parameter regimes ($\lambda > 2$ with $n \geq 500$), delivering reliable estimates with coverage exceeding 90%. However, its efficacy declines near extreme values (e.g., $\eta \approx 0$ or large λ), warranting further investigation. Future work will address these limitations through: (1) Bayesian estimation with informative priors to address identifiability challenges, (2) reparameterization (e.g., θ/λ ratios) to improve parameter estimation, and (3) adaptive optimization methods to enhance convergence in extreme cases. We also intend to extend the model to multivariate survival frameworks and test its applicability across diverse real-world datasets, leveraging the current study's findings. This simulation study provides a solid foundation for these advancements, with results available for further analysis in the supplementary materials.

6. Applications

This section evaluates the proposed unified survival model, which generalizes exponential, EG, and Lomax distributions, across three benchmark datasets: Wheaton river exceedances (right-skewed hydrology), aircraft air-conditioning failure intervals (heterogeneous repairable systems), and bladder cancer remission times (homogeneous survival). The unified model and its special cases are fitted using maximum likelihood estimation (MLE) with the L-BFGS-B algorithm, employing a convergence tolerance of 10^{-8} and a maximum of 2000 iterations. Initial parameter estimates are data-driven: $\theta_0 = 1/\overline{y}$ (inverse mean as a baseline rate), $\lambda_0 = \max(1.1, \overline{y}^2/s_y^2)^*$, and $\eta_0 = 0.5$ (a neutral midpoint for the mixing parameter). Standard errors are derived from the numerical Hessian, and 95% confidence intervals are constructed under asymptotic normality.

Model performance is assessed using information criteria (AIC, BIC, HQIC, CAIC), AIC weights, likelihood-ratio (LR) tests for nested models, and Vuong tests for non-nested comparisons. Goodness-of-fit is evaluated with Cramér-von Mises (W_n^2), Anderson-Darling (A_n^2), Watson (U_n^2), and a Kolmogorov-Smirnov-like (L_n) statistic, supplemented by diagnostic panels including probability density functions (PDFs), survival functions, P-P plots, and total time on test (TTT) curves. Table 2 summarizes the datasets.

Table 2. Summary measures for various datasets.

Data Sets	n	Min	Q1	Median	Mean	Q3	Max
Bladder Cancer	115	0.08	3.34	6.25	9.51	11.8	79.1
Air Conditioning	125	1	18	50	91.9	111	603
Wheaton River	72	0.10	2.13	9.50	12.2	20.1	64.0

Information criteria are defined as:

$$\text{AIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2q, \quad \text{BIC} = -2\ell(\hat{\boldsymbol{\theta}}) + q\log n,$$

$$\text{HQIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2q\log\log n, \quad \text{CAIC} = -2\ell(\hat{\boldsymbol{\theta}}) + q\frac{n\log n}{n-q-1},$$

where $\ell(\hat{\boldsymbol{\theta}})$ is the maximized log-likelihood, q is the number of parameters, and n is the sample size. Lower values indicate superior parsimony-adjusted fit. Goodness-of-fit statistics, comparing empirical and fitted cumulative distribution functions (CDFs), are:

- *Cramér-von Mises (W_n^2)*: $\frac{1}{12n} + \sum_{i=1}^{n}\left[F(y_{(i)};\hat{\boldsymbol{\theta}}) - \frac{2i-1}{2n}\right]^2$,
- *Anderson-Darling (A_n^2)*: $-n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln F(y_{(i)};\hat{\boldsymbol{\theta}}) + \ln(1 - F(y_{(n+1-i)};\hat{\boldsymbol{\theta}}))\right]$,
- *Watson (U_n^2)*: $\sum_{i=1}^{n}\left[F(y_{(i)};\hat{\boldsymbol{\theta}}) - \frac{2i-1}{2n} - \bar{F} + \frac{1}{2}\right]^2 + \frac{1}{12n}$, where $\bar{F} = \frac{1}{n}\sum_{i=1}^{n}F(y_{(i)};\hat{\boldsymbol{\theta}})$,
- *Kolmogorov-Smirnov-like (L_n)*: $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[F(y_{(i)};\hat{\boldsymbol{\theta}}) - \frac{i}{n}\right]^2}$.

Smaller values of these statistics indicate better model-data alignment.

*The initial λ_0 leverages the method-of-moments estimate for the gamma shape parameter ($\overline{y}^2/s_y^2$, where s_y^2 is the sample variance), with a minimum value of 1.1 to ensure a finite mean ($\lambda > 1$) and enhance numerical stability during optimization.

Table 3. Parameter estimates with standard errors and 95% confidence intervals.

Distribution	θ (SE, 95% CI)	λ (SE, 95% CI)	η (SE, 95% CI)
	Wheaton river exceedances		
Unified	0.0901 (0.00071, [0.0887, 0.0914])	6.33 (0.0007, [6.33, 6.34])	0.0849 (0.00071, [0.0836, 0.0864])
EG	0.0687 (0.0180, [0.0410, 0.1149])	–	0.303 (0.303, [0.0254, 0.879])
Lomax	0.0836 (0.0109, [0.0648, 0.1078])	50.0 (128, [0.327, 7645])	–
Exp.	0.0819 (0.0096, [0.0650, 0.1032])	–	–
	Air conditioning		
Unified	0.0106 (0.0005, [0.0096, 0.0117])	5.58 (2.87, [2.04, 15.3])	0.333 (0.117, [0.151, 0.583])
EG	0.0070 (0.0017, [0.0044, 0.0111])	–	0.571 (0.146, [0.292, 0.811])
Lomax	0.0143 (0.0022, [0.0105, 0.0194])	3.99 (2.06, [1.45, 10.97])	–
Exp.	0.0109 (0.0010, [0.0091, 0.0130])	–	–
	Bladder cancer		
Unified	0.116 (0.0150, [0.0901, 0.1497])	10.8 (10.7, [1.57, 74.8])	0.000 (–[†], [0, 0.001])
EG	0.101 (0.0201, [0.0687, 0.1495])	–	0.0681 (0.302, [0.0000, 0.999])
Lomax	0.116 (0.0150, [0.0901, 0.1497])	10.8 (10.7, [1.57, 74.8])	–
Exp.	0.105 (0.0098, [0.0876, 0.1263])	–	–

Table 4. Model fit comparison: Information criteria and goodness-of-fit statistics.

Distribution	$-\log L$	AIC	BIC	HQIC	CAIC	W_n^2	A_n^2	U_n^2	L_n	AIC Weight
					Wheaton river exceedances					
Unified	252.5	511	518	513.7	511.4	0.188	1.09	0.188	0.0515	0.046
EG	251.8	507.5	512.1	509.2	508.0	0.185	1.06	0.185	0.0510	0.267
Lomax	252.1	508.3	512.8	510.0	509.8	0.225	1.41	0.213	0.0577	0.184
Exp.	252.1	506.3	508.5	507.2	508.6	0.231	1.46	0.216	0.0586	0.503
					Air conditioning					
Unified	687.0	1380	1388	1383	1379	0.0428	0.319	0.0425	0.0191	0.123
EG	686.3	1376.7	1382.3	1379.0	1377.6	0.0373	0.285	0.0373	0.0176	0.634
Lomax	687.5	1379.0	1384.6	1381.3	1379.9	0.0685	0.451	0.0640	0.0246	0.204
Exp.	690.1	1382.3	1385.1	1383.4	1385.2	0.417	2.26	0.232	0.0605	0.039
					Bladder cancer					
Unified	373.3	752.5	760.8	755.9	751.4	0.181	1.15	0.166	0.0385	0.115
EG	373.95	751.9	757.4	754.1	752.8	0.167	1.06	0.167	0.0383	0.157
Lomax	373.3	750.5	756.0	752.8	751.4	0.181	1.15	0.166	0.0385	0.313
Exp.	374.0	749.95	752.7	751.1	752.8	0.160	1.01	0.159	0.0379	0.416

6.1. Wheaton river exceedances data

This dataset ($n = 72$) comprises exceedances (m^3/s above threshold) of flood peaks from 1958–1984, rounded to one decimal place [25], exhibiting right-skewness (skewness = 1.5, kurtosis = 3.19):

> 1.7, 1.4, 0.6, 9.0, 5.6, 1.5, 2.2, 18.7, 2.2, 1.7, 30.8, 2.5, 14.4, 8.5, 39.0, 7.0, 13.3, 27.4, 1.1, 25.5, 0.3, 20.1, 4.2, 1.0, 0.4, 11.6, 15.0, 0.4, 25.5, 27.1, 20.6, 14.1, 11.0, 2.8, 3.4, 20.2, 5.3, 22.1, 7.3, 14.1, 11.9, 16.8, 0.7, 1.1, 22.9, 9.9, 21.5, 5.3, 1.9, 2.5, 1.7, 10.4, 27.6, 9.7, 13.0, 14.4, 0.1, 10.7, 36.4, 27.5, 12.0, 1.7, 1.1, 30.0, 2.7, 2.5, 9.3, 37.6, 0.6, 3.6, 64.0, 27.0

The exponential model provides the best fit ($\hat{\theta} = 0.0819$, SE = 0.0096, 95% CI = [0.0650, 0.1032], AIC = 506.3), followed by EG ($\hat{\theta} = 0.0687$, $\hat{\eta} = 0.303$, 95% CI = [0.0254, 0.879], AIC = 507.5), Lomax ($\hat{\theta} = 0.0836$, $\hat{\lambda} = 50.0$, 95% CI = [0.327, 7645], AIC = 508.3), and Unified ($\hat{\theta} = 0.0901$, $\hat{\lambda} = 6.33$, $\hat{\eta} = 0.0849$, 95% CI = [0.0836, 0.0864], AIC = 511). Likelihood-ratio (LR) tests (EG vs. Exponential, $\Lambda = 0.735$, $p = 0.391$) and Vuong tests (Unified vs. Exponential, $V = -0.348$, $p = 0.728$) support the exponential model, with no significant improvement from Lomax vs. Unified ($\Lambda = -0.761$, $p = 1.000$). AIC weights are 0.503 (Exponential), 0.267 (EG), 0.184 (Lomax), and 0.046 (Unified). The TTT curve's convexity indicates a decreasing failure rate (DFR), aligning with

Exponential and EG (Figure 6).

Diagnostic plots reveal Exponential's strong alignment with the right-skewed histogram and rapid survival function decay, matching the data's drop-off. Unified and EG effectively capture the upper tail, while Lomax shows minor misfit at lower values. The P-P plot confirms overall goodness-of-fit across models, with points near the diagonal, though Lomax deviates slightly at extremes.

6.2. Air conditioning failure time data

This dataset ($n = 125$) records inter-failure times (hours) from seven Boeing 720 aircraft [26, 27]. Representative values include:

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

The EG model offers the best fit ($\hat{\theta} = 0.0070$, SE = 0.0017, 95% CI = [0.0044, 0.0111], $\hat{\eta} = 0.571$, SE = 0.146, 95% CI = [0.292, 0.811], AIC = 1376.7), followed by Unified ($\hat{\theta} = 0.0106$, $\hat{\lambda} = 5.58$, 95% CI = [2.04, 15.3], $\hat{\eta} = 0.333$, 95% CI = [0.151, 0.583], AIC = 1380), Lomax ($\hat{\theta} = 0.0143$, $\hat{\lambda} = 3.99$, 95% CI = [1.45, 10.97], AIC = 1379.0), and Exponential ($\hat{\theta} = 0.0109$, 95% CI = [0.0091, 0.0130], AIC = 1382.3). The LR test (EG vs. Exponential, $\Lambda = 7.57$, $p = 0.006$) and Vuong test (Unified vs. Exponential, $V = 1.247$, $p = 0.212$) favor EG, with no significant difference between Lomax and Unified ($\Lambda = 0.990$, $p = 0.320$). AIC weights are 0.634 (EG), 0.204 (Lomax), 0.123 (Unified), and 0.039 (Exponential). The TTT curve's convexity indicates a DFR, supporting EG and Unified (Figure 7).

The PDF plot demonstrates EG's ability to capture the histogram's peak and heterogeneous tail, while Exponential overestimates the middle range. The survival function plot shows EG's slower decay, fitting longer intervals (e.g., 502, 603), with Unified closely aligned. The P-P plot confirms good fit for EG and Unified, with Exponential deviating in upper quantiles.

6.3. Bladder cancer remission time data

This dataset ($n = 115$) records remission times (months) from [28], ranging from 0.08 to 79.1 months:

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.1, 23.6, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.3, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.8, 25.7, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.2, 25.8, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.8, 26.3, 0.81, 2.62, 3.82, 5.32, 7.32, 10.1, 14.8, 32.2, 2.64, 3.88, 5.32, 7.39, 10.3, 14.8, 34.3, 0.90, 2.69, 4.18, 5.34, 7.59, 10.7, 16.0, 36.7, 1.05, 2.69, 4.23, 5.41, 7.62, 10.8, 16.6, 43.0, 1.19, 2.75, 4.26, 5.41, 7.62, 17.1, 46.1, 1.26, 2.83, 4.33, 5.49, 7.66, 11.3, 17.1, 79.1, 1.35, 2.87, 5.62, 7.87, 11.6, 17.4, 1.40, 3.02, 4.34, 5.71, 7.93, 12.0, 2.02, 3.31, 4.50, 6.25, 8.37, 12.0, 20.3, 2.02, 3.36, 6.76, 12.1, 21.7, 2.07, 3.36, 6.93, 8.65, 12.6, 22.7

The Exponential model provides the best fit ($\hat{\theta} = 0.105$, SE = 0.0098, 95% CI = [0.0876, 0.1263], AIC = 749.95), closely followed by Lomax ($\hat{\theta} = 0.116$, $\hat{\lambda} = 10.8$, 95% CI = [1.57, 74.8], AIC = 750.5), with Unified ($\hat{\theta} = 0.116$, $\hat{\lambda} = 10.8$, $\hat{\eta} = 0.000$, 95% CI = [0, 0.001], AIC = 752.5) and EG ($\hat{\theta} = 0.101$, $\hat{\eta} = 0.0681$, 95% CI = [0.0000, 0.999], AIC = 751.9) trailing. The LR test (EG vs. Exponential, $\Lambda = 0.048$, $p = 0.827$) and Vuong test (Unified vs. Exponential, $V = 0.542$, $p = 0.588$) support Exponential, with no significant difference between Lomax and Unified ($\Lambda = -0.000$, $p = 1.000$). The MLE of $\eta = 0$ for Unified reduces it to Lomax. This boundary MLE is admissible and corresponds to the Lomax face discussed in Appendix B. AIC weights are 0.416 (Exponential), 0.313 (Lomax), 0.157 (EG), and 0.115 (Unified). The TTT curve's convexity indicates a DFR, aligning with Exponential and Lomax (Figure 8).

The PDF plot shows Exponential fitting the histogram's peak and tail well, with Unified and Lomax matching the upper tail, while EG underperforms at lower values. The survival function plot reflects Exponential's rapid decay, consistent with the data, and Unified's slower decline. The P-P plot confirms good fit across models, with EG showing minor deviations at lower quantiles.

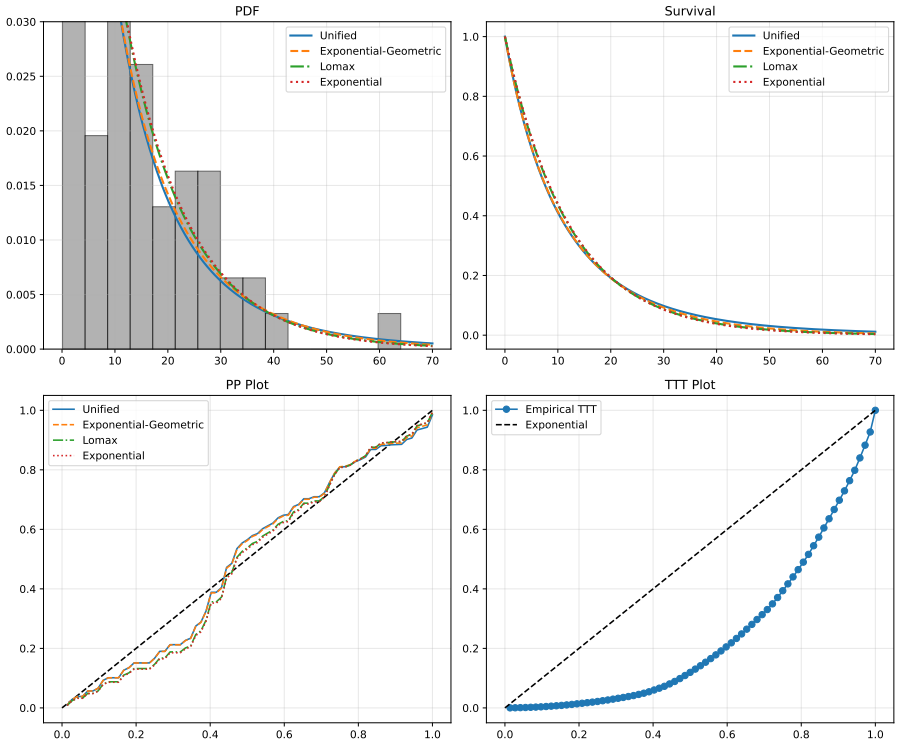

Figure 6. Survival function comparisons for Wheaton river exceedances.

The analysis highlights context-specific model strengths: Exponential's simplicity excels for homogeneous data (Wheaton river AIC = 506.3, Bladder cancer AIC = 749.95), while EG's flexibility addresses heterogeneous, heavy-tailed data (Air conditioning AIC = 1376.7). The unified model demonstrates adaptability, recovering limiting cases (e.g., Lomax when $\eta = 0$, Exponential for homogeneity), and supports applications in hydrology, engineering, and medical contexts with interpretable parameters.

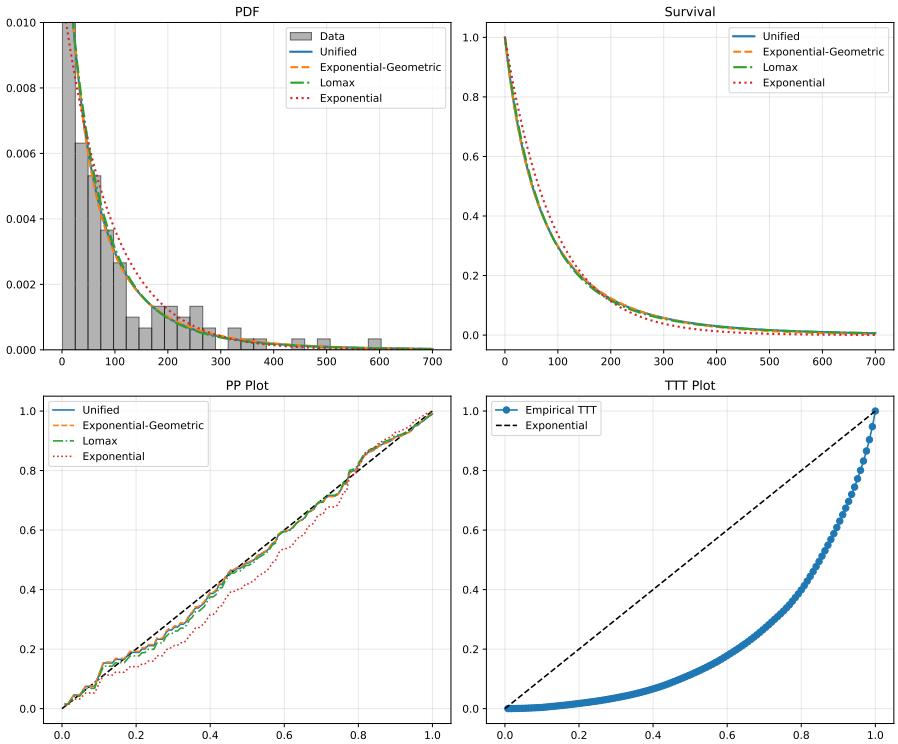

Figure 7. Survival function comparisons for air conditioning failure times.

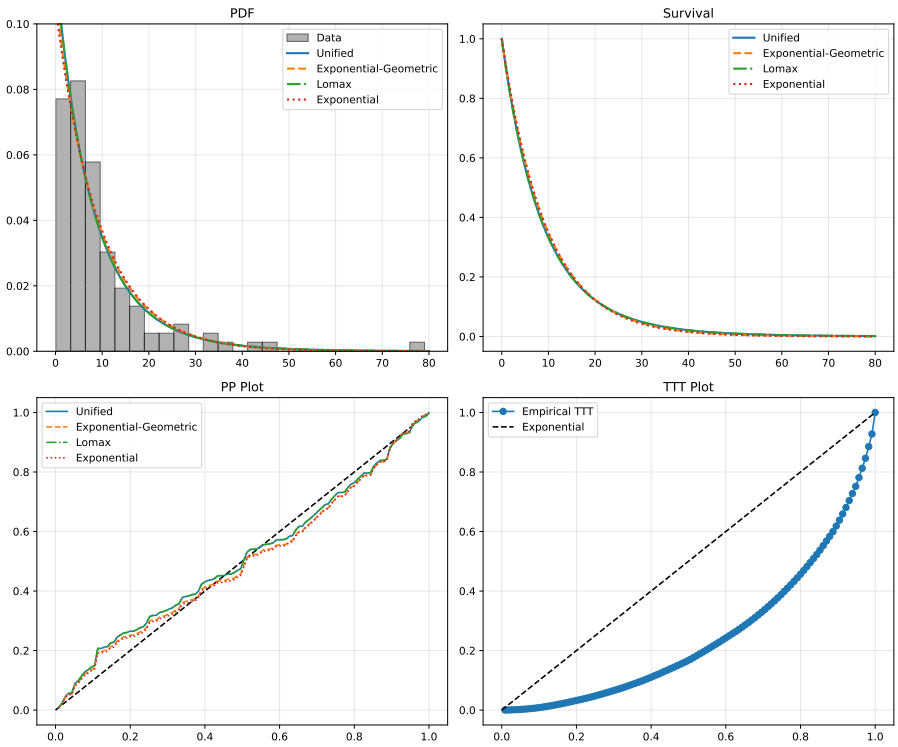

Figure 8. Survival function comparisons for bladder cancer remission times.

7. Conclusions

We proposed a flexible survival model for first-failure times that unifies the exponential-geometric and Lomax distributions through a gamma frailty and random system size. The model is interpretable, mathematically tractable, and generalizes classical distributions. Applications to real-world datasets demonstrate the versatility of the unified model as an alternative fit. The model effectively captures the homogeneous nature of the bladder cancer data with the exponential limit and the heterogeneity of the air conditioning data with the EG structure, validating its adaptability. Future work includes regression extensions, multivariate generalizations, and applications to clustered survival data.

Use of Generative-AI tools declaration

The author declares that the Generative-AI tools were not used 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. KFU254035].

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. J. Guo, X. Kong, N. Wu, L. Xie, Evaluating the lifetime distribution parameters and reliability of products using successive approximation method, *Qual. Reliab. Eng. Int.*, **40** (2024), 3280–3303. http://doi.org/10.1002/qre.3559
2. M. Xu, H. Mao, q-weibull distributions: Perspectives and applications in reliability engineering, *IEEE Trans. Reliab.*, **74** (2025), 3112–3125. http://doi.org/10.1109/tr.2024.3448289
3. Y. Yang, W. Wang, X. Zhang, Y. Xiong, G. Wang, Lifetime data modelling and reliability analysis based on modified weibull extension distribution and bayesian approach, *J. Mech. Sci. Technol.*, **32** (2018), 5121–5126. http://doi.org/10.1007/s12206-018-1009-8
4. M. E. Bakr, O. S. Balogun, A. A. El-Toony, A. M. Gadallah, Reliability analysis for unknown age class of lifetime distribution with real applications in medical science, *Symmetry*, **16** (2024), 1514. http://doi.org/10.3390/sym16111514
5. N. Feroze, U. Tahir, M. Noor-ul Amin, K. S. Nisar, M. S. Alqahtani, M. Abbas, et al., Applicability of modified weibull extension distribution in modeling censored medical datasets: A bayesian perspective, *Sci. Rep.*, **12**. http://doi.org/10.1038/s41598-022-21326-w
6. I. Daigo, K. Iwata, M. Oguchi, Y. Goto, Lifetime distribution of buildings decided by economic situation at demolition: D-based lifetime distribution, *Proc. CIRP*, **61** (2017), 146–151. http://doi.org/10.1016/j.procir.2016.11.221

7. C. Kleiber, S. Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences*, Hoboken: John Wiley & Sons, Inc, 2003. http://doi.org/10.1002/0471457175
8. M. Bhattacharjee, 11 Reliability ideas and applications in economics and social sciences, In: *Handbook of Statistics*, Elsevier, 1988, 175–213. http://doi.org/10.1016/s0169-7161(88)07013-0
9. N. M. Alfaer, A. M. Gemeay, H. M. Aljohani, A. Z. Afify, The extended log-logistic distribution: Inference and actuarial applications, *Mathematics*, **9** (2021), 1386. http://doi.org/10.3390/math9121386
10. S. Chowdhury, A. K. Nanda, A new lifetime distribution with applications in inventory and insurance, *Int. J. Qual. Reliab. Manag.*, **35** (2018), 527–544. http://doi.org/10.1108/ijqrm-12-2016-0227
11. F. H. Riad, A. Radwan, E. M. Almetwally, M. Elgarhy, A new heavy tailed distribution with actuarial measures, *J. Radiat. Res. Appl. Sci.*, **16** (2023), 100562. http://doi.org/10.1016/j.jrras.2023.100562
12. M. Rahmouni, D. Ziedan, The weibull-generalized shifted geometric distribution: properties, estimation, and applications, *AIMS Mathematics*, **10** (2025), 9773–9804. http://doi.org/10.3934/math.2025448
13. Y. Wang, X. Liu, H. Wang, X. Wang, X. Wang, Improved fatigue failure model for reliability analysis of mechanical parts inducing stress spectrum, *Proc. Inst. Mech. Eng.,: J. Risk Reliab.*, **235** (2021), 973–981. http://doi.org/10.1177/1748006x20987402
14. L. Yao, T. Cheng, J. Luo, X. Liu, Weibull distribution-based reliability evaluation of cutting tool via improved bayesian-bootstrap method, *Proc. Inst. Mech. Eng.,: J. Risk Reliab.*, 2025. http://doi.org/10.1177/1748006x251350513
15. M. Rahmouni, A new compound family of weibull order statistics and left k-truncated power series distributions, *Int. J. Anal. Appl.*, **23** (2025), 106. http://doi.org/10.28924/2291-8639-23-2025-106
16. K. Adamidis, S. Loukas, A lifetime distribution with decreasing failure rate, *Stat. Probab. Lett.*, **39** (1998), 35–42. http://doi.org/10.1016/s0167-7152(98)00012-1
17. M. Rahmouni, A. Orabi, The exponential-generalized truncated geometric (EGTG) distribution: A new lifetime distribution, *Int. J. Stat. Probab.*, **7** (2018), 1. http://doi.org/10.5539/ijsp.v7n1p1
18. D. D. Hanagal, Frailty models in public health, In: *Disease Modelling and Public Health*, Elsevier, 2017, 209–247. http://doi.org/10.1016/bs.host.2017.07.004
19. P. Hougaard, Frailty models for survival data, *Lifetime Data Anal.*, **1** (1995), 255–273. http://doi.org/10.1007/bf00985760
20. C. Lai, M. Izadi, Generalized logistic frailty model, *Stat. Probab. Lett.*, **82** (2012), 1969–1977. http://doi.org/10.1016/j.spl.2012.06.012
21. T. A. Balan, H. Putter, A tutorial on frailty models, *Stat. Methods Med. Res.*, **29** (2020), 3424–3454. http://doi.org/10.1177/0962280220921889
22. K. S. Lomax, Business failures: Another example of the analysis of failure data, *J. Amer. Stat. Assoc.*, **49** (1954), 847–852. http://doi.org/10.1080/01621459.1954.10501239

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

24. G. J. McLachlan, D. Peel, *Finite Mixture Models*, New York: John Wiley & Sons, 2000.

25. V. Choulakian, M. A. Stephens, Goodness-of-Fit Tests for the Generalized Pareto Distribution, *Technometrics*, **43** (2001), 478–484. http://doi.org/10.1198/00401700152672573

26. M. Kayid, A. Abouammoh, G. Alomani, Scaled-invariant extended quasi-lindley model: Properties, estimation, and application, *Symmetry*, **15** (2023), 1780. http://doi.org/10.3390/sym15091780

27. F. Proschan, Theoretical explanation of observed decreasing failure rate, *Technometrics*, **5** (1963), 375–383. http://doi.org/10.1080/00401706.1963.10490105

28. E. T. Lee, J. W. Wang, *Statistical Methods for Survival Data Analysis*, Hboken: John Wiley & Sons, 2003. http://doi.org/10.1002/0471458546

29. J. Wolfowitz, On Wald's proof of the consistency of the maximum likelihood estimate, *Ann. Math. Stat.*, **20** (1949), 601–602. http://doi.org/10.1214/aoms/1177729953

30. W. K. Newey, D. McFadden, *Large sample estimation and hypothesis testing*, In: *Handbook of Econometrics, vol. 4*, Elsevier, 2111–2245, 1994. http://doi.org/10.1016/s1573-4412(05)80005-4

Appendix A. Truncation and numerical stabilization (simple rules)

A.1. Series template

The series $\mathcal{S}(y;\boldsymbol{\vartheta}) = \sum_{n=1}^{\infty} T_n(y;\boldsymbol{\vartheta})$ represents a monotone, positive infinite series, where each term is given by $T_n(y;\boldsymbol{\vartheta}) = c(y;\boldsymbol{\vartheta})\frac{\eta^{n-1}}{(\lambda+n\theta y)^{\alpha}}$, with parameters $\boldsymbol{\vartheta} = (\theta, \lambda, \eta)$, constraints $0 \le \eta < 1$, $\alpha > 0$, $y > 0$, and $c(y;\boldsymbol{\vartheta}) > 0$ independent of n. For fixed $(y, \boldsymbol{\vartheta})$, $T_n > 0$ and $T_{n+1} \le T_n$ due to the decaying η^{n-1} and increasing denominator.

For the marginal density $f_Y(y)$, $c = (1-\eta)\theta\lambda^{\lambda+1}$ and $\alpha = \lambda + 1$; for the survival function $S_Y(y)$, $c = (1-\eta)\lambda^{\lambda}$ and $\alpha = \lambda$.

A.2. Tail bounds

Let $S_M = \sum_{n=1}^{M} T_n$ denote the partial sum and $R_M = \sum_{n=M+1}^{\infty} T_n$ the remainder.

Ratio bound (conservative). The ratio $\rho_M = \frac{T_{M+1}}{T_M} = \eta\left(\frac{\lambda+M\theta y}{\lambda+(M+1)\theta y}\right)^{\alpha} \in (0,1)$ provides a conservative upper bound for the remainder:

$$R_M \le \frac{T_{M+1}}{1-\rho_M} = \frac{T_M\rho_M}{1-\rho_M}.$$

As $M \to \infty$, $\rho_M \uparrow \eta$, ensuring the bound remains valid and slightly conservative.

Integral-test bound (tight for large y or M). A tighter bound for large y or M is:

$$R_M \le \frac{c(y;\boldsymbol{\vartheta})}{1-\eta}\frac{\eta^{M-1}}{(\lambda+M\theta y)^{\alpha}}.$$

A.3. Stopping rule

Select an absolute tolerance $\varepsilon = 10^{-8}$. Increment $M = 1, 2, \ldots$ until the tighter bound satisfies:

$$\min\left\{\frac{T_{M+1}}{1-\rho_M}, \frac{c(y;\boldsymbol{\vartheta})}{1-\eta}\frac{\eta^{M-1}}{(\lambda+M\theta y)^{\alpha}}\right\} \le \varepsilon.$$

For a relative tolerance, substitute ε with $\varepsilon_{\mathrm{rel}}(S_M + \widehat{R}_M)$, where $\widehat{R}_M$ is the current remainder estimate.

A.4. Stable evaluation (log–sum–exp)

To mitigate underflow or overflow, evaluate the series in the log domain:

$$a_n = \log T_n = \log c(y;\boldsymbol{\vartheta}) + (n-1)\log\eta - \alpha\log(\lambda + n\theta y),$$

where $a_{\max} = \max_{1\le n\le M} a_n$. The log-sum is then:

$$\log S_M = a_{\max} + \log\left(\sum_{n=1}^{M} e^{a_n - a_{\max}}\right),$$

and the normalized weights are $\tilde{w}_n = \frac{e^{a_n - a_{\max}}}{\sum_{j=1}^{M} e^{a_j - a_{\max}}}$. For ratios of two series, compute each $\log S_M$ independently and take the difference.

Appendix B. Existence and uniqueness of the MLE

Let $Y_1, \ldots, Y_n > 0$ be independent and identically distributed (i.i.d.) observations from the density

$$f(y \mid \theta, \lambda, \eta) = (1-\eta)\sum_{m=1}^{\infty} \eta^{m-1}\frac{m\theta\lambda^{\lambda+1}}{(\lambda+m\theta y)^{\lambda+1}}, \quad \theta > 0, \quad \lambda > 0, \quad \eta \in [0,1),$$

as defined in Eq (2.1). Denote the parameter vector by $\boldsymbol{\vartheta} = (\theta, \lambda, \eta)$, and let the log-likelihood be $\ell_n(\boldsymbol{\vartheta}) = \sum_{i=1}^{n} \log f(Y_i \mid \boldsymbol{\vartheta})$.

B.1. Existence

Continuity

The series is uniformly convergent on compact subsets of $(0,\infty)\times(0,\infty)\times[0,1)$ for fixed $y > 0$, as established by the tail bounds in Appendix A. Consequently, the density $f(y \mid \boldsymbol{\vartheta})$ and the log-likelihood $\ell_n(\boldsymbol{\vartheta})$ are continuous on this domain.

Coercivity (natural boundaries)

For any dataset with $Y_i > 0$, the log-likelihood $\ell_n(\boldsymbol{\vartheta})$ tends to $-\infty$ as $\theta \downarrow 0$, $\theta \uparrow \infty$, $\lambda \downarrow 0$, or $\eta \uparrow 1$, because $f(Y_i \mid \boldsymbol{\vartheta}) \to 0$ in each case. This occurs due to a vanishing prefactor (e.g., $\theta \downarrow 0$ or $\lambda \downarrow 0$) or a denominator blow-up (e.g., $\theta \uparrow \infty$), ensuring that ℓ_n attains its maximum on any compact subset of the interior $(0,\infty)\times(0,\infty)\times[0,1)$.

Limit faces (EG and Lomax)

The model admits two key limiting cases: (i) $\eta = 0$, reducing to the Lomax distribution, and (ii) $\lambda \to \infty$, corresponding to the exponential-geometric (EG) distribution with no frailty. The log-likelihood ℓ_n extends continuously to these boundary faces. By compactifying the parameter space to include these limits, ℓ_n attains its maximum either in the interior of $(0, \infty) \times (0, \infty) \times [0, 1)$ or on one of these faces, consistent with the data analyses in Section 6. This follows from standard extremum-estimation existence theorems [29, 30].

B.2. Local uniqueness

On compact subsets of the interior, termwise differentiation of the series is justified by dominated convergence (as per Appendix A), rendering the score function and observed information matrix $J_n(\boldsymbol{\vartheta})$ continuous. The Fisher information $I(\boldsymbol{\vartheta}) = \mathbb{E}[J_1(\boldsymbol{\vartheta})]$ is finite for $\lambda > 2$, reflecting the existence of sufficient moments. Under standard regularity conditions—namely, the true parameter lies in the interior, the model is identifiable, the density is thrice differentiable, and integrable envelopes exist—the MLE is consistent and asymptotically normal, with a non-singular $I(\boldsymbol{\vartheta}_0)$. With probability approaching one, ℓ_n exhibits strict concavity in a neighborhood of the true $\boldsymbol{\vartheta}_0$, ensuring the MLE is the unique maximizer locally.

AIMS Press