AIMS Mathematics, 11(8): 24445–24489.
DOI: 10.3934/math.2026987
Received: 17 June 2026
Revised: 16 July 2026
Accepted: 29 July 2026
Published: 10 August 2026

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

Research article

A one-parameter decreasing-hazard lifetime model with exact inference on a bounded-below support

Khudhayr A. Rashedi[1], **L. S. Diab**[2], **Abdullah H. Alenezy**[1] **and Ghareeb A. Marei**[3,*]

[1] Department of Mathematics, College of Science, University of Ha'il, Ha'il, Saudi Arabia

[2] Department of Mathematics and Statistics, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia

[3] Faculty of Computers and Information Systems, Egyptian Chinese University, Nasr City, Egypt

* **Correspondence:** Email: ghareeb.adel@ecu.edu.eg.

Abstract: We introduce the *modified exponential root distribution* (MERD), a one-parameter lifetime model on $[1, \infty)$ defined as the square of a unit-shifted exponential variate. Being a monotone transform of an exponential variable, MERD admits closed-form distribution, survival, hazard (strictly decreasing), cumulative and reversed hazard, quantile, and mean-residual-life functions, along with closed-form moments, quantile-based shape measures, order statistics, entropies, and stress-strength reliability via a single root substitution. Since the transformed data are exactly exponential, the pivot $2\lambda S \sim \chi^2_{2n}$ gives exact confidence intervals for the rate parameter at any sample size, unlike most comparably simple lifetime models. We derive closed-form maximum-likelihood and conjugate-gamma Bayesian estimators and evaluate their bias, mean squared error (MSE), and coverage via simulation. The MERD exhibits competitive performance with respect to one- and two-parameter models on the parsimony-penalized information criteria when evaluated on empirical datasets of biomedical survival and actuarial science, while also providing exact inference. We emphasize that these findings are dataset-specific and do not constitute a claim of universal superiority over all competing models; rather, they illustrate the practical utility of the MERD as a tractable and exactly-inferable option for bounded decreasing-hazard data. To address the fixed lower bound and scale, we develop an identifiable three-parameter location-scale extension, show its threshold is estimated by the sample minimum, and examine how far the exact-inference properties persist once threshold and scale are estimated.

Keywords: one-parameter lifetime model; decreasing hazard rate; exact confidence interval; transformed exponential; Bayesian estimation
Mathematics Subject Classification: 62E15, 62F10, 62N05

1. Introduction

One-parametric lifetime distributions are appreciated for their simplicity: a single parameter decreases complexity for both estimating and interpreting the model and decreases over-fitting risk, particularly for small sample sizes. A couple of traditional examples are the exponential distribution (characterized by a constant hazard) and the Rayleigh distribution (characterized by an increasing hazard). Renewed interest in flexible one-parameter families has come from the Lindley distribution and its several relatives [1–5], which we survey in detail in Section 2. Models with a strictly *decreasing* failure rate (DFR) remain comparatively scarce, yet they are important for systems dominated by early failures ("burn-in") and for waiting and relief times concentrated near a known lower bound [6, 7], a gap that recent one-parameter proposals such as the extended Lindley model [8] have also begun to address.

A large number of one-parameter lifetime models have been produced in recent years by transforming, mixing, or compounding a baseline distribution. It is now widely recognized that such constructions add value only when they bring a genuine analytical or inferential advantage rather than merely a new name, and that fit comparisons on a handful of small, much-reused datasets are weak evidence of practical superiority. We therefore state plainly what the modified exponential root distribution (MERD) does and does not offer. By construction, the model is the square of a unit-shifted exponential variate, so it inherits the exponential's tractability and every functional form below reduces to an exponential integral; it is not a new mixture or compounded family. Its distinguishing features are threefold: (i) a strictly decreasing hazard on a known positive lower bound, matching early-failure and relief-time data whose support begins above zero; (ii) closed-form maximum-likelihood and Bayesian estimators; (iii) the feature we regard as most useful, *exact* pivotal inference, since the transformed data are exactly exponential and $2\lambda S \sim \chi^2_{2n}$, giving exact confidence intervals at any sample size, including the small samples typical of reliability work. We do not claim that the MERD dominates richer models; we present it as a fully worked, exactly-inferable option for a specific and common data shape.

Target data characteristics and empirical scope. The MERD framework is designed to address lifetime data with a known positive lower bound (support $[1, \infty)$) where the hazard rate exhibits a strictly decreasing trend. Examples of such data include early failure data, data arising from burn-in processes, and data pertaining to wait times. To assess the framework's practical utility, two diverse empirical cases are analyzed: (i) biomedical survival data (acute myelogenous leukemia) and (ii) skewed actuarial and/or insurance data (vehicle insurance complaint ratios), demonstrating its scope across health and finance domains.

The rest of this paper is structured in the following manner. In Section 2, we place the model in the context of the literature on transformed exponential families and decreasing-hazard families. In Section 3, we construct the model and collect its functional forms. Section 4 is deliberately placed next, ahead of the distributional-properties catalogue, because the exact-inference machinery it develops, the maximum-likelihood and Bayesian estimation procedures and the exact Bayesian pivot, is the paper's central contribution rather than a downstream consequence of the properties that follow. Section 5 then collects the model's distributional properties. Section 6 extends the construction to an unknown lower threshold and scale. We describe the simulation study in Section 7, and Section 8 extends it with a predictive comparison against one-parameter competitors and a study of robustness under

misspecification, contamination, and censoring. In Section 9, we demonstrate the model on real data with a deliberately conservative interpretation of the evidence. Finally, we provide a summary in Section 10.

2. Related work

The MERD belongs to the broad effort to enrich the exponential distribution so as to accommodate non-constant hazards while keeping estimation simple. It is useful to locate it against four strands of that literature.

One-parameter Lindley-type families. The Lindley distribution [1, 2] and the subsequent Akash, Shanker, and Sujatha models [3–5] are one-parameter mixtures of exponential and gamma components. They achieve flexibility with a single parameter but place the mode in the interior of the support, so they are ill-suited to data with a spike at the lower bound; this is visible in our applications (Section 9), where these models systematically misfit the early peak. The MERD shares their parsimony but, being monotone-decreasing, captures the boundary mode directly. Recent extensions such as the extended Lindley model [8] further illustrate the ongoing interest in one-parameter families that balance tractability and flexibility, though they remain mode-interior and do not provide an exact pivot.

Compounding and mixing constructions with decreasing hazard. A large family of two-parameter DFR models is obtained by compounding the exponential with a discrete law: the exponential-geometric distribution [9], its generalization [10], the exponential-Poisson [11], the exponential-logarithmic [12], the power-series compounding framework that unifies them [13], and the exponential Poisson-Lindley model [14]. These models are flexible and motivated by physical "competing-risk" or "first-activation" mechanisms, but they require numerical (often EM-based) estimation and do not admit an exact pivot for the rate. The MERD trades their two-parameter flexibility for closed-form, exactly-inferable inference on a fixed positive support.

Transformation-based generators. Rather than mixing, one may apply a deterministic generator to a baseline cumulative distribution function (CDF) or variate. Examples include the exponentiated (power) transformation that produced the generalized exponential [15], the alpha-power transformation, and its exponential extensions [16], and the more recent Dinesh–Umesh–Sanjay generator applied to the exponential baseline [17, 18]. The MERD is of this transformation type, but it is generated by a particularly elementary map, the square of a unit-shifted exponential, chosen precisely so that the inverse transform returns the data to an exact exponential sample. This is what distinguishes it operationally from the generators above: most transformation generators improve fit at the cost of analytic and inferential tractability, whereas here the transformation is invertible to a known-rate exponential, preserving exact inference.

Threshold (shifted) lifetime models. Reliability data whose support begins at a known positive value are classically handled by the two-parameter shifted exponential and, more generally, by threshold (three-parameter) Weibull and gamma models [7]. These introduce a location parameter at the cost of an extra parameter and boundary-estimation difficulties. The MERD instead fixes the threshold at unity and devotes its single parameter to the hazard, which is appropriate when the lower bound is known a priori; Section 6 derives an identifiable location-scale generalization for the case where the bound and

scale are themselves unknown, and characterizes its estimation and the limits of exact inference in that setting.

Stress-strength and record/censoring-based inference. Beyond decreasing-hazard modeling per se, a related and active strand of the reliability literature develops stress-strength reliability estimation and inference under record and progressive/hybrid censoring schemes for one- and few-parameter lifetime families [19–23]. The stress-strength results we derive for the MERD in Section 5.7 connect directly to this body of work, and the threshold-estimation and censoring results of Sections 6 and 8 extend the present model in a compatible direction.

Against this background, the contribution of the present paper is deliberately narrow: not another flexible multi-parameter family, but a fully worked, single-parameter, decreasing-hazard model on a known positive support whose square-root transform yields exact, finite-sample inference. We are explicit (here and in Section 9) that flexibility is not the selling point and that richer two-parameter models can fit better when the threshold and scale are unknown.

3. Mathematical construction and functional forms

3.1. Construction

Let $T \sim \text{Exp}(\lambda)$ with rate parameter $\lambda > 0$, so that $f_T(t) = \lambda e^{-\lambda t}$ for $t \geq 0$. We define the transformation

$$X = (T + 1)^2. \tag{3.1}$$

Since $T \geq 0$, it follows that $X \geq 1$. The mapping $t \mapsto (t + 1)^2$ is strictly increasing on $[0, \infty)$ with inverse $T = \sqrt{X} - 1$.

Definition 1. *A random variable X is said to follow the MERD with parameter $\lambda > 0$, denoted $X \sim$ MERD(λ), if its CDF is*

$$F(x) = \Pr(X \leq x) = \Pr(T \leq \sqrt{x} - 1) = 1 - e^{-\lambda(\sqrt{x}-1)}, \qquad x \geq 1. \tag{3.2}$$

Differentiating Eq (3.2) with respect to x yields the probability density function (PDF):

$$f(x) = \frac{\lambda}{2\sqrt{x}}\, e^{-\lambda(\sqrt{x}-1)}, \qquad x \geq 1. \tag{3.3}$$

Throughout, we use the transformation

$$t = \sqrt{x} - 1 \quad \Longleftrightarrow \quad x = (t + 1)^2, \qquad \mathrm{d}x = 2(t + 1)\,\mathrm{d}t, \tag{3.4}$$

which maps the MERD to a standard Exp(λ) distribution on t. We call this the *root substitution*.

3.2. Reliability and distributional functions

Based on Eqs (3.2) and (3.3), for $x \geq 1$, the survival function $S(x)$, hazard rate $h(x)$, and cumulative hazard $H(x)$, and for $x > 1$, the reversed hazard rate $\tau(x)$, are

$$S(x) = e^{-\lambda(\sqrt{x}-1)}, \tag{3.5}$$

$$h(x) = \frac{f(x)}{S(x)} = \frac{\lambda}{2\sqrt{x}}, \tag{3.6}$$

$$H(x) = -\ln S(x) = \lambda(\sqrt{x} - 1), \tag{3.7}$$

$$\tau(x) = \frac{f(x)}{F(x)} = \frac{\lambda\, e^{-\lambda(\sqrt{x}-1)}}{2\sqrt{x}\,(1 - e^{-\lambda(\sqrt{x}-1)})}, \qquad x > 1, \tag{3.8}$$

since $F(1) = 0$ makes $\tau(x)$ undefined at the lower endpoint $x = 1$; $S(x)$, $h(x)$, and $H(x)$ remain well defined there ($S(1) = 1$, $h(1) = \lambda/2$, $H(1) = 0$). By Eq (3.6), $h(x)$ is strictly decreasing in x, so the MERD has a DFR. The quantile and mean-residual-life functions, derived in Section 5.1 and Proposition 10, are recorded here for convenience:

$$Q(p) = \left(1 - \tfrac{1}{\lambda}\ln(1-p)\right)^2, \qquad 0 < p < 1, \tag{3.9}$$

$$m(x) = \mathbb{E}[X - x \mid X > x] = \frac{2\sqrt{x}}{\lambda} + \frac{2}{\lambda^2}, \qquad x \geq 1. \tag{3.10}$$

As illustrated in Figure 1, the PDF exhibits a reverse-J shape with its peak at the lower bound $x = 1$. Larger λ concentrates the probability mass nearer the lower bound, steepening the CDF so that probability accumulates almost entirely just above $x = 1$.

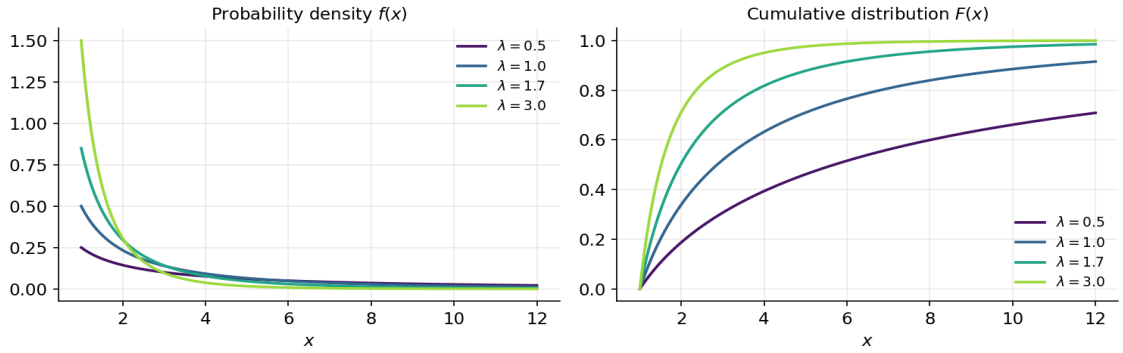


Figure 1. Density (left) and CDF (right) of the MERD(λ).

Figure 2 confirms the DFR property visually: The hazard rate strictly decreases as x increases, and the decline accelerates with larger λ while the survival curve drops more sharply.

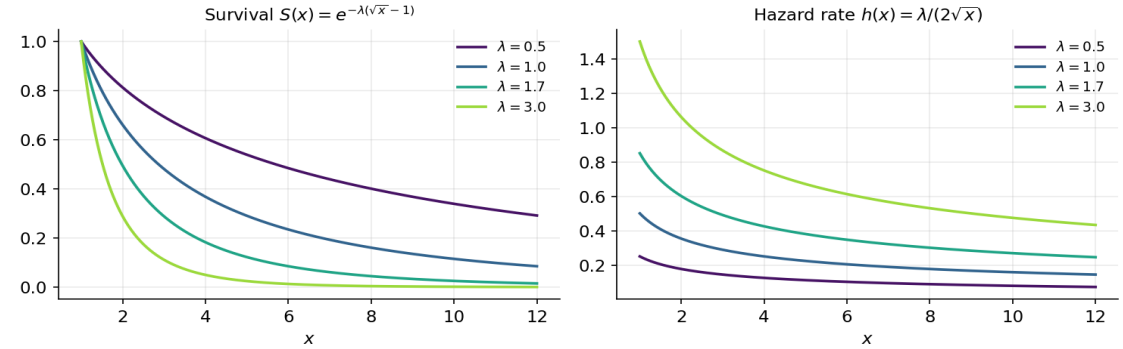


Figure 2. Survival function Eq (3.5) (left) and hazard rate Eq (3.6) (right).

Figure 3 demonstrates that $H(x)$ is concave with respect to x (being linear in $\sqrt{x}$), which is the analytical signature of a DFR model, and that the reversed hazard shows a downward trend when moving away from the lower bound due to the early concentration of mass.

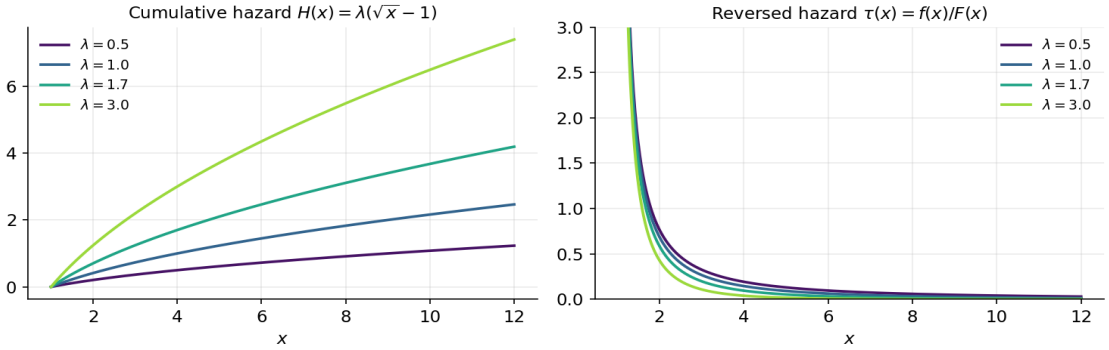


Figure 3. Cumulative hazard Eq (3.7) (left) and reversed hazard Eq (3.8) (right).

Remark 1. *(Relationship to the exponential distribution) By construction* $\sqrt{X} - 1 = T \sim \text{Exp}(\lambda)$*, so the MERD is a strictly monotone transformation of an exponential variate rather than a new mixture or compounded family. This is precisely why every functional form above is elementary: Under the root substitution Eq* (3.4)*, each MERD expectation reduces to a standard exponential integral. Two consequences deserve emphasis. First, the transformed sample* $t_i = \sqrt{x_i} - 1$ *is* exactly *a sample from* Exp(λ)*, which is what makes exact inference available (Section 4.1). Second, the lower bound at* $x = 1$ *and the unit scale are fixed structural features of the basic model, appropriate when the lower bound is known a priori.*

4. Parameter estimation

We place this section immediately after the construction (Section 3) and ahead of the distributional-properties catalogue of Section 5, since the exact-inference results developed here are the paper's principal contribution rather than a routine consequence of the properties collected afterward; Section 5 may be skipped on a first reading without loss of continuity. We estimate the rate λ in both the frequentist and the Bayesian framework. Throughout, write

$$t_i = \sqrt{x_i} - 1, \quad S = \sum_{i=1}^{n} t_i. \tag{4.1}$$

By Remark 1, the transformed observations t_i are an exact i.i.d. sample from Exp(λ). Hence, S is a sum of n independent exponentials,

$$S \sim \text{Gamma}(n, \lambda) \quad (\text{shape } n, \text{ rate } \lambda), \quad \text{so} \quad 2\lambda S \sim \chi^2_{2n}, \tag{4.2}$$

the pivot holding for *every* sample size n, not merely asymptotically. This single distributional fact, a direct consequence of the root substitution, underlies all of the exact inference developed below, and is, in our view, the model's principal inferential advantage over comparably simple one-parameter competitors.

4.1. Maximum likelihood estimation

The log-likelihood is

$$\ell(\lambda) = n \ln \lambda - n \ln 2 - \tfrac{1}{2} \sum_{i=1}^{n} \ln x_i - \lambda S. \tag{4.3}$$

The score is $\ell'(\lambda) = n/\lambda - S$ and the observed information $\ell''(\lambda) = -n/\lambda^2 < 0$, so ℓ is strictly concave, and the likelihood equation $\ell'(\lambda) = 0$ admits the unique maximizer

$$\hat{\lambda}_{\mathrm{ML}} = \frac{n}{S} = \frac{n}{\sum_{i=1}^{n} (\sqrt{x_i} - 1)}. \tag{4.4}$$

The Fisher information per observation is $I(\lambda) = -\mathbb{E}[\ell''(\lambda)]/n = 1/\lambda^2$, whence $\sqrt{n}\,(\hat{\lambda}_{\mathrm{ML}} - \lambda) \xrightarrow{d} \mathcal{N}(0, \lambda^2)$ and the approximate 95% Wald interval is $\hat{\lambda}_{\mathrm{ML}} \pm 1.96\,\hat{\lambda}_{\mathrm{ML}}/\sqrt{n}$. The exponential structure, however, lets us go well beyond this asymptotic statement and characterize the estimator *exactly* at any sample size.

Exact sampling distribution and finite-sample optimality. Because $\hat{\lambda}_{\mathrm{ML}}$ is a deterministic function of the Gamma-distributed sufficient statistic S, its entire small-sample behavior is available in closed form.

Proposition 1. *(Exact distribution and finite-sample properties of the maximum likelihood estimation (MLE)) Let $n \geq 3$. With $S \sim \mathrm{Gamma}(n, \lambda)$ as in Eq (4.2), the MLE Eq (4.4) follows the inverse-gamma law*

$$\hat{\lambda}_{\mathrm{ML}} \sim \text{Inv-Gamma}(n,\, n\lambda), \qquad f_{\hat{\lambda}_{\mathrm{ML}}}(y) = \frac{(n\lambda)^n}{\Gamma(n)}\, y^{-(n+1)}\, e^{-n\lambda/y}, \quad y > 0, \tag{4.5}$$

or, equivalently, $2n\lambda/\hat{\lambda}_{\mathrm{ML}} = 2\lambda S \sim \chi^2_{2n}$. Its exact mean, bias, variance, and mean squared error are

$$\begin{aligned} \mathbb{E}[\hat{\lambda}_{\mathrm{ML}}] &= \frac{n}{n-1}\lambda, \\ \mathrm{Bias}(\hat{\lambda}_{\mathrm{ML}}) &= \frac{\lambda}{n-1}, \\ \mathrm{Var}(\hat{\lambda}_{\mathrm{ML}}) &= \frac{n^2\lambda^2}{(n-1)^2(n-2)}, \\ \mathrm{MSE}(\hat{\lambda}_{\mathrm{ML}}) &= \frac{(n+2)\,\lambda^2}{(n-1)(n-2)}. \end{aligned} \tag{4.6}$$

Proof. For $S \sim \mathrm{Gamma}(n, \lambda)$ and any integer $k < n$,

$$\mathbb{E}\left[S^{-k}\right] = \frac{\lambda^n}{\Gamma(n)} \int_0^\infty s^{n-k-1} e^{-\lambda s}\, ds = \lambda^k\, \frac{\Gamma(n-k)}{\Gamma(n)}.$$

In particular, $\mathbb{E}[S^{-1}] = \lambda/(n-1)$ and $\mathbb{E}[S^{-2}] = \lambda^2/[(n-1)(n-2)]$. Since $\hat{\lambda}_{\mathrm{ML}} = n/S$, the change of variable $s = n/y$ in the Gamma density gives Eq (4.5); the equivalence with χ^2_{2n} is immediate from Eq (4.2). Then,

$$\mathbb{E}[\hat{\lambda}_{\mathrm{ML}}] = n\,\mathbb{E}[S^{-1}] = \frac{n\lambda}{n-1}, \qquad \mathbb{E}[\hat{\lambda}^2_{\mathrm{ML}}] = n^2\,\mathbb{E}[S^{-2}] = \frac{n^2\lambda^2}{(n-1)(n-2)},$$

and the bias, variance, and MSE = Var + Bias^2 in (4.6) follow after using $n^2 + n - 2 = (n+2)(n-1)$. □

The MLE is therefore positively biased by exactly $\lambda/(n-1)$, a bias that vanishes at the parametric rate and is negligible for the sample sizes in our simulation study (Section 7). Two exact corrections follow at no extra cost and reinforce the model's small-sample tractability.

Corollary 1. *(Exactly unbiased and minimum-MSE estimators) For $n \geq 3$, the estimator*

$$\tilde{\lambda} = \frac{n-1}{S}$$

is exactly unbiased, $\mathbb{E}[\tilde{\lambda}] = \lambda$, with $\mathrm{MSE}(\tilde{\lambda}) = \lambda^2/(n-2)$. Moreover, within the class $\{c/S : c > 0\}$, the mean squared error

$$\mathrm{MSE}(c/S) = \lambda^2(n-1)^{-2}[c^2/(n-2) + (c-(n-1))^2]$$

is minimized at $c^* = n-2$, giving

$$\hat{\lambda}^* = \frac{n-2}{S}, \qquad \mathrm{MSE}(\hat{\lambda}^*) = \frac{\lambda^2}{n-1}.$$

Proof. Both statements use

$$\mathbb{E}[S^{-1}] = \lambda/(n-1) \quad \text{and} \quad \mathbb{E}[S^{-2}] = \lambda^2/[(n-1)(n-2)]$$

from Proposition 1. For $\hat{\lambda}_c = c/S$, one has

$$\mathbb{E}[\hat{\lambda}_c] = c\lambda/(n-1) \quad \text{and} \quad \mathrm{Var}(\hat{\lambda}_c) = c^2\lambda^2/[(n-1)^2(n-2)],$$

so the displayed $\mathrm{MSE}(c/S)$ follows; setting $c = n-1$ gives unbiasedness, and differentiating in c yields the optimizer $c^* = n-2$. □

Thus, $\mathrm{MSE}(\hat{\lambda}^*) < \mathrm{MSE}(\tilde{\lambda}) < \mathrm{MSE}(\hat{\lambda}_{\mathrm{ML}})$ for all $n \geq 3$, with all three ratios tending to 1 as $n \to \infty$. We retain $\hat{\lambda}_{\mathrm{ML}}$ as the default in the applications of Section 9 for comparability with the competing fits, but report that the bias-free and minimum-MSE variants are available in identical closed form should small-sample accuracy be the priority.

Exact confidence interval. Inverting the pivot Eq (4.2) gives, for any n and any level $1-\alpha$, the exact equal-tailed confidence interval

$$\left[\frac{\chi^2_{2n,\,\alpha/2}}{2S},\ \frac{\chi^2_{2n,\,1-\alpha/2}}{2S}\right], \tag{4.7}$$

where $\chi^2_{2n,q}$ denotes the q-quantile of the χ^2_{2n} law. Unlike the Wald interval, Eq (4.7) attains its nominal coverage *by construction* at every sample size, a property confirmed numerically in Section 7 down to $n = 20$. The same pivot delivers an exact test of H_0: $\lambda = \lambda_0$ against any alternative, with p-value read directly from the χ^2_{2n} distribution of $2\lambda_0 S$. No comparable exact pivot is available for the Lindley, Akash, Shanker, or Sujatha models, whose inference is necessarily asymptotic.

4.2. Bayesian estimation

The likelihood is proportional to $\lambda^n e^{-\lambda S}$, so the gamma family is conjugate. With prior $\lambda \sim \text{Gamma}(a, b)$, $\pi(\lambda) \propto \lambda^{a-1} e^{-b\lambda}$, the posterior is

$$\lambda \mid \mathbf{x} \sim \text{Gamma}(a + n,\ b + S). \tag{4.8}$$

Under squared-error loss, the Bayes estimator is the posterior mean $\hat{\lambda}_{\text{SE}} = (a + n)/(b + S)$; under the LINEX loss $L(\Delta) = e^{c\Delta} - c\Delta - 1$, it is

$$\hat{\lambda}_{\text{BL}} = \tfrac{a+n}{c} \ln(1 + \tfrac{c}{b+S}).$$

The equal-tailed $100(1 - \alpha)\%$ credible interval is $[G^{-1}_{a+n,\,b+S}(\alpha/2),\ G^{-1}_{a+n,\,b+S}(1 - \alpha/2)]$, where $G^{-1}_{\alpha',\beta'}$ is the Gamma(α', β') quantile function; the HPD interval is obtained by the usual equal-density-endpoint condition on Eq (4.8).

For the data analyses in Section 9, we adopt the Jeffreys prior $\pi(\lambda) \propto \sqrt{I(\lambda)} = 1/\lambda$ (the limit $a, b \to 0$), giving the posterior $\lambda \mid \mathbf{x} \sim \text{Gamma}(n, S)$ whose mean n/S coincides with the MLE Eq (4.4) while supplying a full posterior for credible statements. This choice has a sharper justification than mere objectivity.

Proposition 2. *(Exact probability matching) Under the Jeffreys prior $\pi(\lambda) \propto 1/\lambda$, the posterior Eq* (4.8) *is* Gamma(n, S)*, so that $2S\lambda \mid \mathbf{x} \sim \chi^2_{2n}$. Consequently, the equal-tailed* $100(1 - \alpha)\%$ *Bayesian credible interval coincides* numerically and exactly *with the frequentist interval Eq* (4.7)*, namely $[\chi^2_{2n,\alpha/2}/(2S),\ \chi^2_{2n,1-\alpha/2}/(2S)]$. The Jeffreys prior is therefore an exact probability-matching prior for λ.*

Proof. With $a, b \to 0$ in Eq (4.8) the posterior is Gamma(n, S), for which the scale change $\lambda \mapsto 2S\lambda$ gives

$$2S\lambda \mid \mathbf{x} \sim \text{Gamma}(n, \tfrac{1}{2}) = \chi^2_{2n}.$$

The posterior quantiles of λ are thus $\chi^2_{2n,q}/(2S)$, identical to the endpoints of Eq (4.7). Because the frequentist pivot $2\lambda S \sim \chi^2_{2n}$ is itself exact, the agreement is exact rather than first-order. □

This coincidence is a structural benefit of the construction: The same χ^2_{2n} object serves simultaneously as the frequentist pivot and the posterior law, so the two inferential schools return the same intervals at every sample size.

Remark 2. *(Exact posterior prediction) Mixing the exponential observation density over the Jeffreys posterior gives, for a future transformed value*

$$T_{new} = \sqrt{X_{new}} - 1,$$

the Lomax predictive law

$$\Pr(T_{new} > t \mid \mathbf{x}) = \{S/(S + t)\}^n.$$

Translating through $t = \sqrt{x} - 1$ yields the posterior predictive survival of a future MERD observation in closed form,

$$\Pr(X_{new} > x \mid \mathbf{x}) = \left(\frac{S}{S + \sqrt{x} - 1}\right)^n, \qquad x \geq 1,$$

from which exact Bayesian prediction intervals follow by inversion.

4.3. Prior sensitivity analysis

The closed-form posterior Eq (4.8) makes it straightforward to examine how much the choice of prior hyperparameters (a, b) actually moves the inference, rather than merely asserting robustness. We illustrate this on the acute myelogenous leukemia (AML) dataset of Section 9 ($n = 32, S = 141.970, MLE\hat{\lambda}_{\mathrm{ML}} = 0.2254$), comparing the Jeffreys/vague choice already used in Section 9 against informative Gamma(a, b) priors with prior mean $a/b = \lambda_0$ at increasing concentration a. Table 1 reports the resulting posterior means and credible intervals.

Table 1. Posterior mean and 95% equal-tailed credible interval for λ under alternative priors, AML dataset ($n = 32$, $S = 141.970$).

Prior	(a, b)	Posterior mean	95% credible interval (width)
Jeffreys, $\pi(\lambda) \propto 1/\lambda$	$(0, 0)$	0.2254	[0.1542, 0.3099] (0.1558)
Vague Gamma$(0.001, 0.001)$	$(0.001, 0.001)$	0.2254	[0.1542, 0.3099] (0.1558)
Informative prior centered near the MLE, $\lambda_0 = 0.20$:			
Weak	(1, 5.00)	0.2245	[0.1546, 0.3074] (0.1528)
Moderate	(5, 25.00)	0.2216	[0.1560, 0.2985] (0.1425)
Strong	(10, 50.00)	0.2188	[0.1577, 0.2897] (0.1321)
Very strong	(20, 100.00)	0.2149	[0.1605, 0.2771] (0.1166)
Informative prior in mild conflict with the MLE, $\lambda_0 = 0.35$:			
Weak	(1, 2.86)	0.2279	[0.1568, 0.3119] (0.1551)
Moderate	(5, 14.29)	0.2368	[0.1667, 0.3190] (0.1522)
Strong	(10, 28.57)	0.2463	[0.1775, 0.3261] (0.1487)
Very strong	(20, 57.14)	0.2612	[0.1950, 0.3368] (0.1417)

Two patterns emerge. First, when the prior is centered near the MLE ($\lambda_0 = 0.20$), even a fairly concentrated prior ($a = 20$) shifts the posterior mean by only about 4.7% relative to the Jeffreys/vague answer, and shortens the credible interval by roughly the same proportion as it sharpens the prior; the inference is not brittle to reasonable, data-compatible prior choices. Second, when the prior mean is set further from the data ($\lambda_0 = 0.35$), the posterior is visibly pulled toward the prior as concentration increases, reaching an 16% shift at $a = 20$; this is the expected and desired behavior of a proper Bayes procedure, and confirms that the vague/Jeffreys choice adopted for the main analyses in Section 9 is not silently hiding sensitivity to an unstated informative prior. In line with Proposition 2, the Jeffreys and vague-Gamma rows are numerically indistinguishable in this table, as expected since both are within numerical precision of the exact probability-matching prior.

Figure 4 makes the asymmetry between the two scenarios visually explicit: The near-MLE prior (blue) keeps the posterior mean close to the dashed reference line across the whole range of a, while the conflicting prior (red) pulls the posterior mean steadily upward as a increases, with the credible band narrowing around the (shifted) prior-informed value rather than around the data-driven estimate.

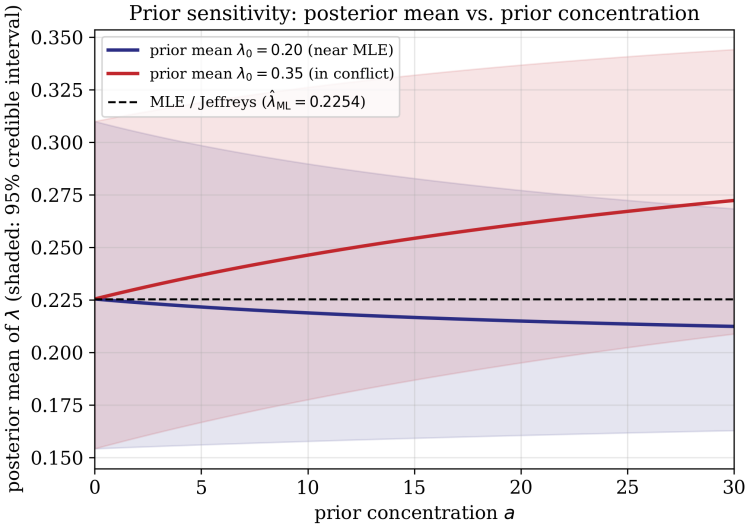


Figure 4. Posterior mean of λ (lines) and 95% credible interval (shaded bands) as a function of prior concentration a, for an informative prior near the MLE ($\lambda_0 = 0.20$) and one in mild conflict with it ($\lambda_0 = 0.35$); the dashed line marks the MLE/Jeffreys value.

5. Statistical properties and measures

This section details the mathematical properties of the MERD, each presented as a distinct subsection.

5.1. Quantile function, Bowley skewness, and Moors kurtosis

Proposition 3. *(Quantile function) For $0 < p < 1$, the quantile function of the MERD is $Q(p) = (1 - \lambda^{-1}\ln(1-p))^2$, as in Eq (3.9).*

Proof. Setting $F(x) = p$ in Eq (3.2) yields $e^{-\lambda(\sqrt{x}-1)} = 1 - p$. Taking logarithms gives $-\lambda(\sqrt{x} - 1) = \ln(1 - p)$, i.e., $\sqrt{x} = 1 - \lambda^{-1}\ln(1 - p)$. Since $0 < 1 - p < 1$, $\ln(1 - p) < 0$ and $\sqrt{x} > 1 > 0$. Squaring yields the result. □

The median and quartiles follow by substituting $p = 1/2, 1/4, 3/4$ into Eq (3.9). Equation (3.9) also provides a direct simulation rule: If $U \sim \text{Uniform}(0, 1)$, then

$$X = \left(1 - \tfrac{1}{\lambda}\ln(1 - U)\right)^2 \sim \text{MERD}(\lambda), \tag{5.1}$$

used in the simulation study (Section 7).

To describe shape robustly we employ the quantile-based *Bowley skewness* and *Moors kurtosis*:

$$B = \frac{Q(\frac{3}{4}) + Q(\frac{1}{4}) - 2Q(\frac{1}{2})}{Q(\frac{3}{4}) - Q(\frac{1}{4})}, \tag{5.2}$$

$$M = \frac{Q(\frac{7}{8}) - Q(\frac{5}{8}) + Q(\frac{3}{8}) - Q(\frac{1}{8})}{Q(\frac{6}{8}) - Q(\frac{2}{8})}. \tag{5.3}$$

Setting $s = 1/\lambda$ and noting $\sqrt{Q(p)} = 1 + c_p s$ with $c_p = -\ln(1-p)$, we have $Q(p) = 1 + 2c_p s + c_p^2 s^2$. Substituting into Eq (5.2) and using $\ln 4 = 2\ln 2$, the constant terms cancel and the coefficient of s collapses to $\ln 4 + \ln\frac{4}{3} - 2\ln 2 = \ln\frac{4}{3}$ in the numerator and $\ln 4 - \ln\frac{4}{3} = \ln 3$ in the denominator; multiplying numerator and denominator by $\lambda^2 = 1/s^2$ gives the closed form

$$B(\lambda) = \frac{2\lambda \ln\frac{4}{3} + 2\ln^2 2 + \ln^2\frac{4}{3}}{2\lambda \ln 3 + 4\ln^2 2 - \ln^2\frac{4}{3}}. \tag{5.4}$$

Remark 3. *An earlier draft of this derivation dropped a factor of* 2 *from the coefficient of* λ *in the numerator; Eq* (5.4) *is the corrected form, confirmed numerically against direct evaluation of* $B(\lambda) = [Q(3/4) + Q(1/4) - 2Q(1/2)]/[Q(3/4) - Q(1/4)]$ *from Eq* (3.9) *(e.g.,* $B(1) \approx 0.4010$*, matching Eq* (5.4) *but not the uncorrected version). Figure 5a has been regenerated from the corrected formula.*

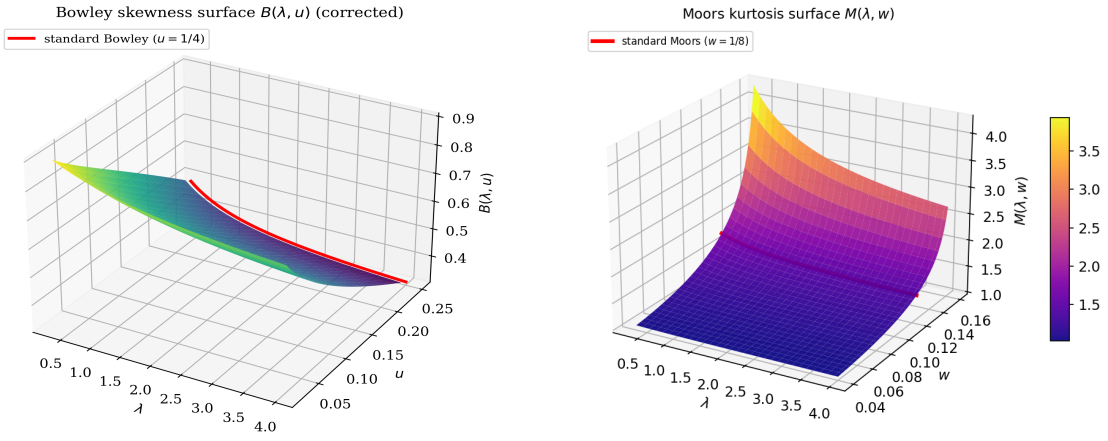
The limits are $B(\lambda) \to \frac{\ln(4/3)}{\ln 3} \approx 0.262$ as $\lambda \to \infty$ and $B(\lambda) \to 0.567$ as $\lambda \to 0$ (the leading-order term in λ is subdominant as $\lambda \to 0$, so the correction leaves this limit unchanged). Since $B(\lambda) > 0$ for all $\lambda > 0$, the MERD is strictly right-skewed.

A parallel derivation applies to the Moors kurtosis Eq (5.3). Writing $c_{1/8} = \ln\frac{8}{7}$, $c_{2/8} = \ln\frac{4}{3}$, $c_{3/8} = \ln\frac{8}{5}$, $c_{5/8} = \ln\frac{8}{3}$, $c_{6/8} = \ln 4$, and $c_{7/8} = \ln 8$, the same expansion gives, after multiplying by λ^2,

$$M(\lambda) = \frac{2\lambda \ln\frac{21}{5} + \ln^2 8 - \ln^2\frac{8}{3} + \ln^2\frac{8}{5} - \ln^2\frac{8}{7}}{2\lambda \ln 3 + 4\ln^2 2 - \ln^2\frac{4}{3}}, \tag{5.5}$$

using $\ln 8 - \ln\frac{8}{3} + \ln\frac{8}{5} - \ln\frac{8}{7} = \ln 3 + \ln\frac{7}{5} = \ln\frac{21}{5}$ for the coefficient of λ. Numerically, $M(\lambda) \to \ln(21/5)/\ln 3 \approx 1.306$ as $\lambda \to \infty$ and $M(\lambda) \to [\ln^2 8 - \ln^2\frac{8}{3} + \ln^2\frac{8}{5} - \ln^2\frac{8}{7}]/[4\ln^2 2 - \ln^2\frac{4}{3}] \approx 1.939$ as $\lambda \to 0$, confirming the two bounds already reported; both exceed the Gaussian benchmark 1.233, so the MERD is leptokurtic throughout its parameter space.

Figure 5 visualizes generalized versions of these measures over λ and varying quantile spacings. The surfaces remain strictly positive, corroborating Eq (5.4), and are stable across spacings, indicating that the right-skewness and leptokurtosis are intrinsic to the model.



(a) Bowley skewness surface $B(\lambda, u)$. **(b)** Moors kurtosis surface $M(\lambda, w)$.

Figure 5. Three-dimensional quantile-based shape measures; the red curves represent the standard slices.

5.2. Moments, mean, variance, skewness, and kurtosis

Proposition 4. *(Raw moments) For any integer $r \geq 1$, the rth raw moment of the MERD is*

$$\mu'_r = \mathbb{E}[X^r] = \sum_{k=0}^{2r} \binom{2r}{k} \frac{k!}{\lambda^k}. \tag{5.6}$$

Proof. See Section 10. □

Corollary 2. *(Mean, variance, skewness, and kurtosis) The first four raw moments yield*

$$\mathbb{E}[X] = 1 + \frac{2}{\lambda} + \frac{2}{\lambda^2}, \qquad \mathrm{Var}(X) = \frac{4(\lambda^2 + 4\lambda + 5)}{\lambda^4}, \tag{5.7}$$

$$\mu_3 = \frac{16(\lambda^3 + 9\lambda^2 + 30\lambda + 37)}{\lambda^6}, \qquad \mu_4 = \frac{48(\lambda^2 + 6\lambda + 17)(3\lambda^2 + 22\lambda + 43)}{\lambda^8}. \tag{5.8}$$

Consequently, $\gamma_1 = \mu_3/\mu_2^{3/2}$ and $\gamma_2 = \mu_4/\mu_2^2 - 3$ are strictly positive for all $\lambda > 0$, so the MERD is right-skewed and leptokurtic.

Proof. See Section 10. □

Remark 4. *Because $X = (T + 1)^2$ is a quadratic transformation of an exponential variate, the moment generating function has no elementary closed form. Nevertheless, the distribution is determined by the moment sequence Eq* (5.6)*, all of whose terms are finite.*

5.3. Mode

Proposition 5. *(Mode) The PDF Eq* (3.3) *is strictly decreasing on* $[1, \infty)$*; the mode is at $x = 1$, with $f(1) = \lambda/2$.*

Proof. $\ln f(x) = \ln(\lambda/2) - \frac{1}{2}\ln x - \lambda(\sqrt{x} - 1)$, so $\frac{\mathrm{d}}{\mathrm{d}x}\ln f(x) = -\frac{1}{2x} - \frac{\lambda}{2\sqrt{x}} < 0$ for $x \geq 1$. Hence, f is strictly decreasing, attaining its maximum at $x = 1$. □

5.4. Incomplete first moment and mean deviations

Proposition 6. *(Incomplete first moment and mean deviations) Let $a = \sqrt{x} - 1$, and*

$$\gamma(s, z) = \int_0^z y^{s-1} e^{-y}\, \mathrm{d}y$$

denote the lower incomplete gamma function. The incomplete first moment is

$$J(x) = \int_1^x w\, f(w)\, \mathrm{d}w = \frac{\gamma(3, \lambda a)}{\lambda^2} + \frac{2\,\gamma(2, \lambda a)}{\lambda} + (1 - e^{-\lambda a}). \tag{5.9}$$

The mean deviations about the mean $\mu = \mathbb{E}[X]$ and the median $\mathcal{M}$ are

$$\delta_1 = 2\mu F(\mu) - 2J(\mu), \qquad \delta_2 = \mu - 2J(\mathcal{M}). \tag{5.10}$$

Proof. See Section 10. □

5.5. Order statistics and closure under minima

Proposition 7. *(Order statistics and closure under minima) For an i.i.d. sample of size n, the PDF of the ith order statistic is*

$$f_{i:n}(x) = \frac{n!}{(i-1)!\,(n-i)!}\,F(x)^{i-1}\{1-F(x)\}^{n-i}f(x), \tag{5.11}$$

and the family is closed under minima, with $X_{(1)} \sim \mathrm{MERD}(n\lambda)$.

Proof. Equation (5.11) is the standard order-statistic formula. For the minimum,

$$\Pr(X_{(1)} > x) = S(x)^n = \left(e^{-\lambda(\sqrt{x}-1)}\right)^n = e^{-n\lambda(\sqrt{x}-1)},$$

which is Eq (3.5) with parameter $n\lambda$; hence, $X_{(1)} \sim \mathrm{MERD}(n\lambda)$. □

5.6. Entropies

Proposition 8. *(Shannon and Rényi entropies) Let* $E_1(\lambda) = \int_\lambda^\infty t^{-1}e^{-t}\,\mathrm{d}t$ *be the exponential integral and* $\Gamma(\cdot,\cdot)$ *the upper incomplete gamma function. The Shannon entropy* $\mathcal{H}$ *and Rényi entropy* $\mathcal{H}_R(\rho)$ *are*

$$\mathcal{H} = 1 + \ln\frac{2}{\lambda} + e^{\lambda}E_1(\lambda), \tag{5.12}$$

$$\mathcal{H}_R(\rho) = \frac{1}{1-\rho}\ln\Big[2\big(\tfrac{\lambda}{2}\big)^{\rho}e^{\rho\lambda}(\rho\lambda)^{\rho-2}\,\Gamma(2-\rho,\rho\lambda)\Big], \qquad 0<\rho<2,\ \rho\neq 1. \tag{5.13}$$

Proof. See Section 10. □

As shown in Figure 6, the Shannon entropy decreases monotonically in λ: Larger λ concentrates the distribution near the lower bound, reducing uncertainty.

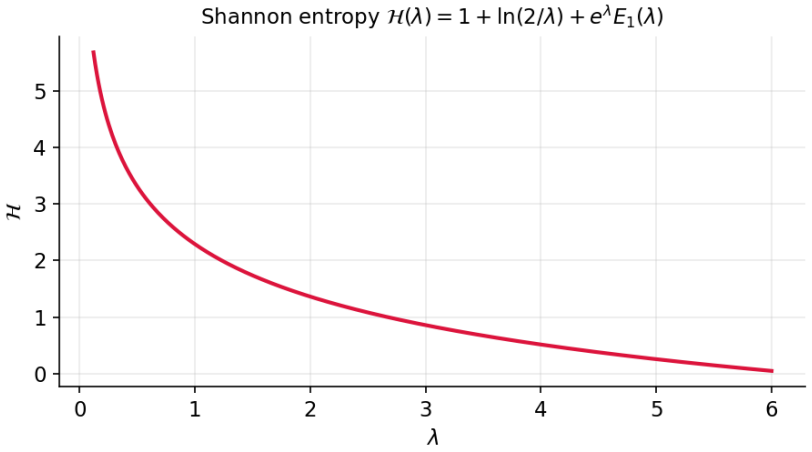


Figure 6. Shannon entropy Eq (5.12) as a function of λ.

5.7. Stress-strength reliability

Proposition 9. *(Stress-strength reliability) If* $X_1 \sim \mathrm{MERD}(\lambda_1)$ *and* $X_2 \sim \mathrm{MERD}(\lambda_2)$ *are independent, then* $R = \Pr(X_1 < X_2) = \lambda_1/(\lambda_1 + \lambda_2)$.

Proof. Conditioning on X_1, $R = \int_1^\infty S_2(x)\, f_1(x)\, \mathrm{d}x$. With the root substitution, $S_2(x) = e^{-\lambda_2 t}$ and $f_1(x)\,\mathrm{d}x = \lambda_1 e^{-\lambda_1 t}\,\mathrm{d}t$, so

$$R = \lambda_1 \int_0^\infty e^{-(\lambda_1+\lambda_2)t}\,\mathrm{d}t = \lambda_1/(\lambda_1 + \lambda_2).$$

This completes the proof. □

This closed form places the MERD in the stress-strength literature surveyed in Section 2 [19–23]: The reliability parameter R depends on the two rates only through their ratio, and, by Proposition 1, plugging in the exact finite-sample distribution of each $\hat{\lambda}_i$ would allow the sampling distribution of $\hat{R}$ to be characterized without relying on asymptotic delta-method approximations, a direction we leave for future work.

5.8. Mean residual life

Proposition 10. *For* $x \geq 1$, *the Mean residual life (MRL) function is* $m(x) = 2\sqrt{x}/\lambda + 2/\lambda^2$, *as in Eq* (3.10).

Proof. $m(x) = S(x)^{-1} \int_x^\infty S(w)\,\mathrm{d}w$. With

$$a = \sqrt{x} - 1, \quad \int_x^\infty S(w)\,\mathrm{d}w = \int_a^\infty 2(t+1)e^{-\lambda t}\,\mathrm{d}t = 2e^{-\lambda a}\left(\frac{a+1}{\lambda} + \frac{1}{\lambda^2}\right).$$

Dividing by $S(x) = e^{-\lambda a}$ and using $a + 1 = \sqrt{x}$ gives the result. Note that $m(x)$ is increasing in x, consistent with the DFR property. □

Figure 7 illustrates the MRL and quantile functions. The increasing MRL mirrors the decreasing hazard, while the quantile curve shows the heavy right tail, more pronounced for smaller λ.

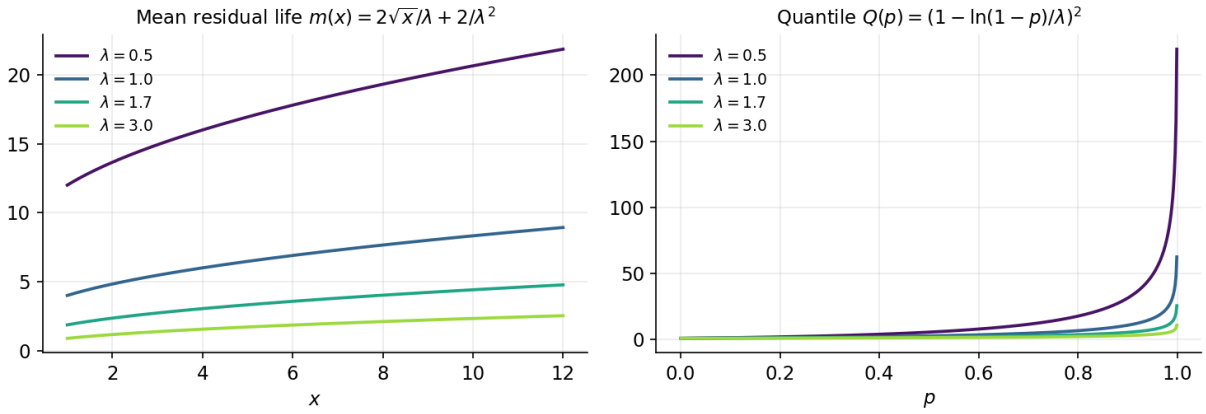


Figure 7. MRL Eq (3.10) (left) and quantile function Eq (3.9) (right).

6. Extension to an unknown threshold and scale

The basic MERD fixes the lower bound at $x = 1$ and the scale implicit in the map $t \mapsto (t + 1)^2$; Section 2 notes a location-scale generalization as a natural direction when these are not known *a priori*. We now derive such a generalization explicitly, establish which parametrization is identifiable, and characterize the resulting estimation problem, rather than leaving the extension unexamined.

6.1. A cautionary remark on parametrization

Scale can be introduced into the construction Eq (3.1) in more than one way, and not every choice yields an identifiable model. If the exponential variate itself is rescaled before the shift and square,

$$X = \tau + (\sigma T + 1)^2 - 1$$

with $T \sim \mathrm{Exp}(\lambda)$, the resulting CDF is

$$F(x) = 1 - \exp\Big[-\tfrac{\lambda}{\sigma}\big(\sqrt{x - \tau + 1} - 1\big)\Big], \quad x \geq \tau,$$

which depends on (σ, λ) only through the ratio λ/σ: any two pairs with the same ratio produce the identical distribution, so σ and λ are not separately identifiable in this parametrization. This is worth recording because it shows why a location-scale extension is not automatic and must be built with some care; it is not evidence against the existence of a genuine extension, only against this particular one.

6.2. An identifiable location-scale construction

Instead, let the scale multiply the entire squared term. For $T \sim \mathrm{Exp}(\lambda)$, threshold $\tau \in \mathbb{R}$, and scale $\sigma > 0$, define

$$X = \tau + \sigma[(T + 1)^2 - 1]. \tag{6.1}$$

Since $T \geq 0$, $X \geq \tau$, attained at $T = 0$; τ is thus the true lower bound of the support, and σ controls how quickly probability mass spreads away from it. Solving Eq (6.1) for T,

$$T = \sqrt{1 + \tfrac{x-\tau}{\sigma}} - 1, \quad x \geq \tau,$$

so that

$$F(x; \tau, \sigma, \lambda) = 1 - \exp\Big[-\lambda\Big(\sqrt{1 + \tfrac{x-\tau}{\sigma}} - 1\Big)\Big], \quad x \geq \tau, \tag{6.2}$$

$$f(x; \tau, \sigma, \lambda) = \frac{\lambda}{2\sigma\sqrt{1 + \frac{x-\tau}{\sigma}}} \exp\Big[-\lambda\Big(\sqrt{1 + \tfrac{x-\tau}{\sigma}} - 1\Big)\Big], \tag{6.3}$$

$$h(x; \tau, \sigma, \lambda) = \frac{\lambda}{2\sigma\sqrt{1 + \frac{x-\tau}{\sigma}}}, \tag{6.4}$$

and the quantile function is

$$Q(p) = \tau + \sigma[(1 - \lambda^{-1} \ln(1 - p))^2 - 1].$$

At $\tau = \sigma = 1$, Eq (6.2) reduces exactly to Eq (3.2), so the basic MERD is the special case with threshold and scale fixed at unity. Because $h(x)$ in Eq (6.4) is, as before, strictly decreasing in x, the DFR property is preserved for every (τ, σ): introducing the threshold and scale changes where the hazard starts and how fast it decays, not its monotone shape.

Proposition 11. *(Identifiability) The map* $(\tau, \sigma, \lambda) \mapsto F(\cdot; \tau, \sigma, \lambda)$ *defined by Eq* (6.2) *is injective on* $\mathbb{R} \times (0, \infty) \times (0, \infty)$.

Proof. Fix τ (identifiable directly as the infimum of the support) and suppose

$$F(x; \tau, \sigma, \lambda) = F(x; \tau, \sigma', \lambda')$$

for all $x \geq \tau$. Differentiating the exponent $-\lambda\sqrt{1 + (x - \tau)/\sigma}$ (up to the additive constant λ) with respect to x and equating gives

$$\lambda / \sqrt{\sigma(\sigma + x - \tau)} = \lambda' / \sqrt{\sigma'(\sigma' + x - \tau)}$$

for all $x \geq \tau$, i.e.,

$$\lambda^2\sigma'(\sigma' + x - \tau) = \lambda'^2\sigma(\sigma + x - \tau).$$

Matching the coefficient of x gives $\lambda^2\sigma' = \lambda'^2\sigma$, and matching the constant term then gives $\lambda^2\sigma'^2 = \lambda'^2\sigma^2$; dividing the second by the first yields $\sigma' = \sigma$, whence $\lambda' = \lambda$. □

Thus, Eq (6.1) is a genuine, identifiable three-parameter extension of the MERD, in contrast to the parametrization of Section 6.1.

6.3. MLE of the threshold

Write

$$t_i(\tau, \sigma) = \sqrt{1 + (x_i - \tau)/\sigma} - 1$$

for the root-substituted observations. As in Section 4, if τ, σ are known, the t_i are exactly i.i.d. Exp(λ) and all of the exact machinery of Sections 4.1 and 4.2 applies unchanged with

$$S = \sum_i t_i(\tau, \sigma).$$

The applied difficulty is that τ (and possibly σ) must ordinarily be estimated, and because the support $[\tau, \infty)$ depends on τ, this is a non-regular estimation problem, structurally similar to threshold estimation in the two- and three-parameter exponential, Weibull, and gamma families [7].

Proposition 12. *(The threshold MLE is the sample minimum) For any fixed* $\sigma > 0$ *and* $\lambda > 0$*, the log-likelihood Eq* (6.3) *is strictly increasing in* τ *throughout its admissible range* $\tau \leq x_{(1)} := \min_i x_i$*. Consequently, the maximum-likelihood threshold estimator is the boundary value*

$$\hat{\tau} = x_{(1)}. \tag{6.5}$$

Proof. Up to terms not involving τ, the log-likelihood is

$$\ell(\tau) = -\lambda \sum_i t_i(\tau) - \tfrac{1}{2} \sum_i \ln\{1 + (x_i - \tau)/\sigma\}.$$

Since $\partial t_i/\partial\tau = -1/[2\sigma(t_i+1)]$,

$$\frac{\partial \ell}{\partial \tau} = \lambda \sum_i \frac{1}{2\sigma(t_i+1)} + \sum_i \frac{1}{2\sigma(t_i+1)^2} > 0$$

for every τ in the admissible range (where $t_i(\tau) \geq 0$ is well defined), regardless of σ and λ. The log-likelihood is therefore strictly increasing on $(-\infty, x_{(1)}]$, so its maximum over the admissible range occurs at the boundary $\tau = x_{(1)}$. □

This mirrors the classical result for the two-parameter exponential, where the threshold MLE is likewise the sample minimum [7]; here it holds for *any* fixed scale σ, which lets estimation proceed in two tractable stages rather than a single three-dimensional search: Set $\hat{\tau} = x_{(1)}$, then maximize the profile log-likelihood

$$\ell_p(\sigma) = n \ln\Big(\frac{n}{S(\hat{\tau}, \sigma)}\Big) - n - n\ln(2\sigma) - \tfrac{1}{2}\sum_i \ln\Big(1 + \tfrac{x_i - \hat{\tau}}{\sigma}\Big)$$

numerically over the single remaining parameter $\sigma > 0$, and finally set $\hat{\lambda} = n/S(\hat{\tau}, \hat{\sigma})$. If instead the scale is fixed at its basic-model value $\sigma = 1$ (a pure threshold extension), Eq (6.5) applies directly with

$$\hat{\lambda} = n/\sum_i [\sqrt{x_i - \hat{\tau} + 1} - 1]$$

in closed form, requiring no numerical search at all.

6.4. Asymptotic behavior and the limits of exact inference

Proposition 13. *(Asymptotic distribution of the threshold estimator) Let $(\tau_0, \sigma_0, \lambda_0)$ denote the true parameters and $\hat{\tau} = x_{(1)}$ as in Eq* (6.5)*. Then,*

$$\frac{2n\lambda_0}{\sigma_0}(\hat{\tau} - \tau_0) \xrightarrow{d} \chi^2_2 \qquad \text{as } n \to \infty.$$

Proof. By Proposition 7 (applied to the transformed observations), the minimum $T_{(1)}$ of n i.i.d. $\text{Exp}(\lambda_0)$ variates satisfies $T_{(1)} \sim \text{Exp}(n\lambda_0)$ exactly for every n, so $nT_{(1)} \sim \text{Exp}(\lambda_0)$ exactly. From Eq (6.1),

$$X_{(1)} - \tau_0 = \sigma_0[T^2_{(1)} + 2T_{(1)}],$$

so

$$\frac{n}{2\sigma_0}(X_{(1)} - \tau_0) = nT_{(1)} + \frac{(nT_{(1)})^2}{2n} \xrightarrow{d} \text{Exp}(\lambda_0)$$

as $n \to \infty$, since $nT_{(1)} = O_p(1)$ makes the second term vanish in probability. Multiplying by $2\lambda_0$ converts the exponential limit to χ^2_2. □

Two consequences temper how far the exact-inference machinery of Section 4 extends to the unknown-threshold case. First, $\hat{\tau}$ is *super-efficient*: It converges at rate n rather than $\sqrt{n}$, exactly as in the classical shifted-exponential threshold problem. Second, and more importantly, the exact pivot $2\lambda S \sim \chi^2_{2n}$ of Eq (4.2) does *not* carry over unchanged once τ is replaced by $\hat{\tau}$: That pivot's

derivation relies on the transformed observations being an exact linear shift of an i.i.d. exponential sample, a property that holds when τ is known but is disturbed by the nonlinear (square-root) dependence of $t_i(\hat{\tau})$ on the estimated threshold. In the classical two-parameter exponential this obstacle does not arise, because the transformation there is linear and the memoryless spacings $X_i - X_{(1)}$ reproduce an exact $\chi^2_{2(n-1)}$ pivot; the square-root map underlying the MERD breaks that linear argument. We therefore do not claim finite-sample exactness for $\hat{\lambda}$ once τ (or σ) is unknown, and recommend a parametric bootstrap, resampling from the fitted LS-MERD($\hat{\tau}, \hat{\sigma}, \hat{\lambda}$), for finite-sample confidence statements in that setting; Proposition 1 and Eq (4.7) remain exactly valid whenever the threshold is fixed by external knowledge rather than estimated from the same sample.

In summary, the fixed lower bound of the basic MERD is not an unexaminable limitation: Eq (6.1) gives an explicit, identifiable three-parameter generalization with closed-form distributional functions, its threshold component has a simple closed-form estimator Eq (6.5) that requires no numerical search, and we have characterized precisely which exact-inference guarantees survive this generalization and which require bootstrap or asymptotic approximation instead.

7. Simulation study

7.1. Design and estimating equations

We run a Monte Carlo study to evaluate the estimators. For each $\lambda \in \{0.5, 1.0, 2.0, 5.0\}$ and $n \in \{20, 30, 50, 100, 200, 400\}$, we generate $R = 20{,}000$ independent samples from MERD(λ) via the inverse-CDF method Eq (5.1). For each sample, we compute the MLE Eq (4.4) and two Bayes estimators from Eq (4.8): a vague prior Gamma$(0.001, 0.001)$ and an informative prior Gamma$(4, 4/\lambda)$. Point performance is assessed by

$$\widehat{\text{bias}} = \frac{1}{R}\sum_{r=1}^{R}(\hat{\lambda}^{(r)} - \lambda), \qquad \widehat{\text{MSE}} = \frac{1}{R}\sum_{r=1}^{R}(\hat{\lambda}^{(r)} - \lambda)^2, \tag{7.1}$$

and the 95% Wald, exact χ^2, and Bayesian credible intervals by

$$\widehat{\text{CP}} = \frac{1}{R}\sum_{r=1}^{R}\mathbf{1}\{\lambda \in [L^{(r)}, U^{(r)}]\}, \qquad \widehat{\text{AL}} = \frac{1}{R}\sum_{r=1}^{R}(U^{(r)} - L^{(r)}). \tag{7.2}$$

7.2. Point and interval results

Table 2 shows consistency: As n increases, bias and MSE approach zero with MSE approximately halved by each doubling of n, a pattern that continues cleanly through the newly added $n = 400$ column. While the vague-prior Bayes estimator and the MLE become nearly indistinguishable for large n (as expected, since the vague prior contributes negligibly once the likelihood dominates), the inclusion of an informative prior in Bayes estimation reduces bias and MSE for small n. In Table 3, all three methods approximate the nominal 95% coverage as n increases, and the intervals become shorter. The exact χ^2 interval provides nominal coverage for n greater than or equal to 20, as is required by its definition, and this remains true at $n = 400$.

Table 2. Bias and MSE of the point estimators (20,000 replications each; expanded to include $n = 400$).

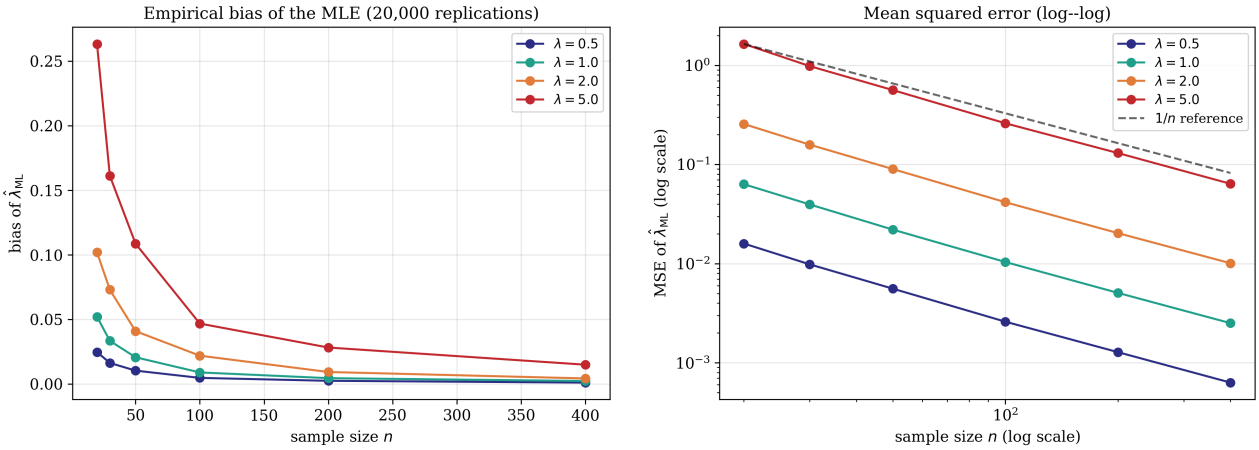
		MLE		Bayes (vague)		Bayes (informative)	
λ	n	Bias	MSE	Bias	MSE	Bias	MSE
0.5	20	0.0248	0.0159	0.0248	0.0159	0.0165	0.0099
0.5	30	0.0164	0.0098	0.0164	0.0098	0.0125	0.0073
0.5	50	0.0104	0.0056	0.0104	0.0056	0.0089	0.0047
0.5	100	0.0048	0.0026	0.0048	0.0026	0.0044	0.0024
0.5	200	0.0025	0.0013	0.0025	0.0013	0.0024	0.0012
0.5	**400**	**0.0012**	**0.0006**	**0.0012**	**0.0006**	**0.0011**	**0.0006**
1.0	20	0.0521	0.0633	0.0521	0.0633	0.0351	0.0394
1.0	30	0.0335	0.0395	0.0335	0.0395	0.0256	0.0292
1.0	50	0.0208	0.0221	0.0208	0.0221	0.0177	0.0185
1.0	100	0.0090	0.0104	0.0090	0.0104	0.0083	0.0096
1.0	200	0.0046	0.0051	0.0046	0.0051	0.0044	0.0049
1.0	**400**	**0.0024**	**0.0025**	**0.0024**	**0.0025**	**0.0023**	**0.0025**
2.0	20	0.1020	0.2559	0.1019	0.2558	0.0683	0.1593
2.0	30	0.0731	0.1584	0.0730	0.1584	0.0565	0.1170
2.0	50	0.0410	0.0900	0.0410	0.0900	0.0349	0.0755
2.0	100	0.0220	0.0417	0.0220	0.0417	0.0204	0.0383
2.0	200	0.0094	0.0203	0.0094	0.0203	0.0090	0.0195
2.0	**400**	**0.0044**	**0.0101**	**0.0044**	**0.0101**	**0.0043**	**0.0099**
5.0	20	0.2634	1.6425	0.2622	1.6402	0.1765	1.0155
5.0	30	0.1612	0.9819	0.1605	0.9810	0.1224	0.7253
5.0	50	0.1087	0.5633	0.1083	0.5630	0.0930	0.4730
5.0	100	0.0469	0.2605	0.0467	0.2605	0.0431	0.2397
5.0	200	0.0283	0.1305	0.0282	0.1304	0.0273	0.1252
5.0	**400**	**0.0150**	**0.0639**	**0.0150**	**0.0639**	**0.0148**	**0.0626**

Table 3. Coverage probability (CP) and average length (AL) of nominal 95% intervals (20,000 replications each; expanded to include $n = 400$ and, for completeness, $\lambda = 1.0$).

		Wald		Exact χ^2		Credible	
λ	n	CP	AL	CP	AL	CP	AL
0.5	20	0.952	0.460	0.949	0.458	0.949	0.458
0.5	30	0.952	0.370	0.950	0.369	0.950	0.369
0.5	50	0.950	0.283	0.949	0.282	0.949	0.282
0.5	100	0.953	0.198	0.952	0.198	0.952	0.198
0.5	200	0.950	0.139	0.949	0.139	0.949	0.139
0.5	**400**	**0.949**	**0.098**	**0.949**	**0.098**	**0.949**	**0.098**
1.0	**20**	**0.955**	**0.922**	**0.952**	**0.918**	**0.952**	**0.918**
1.0	**30**	**0.951**	**0.740**	**0.950**	**0.738**	**0.950**	**0.738**
1.0	**50**	**0.951**	**0.566**	**0.950**	**0.565**	**0.950**	**0.565**
1.0	**100**	**0.951**	**0.396**	**0.951**	**0.395**	**0.951**	**0.395**
1.0	**200**	**0.952**	**0.278**	**0.951**	**0.278**	**0.951**	**0.278**
1.0	**400**	**0.954**	**0.196**	**0.954**	**0.196**	**0.954**	**0.196**
2.0	20	0.954	1.843	0.950	1.834	0.950	1.834
2.0	30	0.953	1.484	0.951	1.479	0.951	1.479
2.0	50	0.950	1.131	0.949	1.129	0.949	1.129
2.0	100	0.952	0.793	0.951	0.792	0.951	0.792
2.0	200	0.953	0.557	0.953	0.557	0.953	0.557
2.0	**400**	**0.951**	**0.393**	**0.951**	**0.393**	**0.951**	**0.393**
5.0	20	0.952	4.614	0.949	4.593	0.949	4.593
5.0	30	0.952	3.694	0.951	3.683	0.951	3.683
5.0	50	0.948	2.832	0.946	2.827	0.946	2.827
5.0	100	0.950	1.978	0.950	1.977	0.950	1.977
5.0	200	0.949	1.394	0.949	1.393	0.949	1.393
5.0	**400**	**0.949**	**0.983**	**0.949**	**0.983**	**0.949**	**0.983**

7.3. Visual diagnostics for the simulation

Figure 8 corroborates the numerical findings. The bias decays to zero for all λ, and the log-log MSE plot is parallel to the $1/n$ reference line, consistent with the $O(n^{-1})$ rate implied by $I(\lambda) = 1/\lambda^2$.



(a) Bias vs. n. **(b)** MSE vs. n (log–log scale).

Figure 8. Monte-Carlo bias Eq (7.1) (left) and MSE (right) of the MLE $\hat{\lambda}_{\mathrm{ML}}$.

Figure 9 shows the empirical coverage fluctuating tightly around 0.95, confirming the Wald interval is well-calibrated even at $n = 20$.

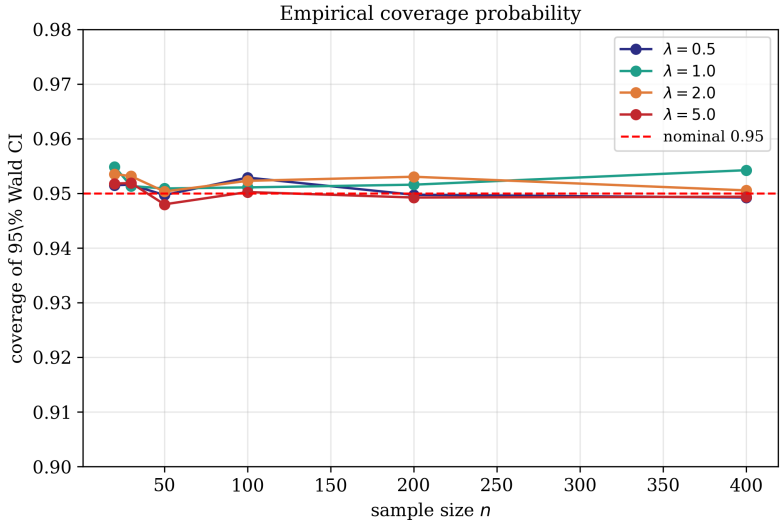


Figure 9. Empirical coverage Eq (7.2) of the 95% Wald interval against n.

Figure 10 exhibits asymptotic normality: The histogram for $n = 30$ shows mild right skewness, and that for $n = 200$ aligns closely with the standard normal distribution, confirming $\sqrt{n}(\hat{\lambda}_{\mathrm{ML}} - \lambda) \xrightarrow{d} \mathcal{N}(0, \lambda^2)$.

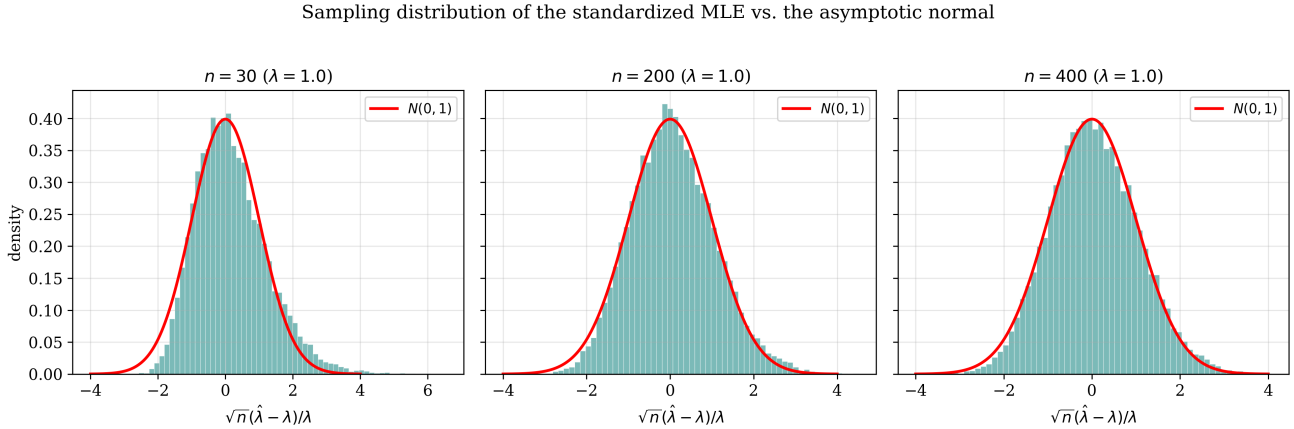


Figure 10. Sampling distribution of the standardized MLE $\sqrt{n}(\hat{\lambda}_{\mathrm{ML}} - \lambda)/\lambda$ against the standard normal density, for $n = 30$ and $n = 200$ at $\lambda = 1$.

8. Robustness, censored inference, and predictive comparison

The simulation study of Section 7 assesses the estimators under the correctly specified MERD model. We now go beyond this in three directions requested in review: a quantitative out-of-sample predictive comparison against the closest one-parameter competitors (Lindley, Akash, Shanker), and a study of robustness under model misspecification and contamination; we additionally show that the exact-inference machinery of Section 4 extends, in closed form, to Type-II censored data.

8.1. Predictive comparison via CRPS

Before a predictive comparison is meaningful, a competitor must first describe the data adequately in sample. Table 4 reports the Kolmogorov–Smirnov statistic and p-value for the MERD, Lindley, Akash, and Shanker fits on both real datasets of Section 9.

Table 4. In-sample Kolmogorov–Smirnov goodness-of-fit screening: MERD vs. Lindley, Akash, and Shanker.

Dataset	Model	KS statistic	KS p-value	In-sample adequacy
AML	MERD	0.116	0.737	not rejected
AML	Lindley	0.309	3.2×10^{-3}	rejected
AML	Akash	0.354	4.4×10^{-4}	rejected
AML	Shanker	0.318	2.2×10^{-3}	rejected
Insurance	MERD	0.058	0.907	not rejected
Insurance	Lindley	0.270	3.1×10^{-6}	rejected
Insurance	Akash	0.348	3.7×10^{-10}	rejected
Insurance	Shanker	0.291	3.8×10^{-7}	rejected

All three one-parameter Lindley-type competitors are rejected at the 5% level on both datasets, several by several orders of magnitude in p-value. Rather than treat this as disqualifying the comparison, we report it explicitly and proceed to a genuine out-of-sample predictive comparison

via 5-fold cross-validation: On each fold, all four models are refit on the training data and scored on the held-out fold by the continuous ranked probability score (CRPS), a strictly proper scoring rule for which lower values indicate better-calibrated, sharper predictive distributions. Table 5 reports the resulting scores.

Table 5. Five-fold cross-validated CRPS (lower is better) for MERD vs. Lindley, Akash, and Shanker.

Dataset	MERD	Lindley	Akash	Shanker
AML	**25.72**	27.57	29.02	27.71
Vehicle Insurance	**8.86**	9.57	10.03	9.64

The MERD attains the lowest (best) CRPS on both datasets, consistent with its in-sample adequacy; the gap tracks the KS results, so the out-of-sample predictive comparison and the in-sample screening agree rather than being in tension. We regard this as informative but not decisive evidence given the small sample sizes ($n = 32, 89$), and note that predictive evaluation against richer, well-fitting multi-parameter competitors (e.g., Weibull, generalized inverse Weibull) remains a valuable further direction.

8.2. Robustness to model misspecification

To assess sensitivity to misspecification of the root distribution, we generalize the construction to $X = (T + 1)^2$ with $T \sim \text{Weibull}(\kappa, 1/\lambda)$, which reduces to the true MERD exactly at $\kappa = 1$. For $\kappa \in \{0.7, 1.0, 1.3, 1.5\}$ and $n = 50$ we generate $R = 5{,}000$ samples with $\lambda = 1$, fit the (misspecified, whenever $\kappa \neq 1$) MERD by MLE, and report bias, MSE, the empirical coverage of the nominal 95% exact χ^2 interval Eq (4.7) (constructed as if the model were correctly specified), and the rejection rate of a KS test of the fitted MERD against the sample. Table 6 summarizes the results.

Table 6. Sensitivity to misspecification of the root distribution ($n = 50$, $R = 5{,}000$, true $\lambda = 1$).

κ (true root shape)	Bias	MSE	Coverage of nominal 95% CI	KS rejection rate
0.7 (misspecified)	−0.180	0.063	0.555	0.438
1.0 (correctly specified)	0.018	0.022	0.947	0.004
1.3 (misspecified)	0.098	0.025	0.970	0.166
1.5 (misspecified)	0.119	0.026	0.975	0.460

At $\kappa = 1$, the exact interval attains its nominal coverage and the KS test rejects at essentially the nominal rate, as expected. Under misspecification, the picture is asymmetric and informative: With $\kappa = 0.7$, the exact interval's coverage collapses to 0.555, a clear failure that is correctly flagged by a KS rejection rate of 44%; with $\kappa = 1.3$ and 1.5, the interval's nominal coverage is (spuriously) preserved or even over-covers, but the KS test still detects the misspecification in 17%–46% of samples. The practical implication is that the exact inference of Section 4 should always be accompanied by the goodness-of-fit diagnostics already reported in Section 9; coverage of the exact interval alone is not a sufficient check of model adequacy.

8.3. Robustness to contamination

Next, we contaminate a fraction $\varepsilon \in \{0, 0.05, 0.10\}$ of each MERD(1) sample ($n = 50$, $R = 5{,}000$) by inflating the contaminated observations' distance from the threshold by a factor of 8 (i.e., $x \mapsto 1+8(x-1)$), a simple gross-outlier scheme, and refit the (uncorrected) MLE. Table 7 reports the resulting bias and MSE.

Table 7. Sensitivity of the MLE to gross-outlier contamination ($n = 50$, $R = 5{,}000$, true $\lambda = 1$).

Contamination fraction ε	Bias	MSE
0.00	0.021	0.022
0.05	−0.098	0.032
0.10	−0.192	0.058

As expected for an MLE with no built-in downweighting mechanism, bias and MSE increase steadily with the contamination fraction. We report this transparently rather than as a strength: the exactness results of Section 4 are properties of the model under correct specification and clean data, not a robustness guarantee, and practitioners working with data suspected to contain gross outliers should pair the MLE with the diagnostic plots of Section 9 or a robust preliminary screen.

8.4. Exact inference under type-II censoring

Reliability data are frequently type-II censored: Only the smallest r of n lifetimes are observed, the remaining $n - r$ being known only to exceed $x_{(r)}$. We show that the exact-inference machinery of Section 4 extends to this setting in closed form.

Proposition 14. *(Exact type-II censored MLE) Let $x_{(1)} < \cdots < x_{(r)}$ be the first r order statistics observed out of n i.i.d.* MERD(λ) *lifetimes, and write $t_{(i)} = \sqrt{x_{(i)}} - 1$. The type-II censored MLE is*

$$\hat{\lambda}_r = \frac{r}{W}, \qquad W := \sum_{i=1}^{r} t_{(i)} + (n - r)\, t_{(r)}, \tag{8.1}$$

and $2\lambda W \sim \chi^2_{2r}$ exactly, giving the exact equal-tailed $100(1-\alpha)\%$ confidence interval

$$\left[\frac{\chi^2_{2r,\alpha/2}}{2W}, \frac{\chi^2_{2r,1-\alpha/2}}{2W}\right].$$

Proof. The type-II censored likelihood is $L(\lambda) \propto \prod_{i=1}^{r} f(x_{(i)}; \lambda)\, S(x_{(r)}; \lambda)^{n-r}$. Since $\ln f(x; \lambda) = \ln \lambda - \ln(2\sqrt{x}) - \lambda(\sqrt{x} - 1)$ and $\ln S(x; \lambda) = -\lambda(\sqrt{x} - 1)$, the λ-dependent part of the log-likelihood is $r \ln \lambda - \lambda W$ with W as in Eq (8.1); maximizing gives $\hat{\lambda}_r = r/W$. Because the root substitution $t_{(i)} = \sqrt{x_{(i)}} - 1$ is monotone increasing, the $t_{(i)}$ are exactly the order statistics of an i.i.d. Exp(λ) sample, so W is precisely the classical total-time-on-test statistic for Type-II censored exponential data, for which $2\lambda W \sim \chi^2_{2r}$ exactly [7]. □

This is a direct extension of the exact pivot Eq (4.2): censoring simply changes n to r degrees of freedom while preserving finite-sample exactness, with no asymptotic approximation at any censoring

level. Table 8 confirms this numerically for censoring proportions $r/n \in \{1.0, 0.9, 0.7, 0.5\}$ ($n = 50$, $R = 5{,}000$, true $\lambda = 1$).

Table 8. Bias, MSE, and exact-interval coverage under type-II censoring ($n = 50$, $R = 5{,}000$, true $\lambda = 1$).

Censoring proportion r/n	r	Bias	MSE	Coverage of exact 95% CI
1.00 (uncensored)	50	0.021	0.022	0.946
0.90	45	0.026	0.026	0.945
0.70	35	0.031	0.033	0.949
0.50	25	0.041	0.048	0.947

As expected, MSE increases as fewer observations are retained, but the exact interval Eq (8.1) attains its nominal 95% coverage at every censoring level considered, down to observing only half the sample, exactness is preserved under censoring, not merely approximately maintained. Figure 11 summarizes the three robustness studies together.

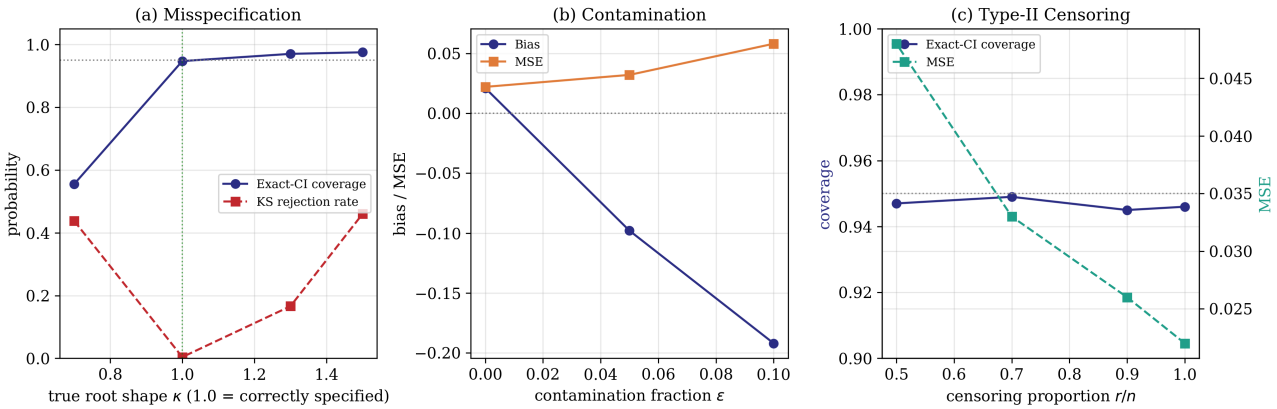


Figure 11. Summary of the robustness study: (a) misspecification of the root distribution's shape, showing exact-interval coverage collapsing away from $\kappa = 1$ while the KS test correctly flags the departure; (b) bias and MSE inflation under gross-outlier contamination; (c) coverage and MSE under type-II censoring, showing exact coverage is maintained even as MSE rises with heavier censoring.

9. Applications to real data

To evaluate empirical behavior, we fit the MERD alongside thirty-three competing distributions, listed in Table 9 to classical lifetime datasets spanning biomedical and actuarial domains. Both datasets analysed in this section are complete (uncensored) observations. All models are estimated by maximum likelihood and compared via the Akaike (AIC), corrected Akaike (CAIC), Bayesian (BIC), and Hannan-Quinn (HQIC) information criteria; the Kolmogorov-Smirnov (KS) statistic and its associated p-value; and the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics with their corresponding p-values.

Table 9. Competing probability distributions, abbreviations, and corresponding references.

Model	Abbreviation	Ref.
Sine Topp–Leone Fréchet distribution	STLFRD	[24]
Odd exponential Lomax distribution	OELD	[25]
Odd generalized exponential Lomax distribution	OGELD	[25]
Power XLindley distribution	PXLD	[26]
Half logistic Zeghdoudi distribution	HLZD	[27]
Power komal distribution	PKD	[28]
Alpha power transformed power Lindley distribution	APTPLD	[29]
Power Lindley distribution	PLD	[30]
Power Zeghdoudi distribution	PZD	[31]
Weibull Lomax distribution	WLD	[25]
Power length-biased new XLindley distribution	PLBNXLD	[32]
Sine exponentiated Weibull exponential distribution	SEWED	[33]
Weibull distribution	WD	[34]
Inverse exponentiated Weibull distribution	IEWD	[35]
Gamma distribution	GD	[36]
Half logistic new-Weibull Pareto distribution	HLNWPD	[37]
Half logistic Weibull distribution	HLWD	[37]
Log-logistic distribution	LOLD	[38]
Alpha power transformed log-logistic distribution	APTLLD	[39]
Sine inverted exponentiated Weibull distribution	SIEWD	[40]
Alpha power inverse Weibull distribution	APIWD	[41]
Inverse power Zeghdoudi distribution	IPZD	[42]
Burr type III distribution	BIIID	[43]
Inverse Burr distribution	IBD	[44]
Inverse power Ishita distribution	IPIshD	[45]
Inverse Weibull distribution	IWD	[41]
Generalized inverse Weibull distribution	GIWD	[46]
Inverse power Gompertz distribution	IPGoD	[47]
Inverse power Lindley distribution	IPLD	[48]
Inverse Kumaraswamy distribution	IKD	[49]
Lomax Distribution	LoD	[50]
Alpha power transformed extended Lindley distribution	APTELD	[51]
Inverse gompertz distribution	IGoD	[52]

9.1. Acute myelogenous leukemia data (biomedical survival)

To evaluate the performance and applicability of the proposed model, a real-life survival dataset concerning patients diagnosed with AML was analyzed. The dataset consists of the survival times (in weeks) of 32 patients suffering from acute myelogenous leukemia.

The observed survival times are 65, 156, 100, 134, 16, 108, 121, 4, 39, 143, 56, 26, 22, 1, 1, 5, 65,

56, 65, 17, 7, 16, 22, 3, 4, 2, 8, 4, 3, 30, 4, 43.

This dataset has become a primary example within the existing literature of survival analysis and lifetime distributions for assessing the adequacy of alternative statistical models. The original data were published in [53], and multiple authors have since analyzed them, including [54].

The summary statistics data is included in Table 10. This study is based on thirty-two (32) individuals, whose survival period ranged between one (1) and one hundred and fifty-six (156) weeks. The mean survival period is reported as 42.06 weeks, while the survival period of the middle value (median) is 22 weeks. This indicates a right-skewed distribution with a heavy upper tail.

Table 10. Descriptive statistics for the acute myelogenous leukemia survival times dataset.

Statistic	n	Min	Q_1	Median	Mean	Q_3	Max
Value	32	1	4	22	42.06	65	156

The substantial disparity between the mean and median values, along with the broad range of data points, indicates right-skew and recommends the application of flexible lifetime distributions for data modeling. Figure 12 presents the corresponding exploratory data analysis.

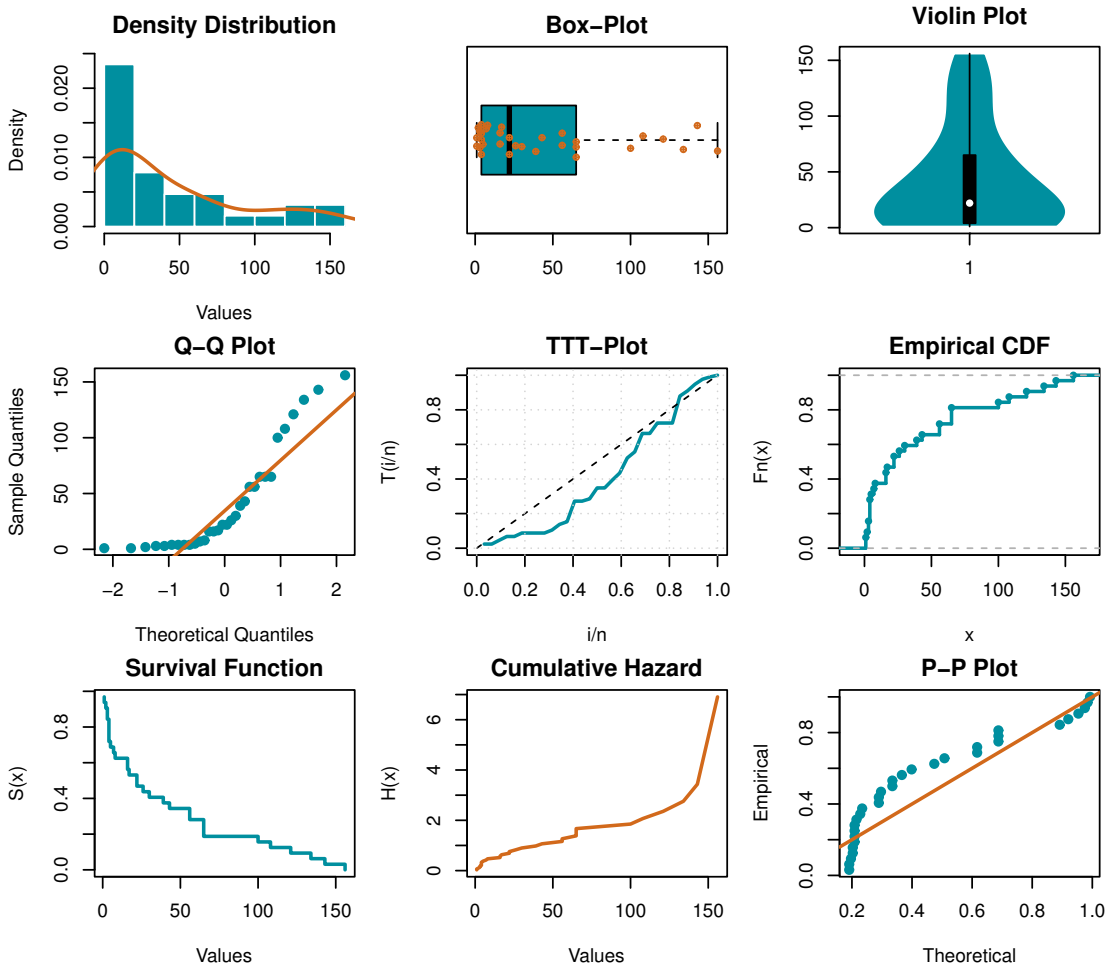


Figure 12. Exploratory data analysis for the acute myeloid leukaemia dataset.

Table 11 reports the fitted parameter estimates and corresponding standard errors for all competing models on the acute myelogenous leukemia survival times dataset.

Table 11. Parameter estimates and corresponding standard errors for all fitted models for the acute myelogenous leukemia survival times dataset.

Model	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\eta}$	SE($\hat{\lambda}$)	SE($\hat{\beta}$)	SE($\hat{\alpha}$)	SE($\hat{\eta}$)
MERD	0.2254				0.0398			
STLFRD	15.1237	0.4395	0.2978		23.2044	0.0954	0.4669	
OELD	1.0523	0.7222	0.0836		2.9283	0.1696	0.2153	
OGELD	0.7966	0.6620	0.0968	1.1350	2.6624	0.3837	0.2387	0.9242
PXLD	0.6268	0.1913			0.0705	0.0530		
HLZD	0.0257	0.2899			0.0068	0.0532		
PKD	0.5986	0.2238			0.0682	0.0605		
APTPLD	0.8009	0.5980	0.2192		1.2804	0.0892	0.1240	
PLD	0.5895	0.2352			0.0699	0.0665		
PZD	0.6088	0.4525			0.1349	0.0558		
WLD	0.2518	0.3512	0.0660	1.7713	0.9432	0.5113	0.1824	1.9327
PLBNXLD	0.4683	0.5524			0.0638	0.1497		
SEWED	0.9691	0.0116	0.6802		0.3198	0.0196	0.3445	
WD	0.7923	36.9841			0.1119	8.7130		
IEWD	0.0899	4561.6336	11.5719		0.0327	13742.7098	2.9750	
GD	0.7058	59.5850			0.1504	17.8402		
HLNWPD	0.4738	0.6753	0.0759		6.2564	0.1015	0.6747	
HLWD	0.6754	0.1257			0.1014	0.0573		
LOLD	19.0457	1.0970			5.4938	0.1564		
APTLLD	0.9980	19.0752	1.0973		2.4637	22.1594	0.1564	
SIEWD	10.2823	0.0982	811.3133		2.4337	0.0338	1990.5680	
APIWD	0.8005	2.8516	8.7261		0.1121	1.6198	17.0459	
IPZD	0.4639	6.6408			0.0605	1.0189		
BIIID	5.7036	0.7553			1.2275	0.0925		
IBD	0.0033	158.4114	140.1253		0.0011	3.7376	54.0633	
IPIshD	0.6919	4.4819			0.0895	0.8748		
IWD	4.2574	0.6848			0.9324	0.0908		
GIWD	0.6849	3.1938	1.5221		0.0908	0.8465	1.6356	
IPGoD	0.6841	0.0092	461.5182		0.0908	0.0102	525.3427	
IPLD	4.9075	0.6749			0.9531	0.0916		
IKD	0.7107	4.4118			0.1178	1.3330		
LoD	91.6520	3.0655			134.5811	3.4372		
APTELD	0.2427	0.0176	0.0007		0.3116	0.0137	0.0144	
IGoD	9.8593	-2.2077			2.3315	0.9737		

Note: Column headers $\hat{\lambda}, \hat{\beta}, \hat{\alpha}, \hat{\eta}$ are generic slots for each model's first, second, third, and fourth parameters respectively, in the order used by the corresponding reference in Table 9; they do not denote a common parameter across models. Blank cells indicate that the parameter is not part of the corresponding model (i.e., that model has fewer than four parameters).

Tables 12 and 13 detail the results of fitting models to the acute myelogenous leukemia survival data. From these, we see that the proposed MERD provides the best overall fit. In particular, among the thirty-three models considered, MERD provides the smallest AIC, BIC, CAIC, and HQIC values and thus the best fit and most parsimonious model.

Table 12. Log-likelihood and information criteria for all fitted models for the acute myelogenous leukemia survival times dataset.

Model	LogLik	AIC	BIC	CAIC	HQIC
MERD	148.014	298.028	299.494	298.161	298.514
STLFRD	151.175	308.351	312.748	309.208	309.808
OELD	149.882	305.765	310.162	306.622	307.222
OGELD	149.869	307.739	313.602	309.220	309.682
PXLD	150.565	305.130	308.061	305.543	306.101
HLZD	150.241	304.483	307.414	304.896	305.454
PKD	150.567	305.134	308.066	305.548	306.106
APTPLD	150.544	307.087	311.485	307.945	308.545
PLD	150.553	305.106	308.037	305.520	306.078
PZD	150.287	304.574	307.505	304.988	305.546
WLD	149.689	307.379	313.242	308.860	309.322
PLBNXLD	150.264	304.528	307.460	304.942	305.500
SEWED	150.275	306.551	310.948	307.408	308.009
WD	150.151	304.302	307.233	304.716	305.274
IEWD	150.096	306.191	310.589	307.049	307.649
GD	150.178	304.357	307.288	304.771	305.328
HLNWPD	150.537	307.074	311.471	307.931	308.531
HLWD	150.537	305.074	308.005	305.487	306.045
LOLD	152.123	308.247	311.178	308.660	309.218
APTLLD	152.123	310.247	314.644	311.104	311.704
SIEWD	150.185	306.370	310.767	307.227	307.828
APIWD	152.505	311.009	315.406	311.866	312.467
IPZD	152.366	308.732	311.664	309.146	309.704
BIIID	152.857	309.714	312.645	310.127	310.685
IBD	149.351	304.703	309.100	305.560	306.161
IPIshD	153.172	310.344	313.276	310.758	311.316
IWD	153.232	310.463	313.395	310.877	311.435
GIWD	153.232	312.463	316.861	313.321	313.921
IPGoD	153.237	312.475	316.872	313.332	313.933
IPLD	153.314	310.627	313.559	311.041	311.599
IKD	153.243	310.486	313.417	310.899	311.457
LoD	151.166	306.331	309.263	306.745	307.303
APTELD	150.750	307.501	311.898	308.358	308.958
IGoD	154.734	313.468	316.400	313.882	314.440

Table 13. Goodness-of-fit statistics for all fitted models for the acute myelogenous leukemia survival times dataset.

Model	KS	KS-p	W*	W-p	A*	A-p
MERD	0.1164	0.7787	0.0758	0.7513	0.4524	0.7250
STLFRD	0.1180	0.7643	0.1129	0.5278	0.7043	0.5541
OELD	0.1214	0.7330	0.0767	0.7148	0.5239	0.7213
OGELD	0.1225	0.7234	0.0796	0.6977	0.5368	0.7083
PXLD	0.1231	0.7173	0.0801	0.6948	0.5610	0.6843
HLZD	0.1233	0.7151	0.0818	0.6846	0.5729	0.6726
PKD	0.1248	0.7009	0.0820	0.6838	0.5697	0.6758
APTPLD	0.1249	0.7004	0.0831	0.6774	0.5741	0.6715
PLD	0.1251	0.6981	0.0824	0.6812	0.5713	0.6742
PZD	0.1259	0.6909	0.0863	0.6593	0.5804	0.6653
WLD	0.1261	0.6886	0.0841	0.6720	0.5599	0.6854
PLBNXLD	0.1266	0.6840	0.0860	0.6612	0.5797	0.6660
SEWED	0.1266	0.6837	0.0849	0.6673	0.5805	0.6653
WD	0.1273	0.6778	0.0823	0.6818	0.5657	0.6797
IEWD	0.1277	0.6736	0.0882	0.6489	0.5884	0.6577
GD	0.1277	0.6734	0.0837	0.6737	0.5756	0.6700
HLNWPD	0.1282	0.6691	0.0851	0.6662	0.5839	0.6621
HLWD	0.1282	0.6691	0.0851	0.6661	0.5841	0.6618
LOLD	0.1283	0.6676	0.1006	0.5848	0.6804	0.5742
APTLLD	0.1285	0.6663	0.1006	0.5847	0.6808	0.5738
SIEWD	0.1291	0.6607	0.0868	0.6569	0.5829	0.6629
APIWD	0.1297	0.6548	0.1244	0.4801	0.7973	0.4819
IPZD	0.1397	0.5598	0.1318	0.4521	0.8143	0.4698
BIIID	0.1405	0.5522	0.1328	0.4484	0.8259	0.4617
IBD	0.1478	0.4866	0.1723	0.3293	0.9527	0.3823
IPIshD	0.1527	0.4447	0.1426	0.4145	0.8874	0.4212
IWD	0.1535	0.4378	0.1436	0.4115	0.8963	0.4156
GIWD	0.1535	0.4377	0.1436	0.4114	0.8966	0.4155
IPGoD	0.1539	0.4348	0.1440	0.4100	0.8985	0.4143
IPLD	0.1548	0.4268	0.1452	0.4061	0.9085	0.4082
IKD	0.1564	0.4142	0.1569	0.3705	0.9374	0.3911
LoD	0.1622	0.3687	0.1425	0.4151	0.9780	0.3683
APTELD	0.1691	0.3194	0.1601	0.3617	1.1191	0.3000
IGoD	0.1883	0.2063	0.3422	0.1029	1.9268	0.1012

Measures of goodness-of-fit in Table 13 also support the results already described. MERD provides the smallest KS and the largest KS-p values. This indicates that in this case the null hypothesis, which states that the model fits the data sufficiently, is not rejected. Similarly, for MERD, W^* and A^* take on the smallest values, while W-p and A-p take on the largest values. These results also support that the MERD model is the most appropriate of those considered for describing the central and tail behaviors of the acute myelogenous leukemia survival data.

Figure 13 provides a detailed graphical evaluation of the fitted MERD model. As for the histogram, when the fitted density curve is added, it is evident that the MERD density is aligned with the empirical data distribution, especially in the region that is highly skewed to the right. The empirical and fitted CDFs almost completely overlap, and therefore, the fitted model is confirmed to be adequate. Also, in the Q–Q and P–P plots, points that are located close to the 45-degree line of reference imply that there is a high level of agreement between the observed and the theoretical quantiles and probabilities, respectively. The survival and cumulative hazard plots also indicate a smooth monotone pattern that is in line with the leukemia survival data. From both the numerical criteria and the graphical diagnostics, it is clear that the MERD model fits the AML dataset well.

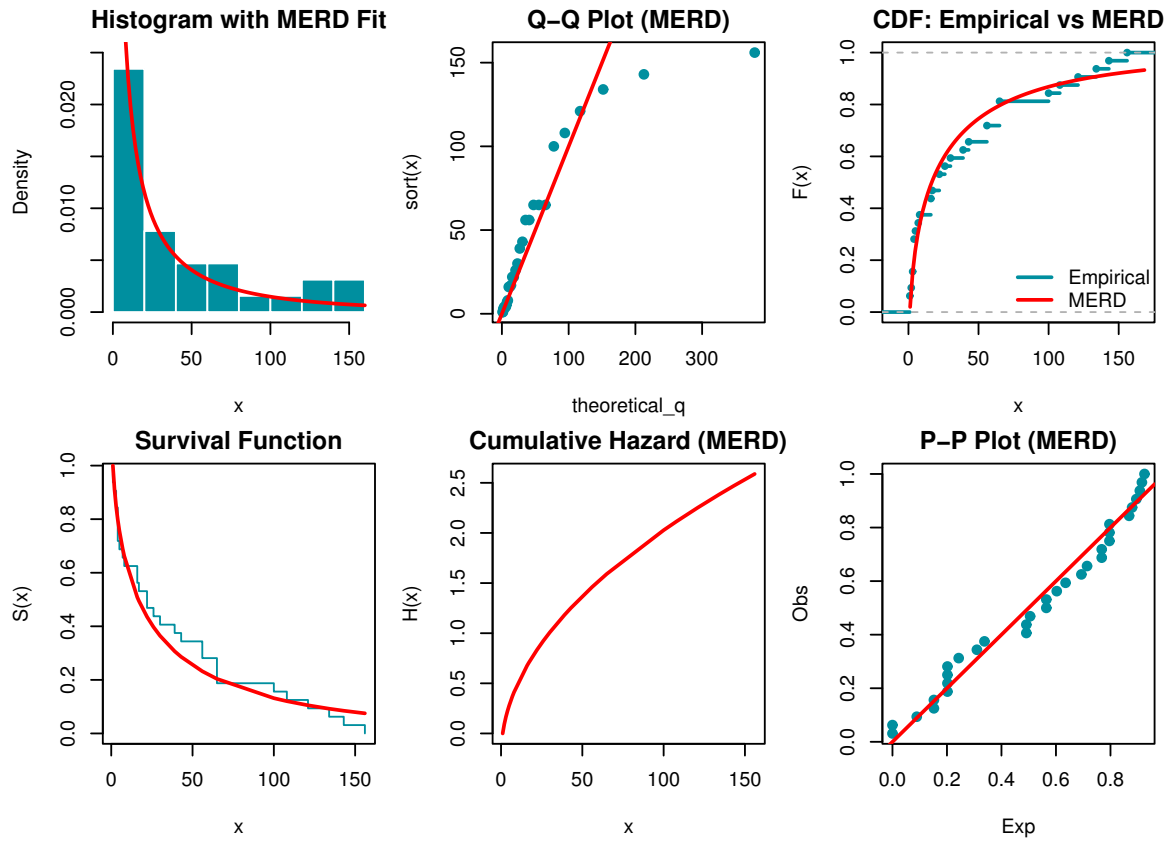


Figure 13. The diagnostic plots for the MERD analysis of the acute myelogenous leukemia survival data include a histogram with an overlaid fitted density, a Q-Q plot, both empirical and fitted CDFs, the survival and cumulative hazard functions, as well as a P-P plot.

9.2. Vehicle insurance complaints data (actuarial/insurance)

The second real dataset consists of the complaint ratios of automobile insurance companies doing business in New York State, as published by the New York State Department of Financial Services for the period beginning in 2009 [55]. The data represent the number of consumer complaints upheld against each insurer as a proportion of their total business over a two-year period. This dataset has been widely used in the actuarial and insurance literature for evaluating the performance of continuous probability distributions; the data were analyzed by Khan et al. [56].

204.173, 84.769, 65.335, 62.505, 46.735, 43.693, 35.072, 32.511, 30.867, 30.397,
29.776, 26.249, 23.911, 23.263, 22.796, 20.956, 19.667, 18.093, 17.419, 16.934,
16.023, 15.572, 15.374, 14.056, 14.044, 14.028, 12.493, 11.733, 11.331, 11.177,
11.008, 10.772, 9.918, 9.881, 9.421, 9.399, 9.336, 9.141, 8.919, 8.253,
8.193, 7.647, 7.589, 7.493, 7.094, 6.635, 6.365, 6.259, 6.217, 5.416,
5.323, 4.405, 4.190, 4.109, 3.764, 3.613, 3.384, 3.106, 3.019, 2.979,
2.890, 2.843, 2.841, 2.770, 2.749, 2.643, 2.616, 2.550, 2.546, 2.433,
2.425, 2.403, 2.392, 2.087, 1.945, 1.833, 1.805, 1.798, 1.741, 1.719,
1.618, 1.584, 1.416, 1.407, 1.364, 1.358, 1.238, 1.199, 1.048.

The observations represent upheld complaints against vehicle insurance companies and exhibit substantial positive skewness, with a few exceptionally large values compared with the majority of observations. Therefore, the data provide an appropriate benchmark for assessing the flexibility of competing statistical models in fitting highly skewed insurance data.

The sample consists of $n = 89$ observations. Summary statistics of the data are reported in Table 14.

Table 14. Descriptive statistics for the vehicle insurance complaints data.

Statistic	n	Min	Q_1	Median	Mean	Q_3	Max
Value	89	1.048	2.616	7.094	14.079	15.374	204.173

It can be observed that the mean (14.079) is considerably larger than the median (7.094), indicating a strong right-skewed distribution. Moreover, the large gap between the maximum observation (204.173) and the third quartile (15.374) suggests the presence of extreme observations and a heavy right tail. Figure 14 presents the corresponding exploratory data analysis.

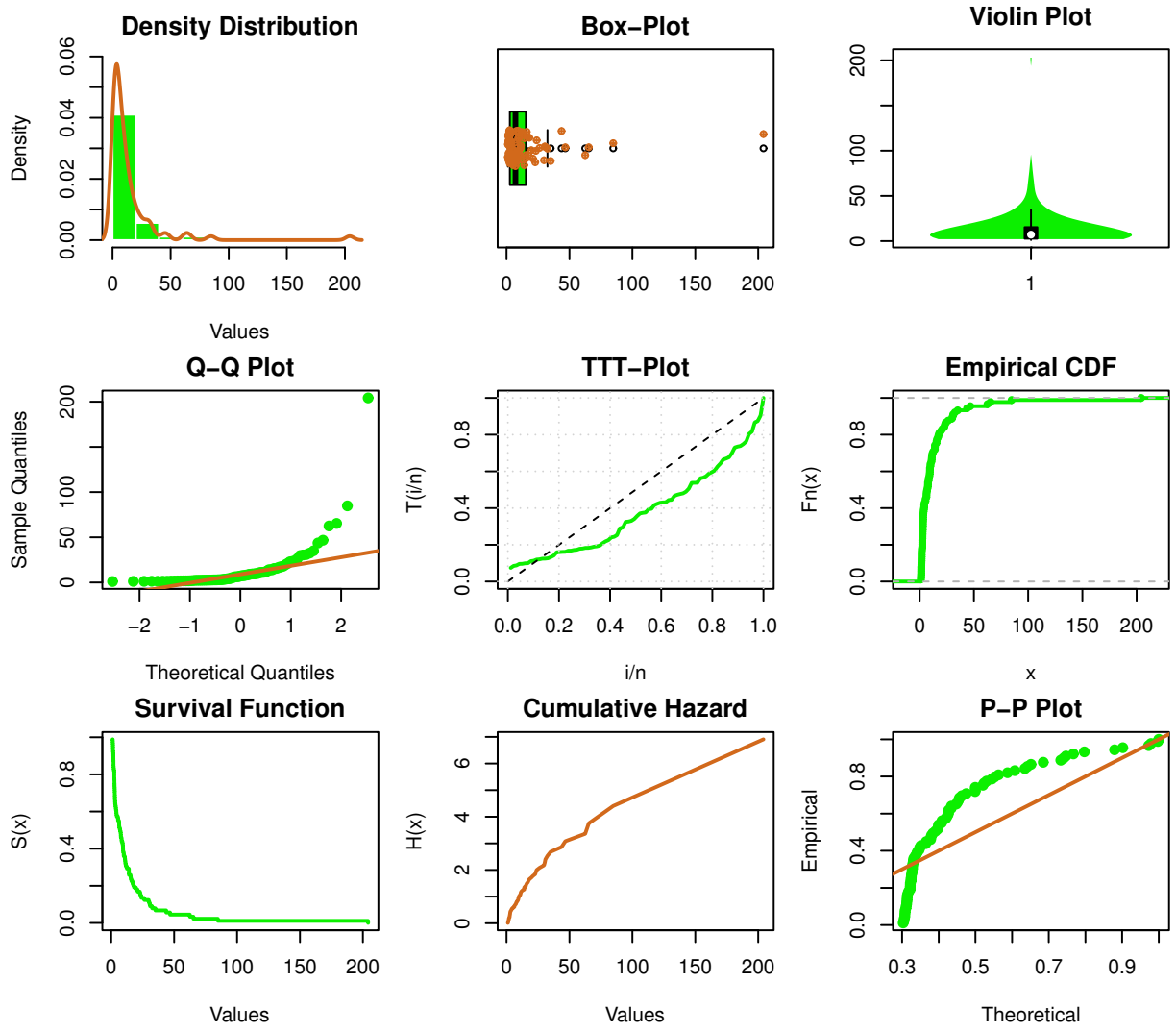


Figure 14. Exploratory data analysis for the vehicle insurance complaints dataset.

Tables 15–17 summarize the estimation and goodness-of-fit results for the vehicle insurance complaints data. The proposed MERD again attains the best fit among the competing models on this dataset.

From Table 16, MERD attains the lowest information criteria values among all fitted distributions, with AIC, BIC, CAIC, and HQIC. Compared to the nearest competitors, these values are substantially smaller, which suggests that MERD presents the most parsimonious and best-fitting model for these extremely skewed insurance data.

The goodness-of-fit statistics in Table 17 support the same conclusion. The model achieves the smallest KS and the largest KS-p among all models, indicating close agreement between the fitted distribution and the empirical data. Likewise, MERD produces the minimum values of W^* and A^*, together with the largest corresponding p-values, on this dataset.

Table 15. Parameter estimates and corresponding standard errors for all fitted models for the vehicle insurance data.

Model	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\eta}$	SE($\hat{\lambda}$)	SE($\hat{\beta}$)	SE($\hat{\alpha}$)	SE($\hat{\eta}$)
MERD	0.4742				0.0503			
IGoD	3.6074	0.4403			0.6050	0.5212		
IPIshD	1.0533	4.4143			0.0859	0.5283		
IPGoD	0.9699	0.5777	5.8477		0.1375	0.8112	10.1079	
IPZD	0.6810	6.3491			0.0555	0.5835		
IKD	1.1813	6.4182			0.1174	1.3288		
IPLD	4.8279	1.0282			0.5774	0.0879		
IBD	833.4050	0.0063	1.0431		575.8355	0.0028	0.0872	
IWD	4.1794	1.0425			0.5643	0.0872		
IEWD	0.8585	1.4083	4.2805		0.3899	1.1574	0.5918	
WD	0.8227	12.3018			0.0609	1.6841		
GD	0.8152	17.2681			0.1056	3.0184		
PZD	0.8170	0.5245			0.0897	0.0365		
PXLD	0.7007	0.2935			0.0445	0.0397		
PLD	0.6629	0.3565			0.0438	0.0488		
PKD	0.6770	0.3327			0.0429	0.0432		
PLBNXLD	0.5278	0.7991			0.0394	0.1023		
APIWD	1.1123	3.7475	2.0246		0.1255	0.8949	2.0758	
GIWD	1.0422	2.7149	1.5121		0.0872	26.0996	13.9489	
SEWED	0.0952	4.6349	32915.9403		0.0089	0.2759	17830.0494	
BIIID	5.4548	1.1340			0.7137	0.0876		
SIEWD	3.9163	1.0378	0.5388		0.5819	0.5364	0.4515	
STLFRD	0.1390	0.3769	239.5722		0.1244	0.0315	405.1695	
HLZD	0.0796	0.3465			0.0121	0.0403		
HLWD	0.6932	0.2650			0.0547	0.0513		
APTLLD	1.0022	6.5011	1.4862		1.7307	3.8654	0.1282	
LOLD	6.5058	1.4860			0.8228	0.1281		
APTELD	0.0095	0.0218			0.0143	0.0098	0.0062	
APTPLD	0.0201	0.7338	0.1444		0.0296	0.0518	0.0340	
LoD	22.9205	2.6471			10.8615	0.9528		
WLD	1.0499	0.0800	25.4428	2.0734	0.9894	2.8108	0.0317	0.4343
OGELD	0.0307	0.1005	6.1024	48.7074	0.0595	0.0904	8.6724	71.9422
OELD	22.6248	0.0091	285.4940		10.7809	0.0061	216.6946	
HLNWPD	0.7241	0.6932	0.2119		23.3259	0.0548	4.7289	

Note: Column headers $\hat{\lambda}, \hat{\beta}, \hat{\alpha}, \hat{\eta}$ are generic slots for each model's first, second, third, and fourth parameters respectively, in the order used by the corresponding reference in Table 9; they do not denote a common parameter across models. Blank cells indicate that the parameter is not part of the corresponding model (i.e., that model has fewer than four parameters).

Table 16. Log-likelihood and information criteria for all fitted models for the vehicle insurance data.

Model	LogLik	AIC	BIC	CAIC	HQIC
MERD	302.437	606.874	609.363	606.920	607.878
IGoD	306.562	617.124	622.101	617.264	619.130
IPIshD	306.836	617.672	622.649	617.811	619.678
IPGoD	306.538	619.076	626.542	619.358	622.085
IPZD	306.606	617.211	622.188	617.351	619.217
IKD	307.509	619.018	623.995	619.158	621.024
IPLD	306.782	617.564	622.542	617.704	619.571
IBD	306.794	619.588	627.054	619.870	622.597
IWD	306.790	617.579	622.557	617.719	619.585
IEWD	306.699	619.398	626.864	619.681	622.408
WD	320.412	644.823	649.800	644.963	646.829
GD	323.065	650.130	655.107	650.270	652.136
PZD	316.105	636.209	641.187	636.349	638.216
PXLD	321.317	646.635	651.612	646.774	648.641
PLD	320.190	644.380	649.357	644.519	646.386
PKD	320.649	645.299	650.276	645.438	647.305
PLBNXLD	316.235	636.470	641.447	636.609	638.476
APIWD	306.556	619.112	626.578	619.394	622.121
GIWD	306.790	619.579	627.045	619.862	622.589
SEWED	306.476	618.953	626.419	619.235	621.962
BIIID	307.529	619.058	624.036	619.198	621.065
SIEWD	306.626	619.252	626.718	619.535	622.262
STLFRD	306.988	619.976	627.442	620.258	622.985
HLZD	323.633	651.266	656.244	651.406	653.273
HLWD	321.914	647.827	652.805	647.967	649.834
APTLLD	311.129	628.258	635.724	628.541	631.268
LOLD	311.129	626.258	631.236	626.398	628.265
APTELD	315.141	636.282	643.748	636.564	639.291
APTPLD	315.990	637.980	645.446	638.263	640.990
LoD	314.736	633.471	638.448	633.611	635.477
WLD	310.082	628.165	638.119	628.641	632.177
OGELD	306.445	620.890	630.845	621.366	624.903
OELD	314.740	635.480	642.946	635.762	638.489
HLNWPD	321.914	649.827	657.293	650.110	652.837

Table 17. Goodness-of-fit statistics for all fitted models for the vehicle insurance data.

Model	KS	KS-p	W*	W-p	A*	A-p
MERD	0.0580	0.9082	0.0547	0.8487	0.3584	0.8885
IGoD	0.1034	0.2774	0.1959	0.2758	1.0087	0.3525
IPIshD	0.0924	0.4079	0.1766	0.3182	0.9331	0.3941
IPGoD	0.1014	0.2989	0.1832	0.3028	0.9579	0.3799
IPZD	0.0956	0.3671	0.1602	0.3604	0.8612	0.4386
IKD	0.0979	0.3389	0.1744	0.3236	0.9471	0.3860
IPLD	0.0952	0.3713	0.1789	0.3129	0.9409	0.3896
IBD	0.0936	0.3926	0.1771	0.3169	0.9340	0.3936
IWD	0.0937	0.3915	0.1772	0.3168	0.9341	0.3935
IEWD	0.0920	0.4134	0.1660	0.3447	0.8862	0.4225
WD	0.1257	0.1100	0.2644	0.1709	1.9732	0.0951
GD	0.1157	0.1704	0.4356	0.0581	2.6477	0.0416
PZD	0.1010	0.3035	0.1911	0.2856	1.4277	0.1947
PXLD	0.1294	0.0925	0.2721	0.1624	2.0303	0.0885
PLD	0.1216	0.1325	0.2521	0.1857	1.8757	0.1077
PKD	0.1234	0.1224	0.2647	0.1705	1.9477	0.0983
PLBNXLD	0.1018	0.2949	0.1851	0.2987	1.4184	0.1972
APIWD	0.0901	0.4400	0.1678	0.3400	0.8963	0.4162
GIWD	0.0936	0.3922	0.1769	0.3175	0.9326	0.3944
SEWED	0.0953	0.3707	0.1583	0.3658	0.8477	0.4475
BIIID	0.0956	0.3669	0.1767	0.3180	0.9557	0.3811
SIEWD	0.0895	0.4475	0.1686	0.3379	0.8988	0.4147
STLFRD	0.1045	0.2668	0.1510	0.3871	0.8442	0.4499
HLZD	0.1235	0.1215	0.5014	0.0393	2.9299	0.0298
HLWD	0.1379	0.0613	0.2654	0.1698	2.0631	0.0850
APTLLD	0.1096	0.2188	0.1511	0.3867	0.9909	0.3618
LOLD	0.1096	0.2191	0.1510	0.3870	0.9905	0.3621
APTELD	0.1030	0.2822	0.1505	0.3885	1.2464	0.2505
APTPLD	0.1026	0.2856	0.1634	0.3516	1.3023	0.2316
LoD	0.1150	0.1754	0.1382	0.4284	1.2555	0.2473
WLD	0.1048	0.2638	0.1166	0.5105	0.8417	0.4516
OGELD	0.0955	0.3682	0.1563	0.3715	0.8384	0.4538
OELD	0.1150	0.1759	0.1383	0.4279	1.2559	0.2471
HLNWPD	0.1379	0.0614	0.2655	0.1697	2.0634	0.0849

The maximum likelihood estimate of the MERD parameter is $\hat{\lambda} = 0.4742$ with a relatively small standard error of 0.0503, suggesting stable parameter estimation despite the presence of extreme observations in the data

The diagnostics presented in Figure 15 highlight the robust nature of MERD. First, the fitted density curve is able to accurately replicate the empirical histogram and is able to capture the heaviness of the

right tail due to the presence of large complaint values. Next, in the Q-Q plot, the empirical quantiles and theoretical quantiles align with each other, with the exception of slight deviations observed due to the extreme upper tail in the quantiles, which, as in many insurance datasets, is expected due to the nature of the data being significantly skewed.

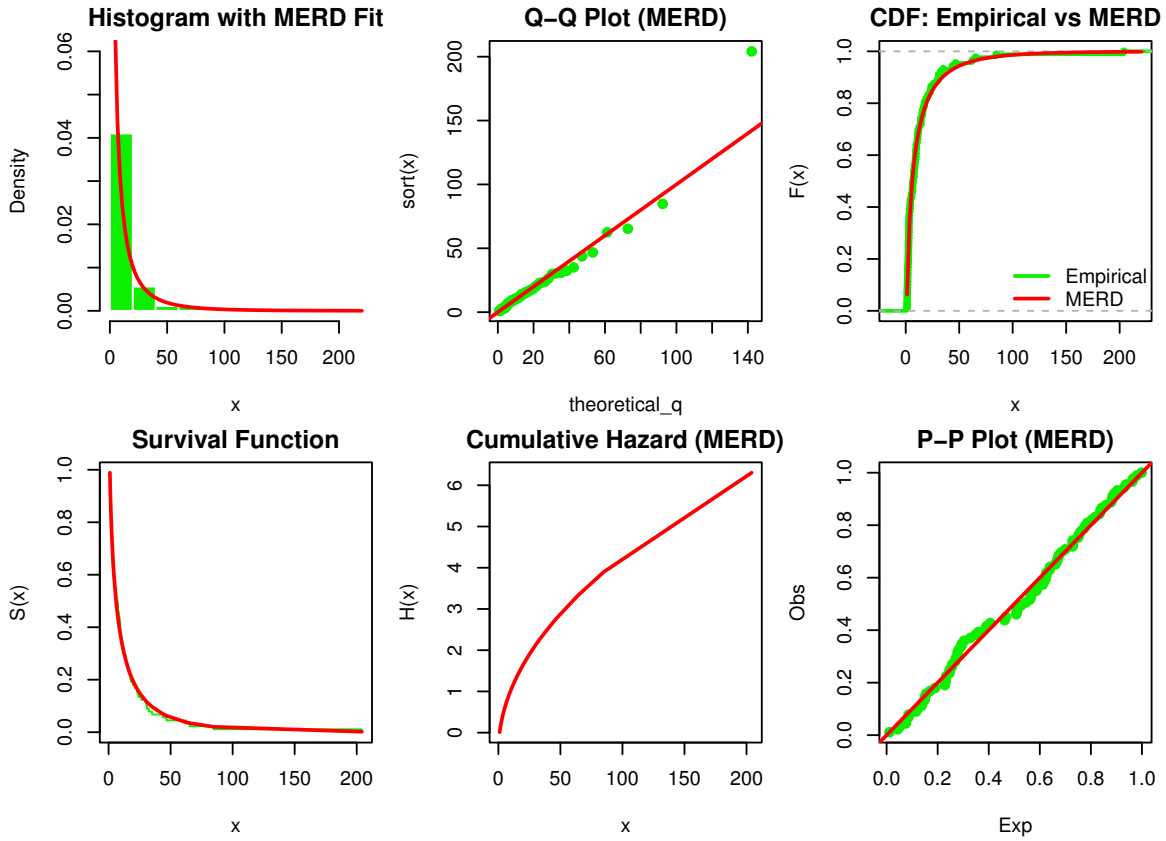


Figure 15. Diagnostic plots for the MERD fitted to the vehicle insurance complaints data, including the fitted density over the histogram, Q–Q plot, empirical and fitted cumulative distribution functions, survival function, cumulative hazard function, and P–P plot.

There is close overlap between the empirical and fitted CDFs across the entire range of data, indicating good agreement of fits. The survival function shows that MERD fits well to the slow tail upper decay. The cumulative hazard function shows a pattern of smooth and consistent increase within the framework of the insurance risk disfunction. In the P–P plot, the data points are distributed closely to the 45 degree line, indicating agreement of observed and fitted probabilities.

These graphical diagnostics, together with the favorable numerical goodness-of-fit measures, indicate that MERD provides a more accurate and flexible representation of the vehicle insurance complaints data than the alternative models considered in this study.

Summary of empirical findings Overall, the empirical analyses based on real datasets spanning the biomedical (Acute Myelogenous Leukemia survival times) and actuarial (vehicle insurance complaints) domains demonstrate the practical usefulness and flexibility of the proposed MERD. Across both datasets, the MERD consistently achieves the lowest values of all information criteria

(AIC, BIC, CAIC, HQIC) and the smallest goodness-of-fit statistics (KS, W*, A*) with the highest corresponding p-values among the 33 competing models considered. These results provide strong empirical evidence for the MERD's competitive fitting performance. However, we emphasize that these findings are dataset-specific and should not be interpreted as a claim of universal superiority over all multi-parameter models in every context. The primary contribution of the MERD remains its unique combination of competitive performance with analytical tractability and exact finite-sample inference, a feature that is particularly valuable in small-sample settings where asymptotic approximations may be unreliable.

Theoretical interpretation of the empirical fit. The strong performance of the MERD on these datasets can be understood through its analytical properties. The density function exhibits a reverse-J shape with its mode at the lower bound $x = 1$, which naturally captures the concentration of observations near the minimum value observed in both datasets. The strictly decreasing hazard function $h(x) = \lambda/(2\sqrt{x})$ reflects the DFR characteristic of early-failure and heavily right-skewed data. The quantile function $Q(p) = (1 - \lambda^{-1}\ln(1-p))^2$ produces a heavy right tail, accommodating the extended survival times in the AML data and the extreme complaint ratios in the insurance data. These theoretical features collectively explain why the MERD provides an excellent fit to the empirical distributions without requiring additional parameters.

10. Conclusions

We have given a complete, self-contained analytic treatment of the MERD, a one-parameter lifetime model on $[1, \infty)$ obtained as the square of a unit-shifted exponential variate. Because the model is a monotone transformation of the exponential, its distribution, reliability, entropy, moment, and shape functions are all elementary, and a single root substitution reduces every computation to a standard exponential integral. The same structure delivers what we regard as the model's most useful practical feature: The transformed observations are exactly exponential, so the chi-square pivot $2\lambda S \sim \chi^2_{2n}$ provides *exact* confidence intervals for the rate parameter at any sample size, alongside closed-form maximum-likelihood and conjugate Bayesian estimators.

Empirical validation across domains. Consistent with the summary given in Section 9, the MERD fit both the AML survival times and the insurance complaint ratios better than the thirty-three competing models considered, combining this empirical performance with its minimal, one-parameter structure and exact-inference machinery.

A Monte Carlo simulation substantiated the claims regarding the consistency and asymptotic normality of the estimators, as well as their interval coverage being well calibrated. With respect to the primary datasets, the MERD performed comparably to one- and two-parameter models with respect to the parsimony-penalized information criteria. We have rather carefully stated this evidence: richer two-parameter models such as the Weibull can attain higher omnibus goodness-of-fit statistics on some datasets, and in such cases the MERD's principal value lies in its exact inference machinery and parsimony rather than in superior raw fit. The model should not be seen as a strong competitor to all other models, but rather a preferable, exactly-inferable model for data that is bounded below and that has early failures with a sharp concentration at the lower boundary.

The primary constraints arise from the structure of the model. The single-parameter model constrains both the lower bound and the scale. Additionally, the moderate-sized datasets ($n = 32$ and

$n = 89$) should be taken into account when considering the reported information-criteria gaps. Section 6 addresses this directly: we derived an identifiable three-parameter location-scale extension, showed that its threshold is estimated by the sample minimum with no numerical search required, and established that this estimator is super-efficient (converging at rate n) while the finite-sample exactness of the λ-pivot does not survive threshold estimation and should instead be handled by parametric bootstrap. This reconnects the model to the threshold-model framework, removes the fixed-bound restriction where the threshold is unknown, and identifies exactly which inferential guarantees are retained and which are lost in doing so. Section 8 further shows that a predictive (CRPS) comparison against the closest one-parameter competitors favors the MERD, that misspecification is reliably flagged by the goodness-of-fit diagnostics already used throughout the paper even when the exact interval's nominal coverage is not itself a reliable warning sign, that gross-outlier contamination degrades the MLE as expected for an uncorrected likelihood method, and that the exact chi-square pivot extends without modification to type-II censored data (Proposition 14). A remaining direction for future work is a broader robustness study against richer contamination and censoring mechanisms and against additional multi-parameter competitors, together with validation on larger datasets.

Author contributions

Khudhayr A. Rashedi: conceptualization, methodology, formal analysis, writing – original draft, writing – review & editing, project administration; L. S. Diab: methodology, formal analysis, software, validation, funding acquisition, writing – review & editing; Abdullah Alenezy: investigation, data curation, software, validation, writing – review & editing; Ghareeb A. Marei: conceptualization, methodology, formal analysis, writing – original draft, writing – review & editing, supervision. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

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

Fundings

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2601).

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. D. V. Lindley, Fiducial distributions and Bayes' theorem, *J. R. Stat. Soc. Ser. B*, **20** (1958), 102–107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x

2. M. E. Ghitany, B. Atieh, S. Nadarajah, Lindley distribution and its application, *Math. Comput. Simul.*, **78** (2008), 493–506. https://doi.org/10.1016/j.matcom.2007.06.007
3. R. Shanker, Akash distribution and its applications, *Int. J. Probab. Stat.*, **4** (2015), 65–75. https://doi.org/10.5923/j.ijps.20150403.01
4. R. Shanker, Shanker distribution and its applications, *Int. J. Probab. Stat.*, **5** (2015), 338–348. https://doi.org/10.5923/j.statistics.20150506.08
5. R. Shanker, Sujatha distribution and its applications, *Stat. Transition*, **17** (2016), 391–410. https://doi.org/10.21307/stattrans-2016-029
6. F. Proschan, Theoretical explanation of observed decreasing failure rate, *Technometrics*, **5** (1963), 375–383. https://doi.org/10.1080/00401706.1963.10490105
7. J. F. Lawless, *Statistical models and methods for lifetime data*, 2 Eds., John Wiley & Sons, Inc., 2003. https://doi.org/10.1002/9781118033005
8. M. Imran, M. A. F. Elbarkawy, F. Jamal, G. A. Marei, An extended Lindley model with applications to entropy-based loss and actuarial risk measures, *Comput. J. Math. Stat. Sci.*, **5** (2026), 1–33.
9. K. Adamidis, S. Loukas, A lifetime distribution with decreasing failure rate, *Stat. Probab. Lett.*, **39** (1998), 35–42. https://doi.org/10.1016/s0167-7152(98)00012-1
10. K. Adamidis, T. Dimitrakopoulou, S. Loukas, On an extension of the exponential-geometric distribution, *Stat. Probab. Lett.*, **73** (2005), 259–269. https://doi.org/10.1016/j.spl.2005.03.013
11. C. Kuş, A new lifetime distribution, *Comput. Stat. Data Anal.*, **51** (2007), 4497–4509. https://doi.org/10.1016/j.csda.2006.07.017
12. 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
13. 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
14. W. Barreto-Souza, H. S. Bakouch, A new lifetime model with decreasing failure rate, *Statistics*, **47** (2013), 465–476. https://doi.org/10.1080/02331888.2011.595489
15. R. D. Gupta, D. Kundu, Generalized exponential distributions, *Aust. N. Z. J. Stat.*, **41** (1999), 173–188.
16. A. S. Hassan, R. E. Mohamed, M. Elgarhy, A. Fayomi, Alpha power transformed extended exponential distribution: properties and applications, *J. Nonlinear Sci. Appl.*, **12** (2019), 239–251. https://doi.org/10.22436/jnsa.012.04.05
17. D. Kumar, U. Singh, S. K. Singh, A method of proposing new distribution and its application to Bladder cancer patients data, *J. Stat. Appl. Probab. Lett.*, **2** (2015), 235–245. https://doi.org/10.12785/jsapl/020306
18. B. Thomas, V. M. Chacko, Power generalized DUS transformation of the exponential distribution, *ArXiv*, 2021. https://doi.org/10.48550/arXiv.2111.14627
19. G. A. Marei, H. M. Aljohani, F. M. Alghamdi, M. O. Mohamed, On estimation of stress-strength model for exponential flexible Weibull extension distribution based on K-records, *Contemp. Math.*, **6** (2025), 2685–2697. https://doi.org/10.37256/cm.6220256450

20. B. Elkalzah, M. O. Mohamed, K. Elsharkawy, E. S. Osman, A. Aldukeel, G. A. Marei, Classical and Bayesian estimation of stress-strength reliability under the discrete Alpha-power Weibull distribution with incomplete and record data: application to high-voltage capacitors, *AIMS Math.*, **11** (2026), 6374–6399. https://doi.org/10.3934/math.2026263

21. A. H. Alenezy, A. Ben Ghorbal, K. A. Rashedi, G. A. Marei, Bridging Markov chain Monte Carlo techniques and Tierney-Kadane approximations for progressively censored Garhy reliability models: simulation insights and a medical application, *Mathematics*, **14** (2026), 1777. https://doi.org/10.3390/math14101777

22. A. H. Alenezy, B. Elkalzah, A. F. I. Alharshan, G. A. Marei, Robust inference for stress-strength reliability of the Garhy distribution under diverse record schemes with engineering and medical applications, *Mathematics*, **14** (2026), 1940. https://doi.org/10.3390/math14111940

23. K. A. Rashedi, L. S. Diab, A. H. Alenezy, G. A. Marei, Inference for stress-strength reliability under unified hybrid censoring: a one-parameter model with applications, *Mathematics*, **14** (2026), 2041. https://doi.org/10.3390/math14122041

24. A. S. Hassan, E. I. Ali, H. Hamdani, A. M. Gemeay, A. W. Shawki, M. Elgarhy, A comparative study of estimation methods for the new Sine Topp-Leone Fréchet distribution, *Sci. Afr.*, **31** (2026), e03142. https://doi.org/10.1016/j.sciaf.2025.e03142

25. A. S. Hassan, M. Abd-Allah, Exponentiated Weibull-Lomax distribution: properties and estimation, *J. Data Sci.*, **16** (2018), 277–298. https://doi.org/10.6339/JDS.201804_16(2).0004

26. B. Meriem, A. M. Gemeay, E. M. Almetwally, Z. Halim, E. Alshawarbeh, A. T. Abdulrahman, et al., The power XLindley distribution: statistical inference, fuzzy reliability, and COVID-19 application, *J. Funct. Spaces*, **2023** (2023), 9094078. https://doi.org/10.1155/2022/9094078

27. M. N. Alshahrani, Half logistic Zeghdoudi distribution: statistical properties, estimation, simulation and applications, *J. Stat. Appl. Probab.*, **14** (2025), 165–182. https://doi.org/10.18576/jsap/140203

28. R. Shanker, M. Ray, H. R. Prodhani, Power Komal distribution with properties and application in reliability engineering, *Reliab. Theory Appl.*, **18** (2023), 591–603. https://doi.org/10.24412/1932-2321-2023-476-591-603

29. A. S. Hassan, M. Elgarhy, R. E. Mohamd, S. Alrajhi, On the alpha power transformed power Lindley distribution, *J. Probab. Stat.*, **2019** (2019), 8024769. https://doi.org/10.1155/2019/8024769

30. M. E. Ghitany, D. K. Al-Mutairi, N. Balakrishnan, L. J. Al-Enezi, Power Lindley distribution and associated inference, *Comput. Stat. Data Anal.*, **64** (2013), 20–33. https://doi.org/10.1016/j.csda.2013.02.026

31. K. Aidi, A. I. Al-Omari, R. Alsultan, The power Zeghdoudi distribution: properties, estimation, and applications to real right-censored data, *Appl. Sci.*, **12** (2022), 12081. https://doi.org/10.3390/app122312081

32. S. Kharvi, M. R. Irshad, A. I. Al-Omari, R. Alsultan, Power length-biased new XLindley distribution: properties and modeling of real data, *Mathematics*, **13** (2025), 1394. https://doi.org/10.3390/math13091394

33. S. A. Alyami, I. Elbatal, N. Alotaibi, E. M. Almetwally, M. Elgarhy, Modeling to factor productivity of the United Kingdom food chain: using a new lifetime-generated family of distributions, *Sustainability*, **14** (2022), 8942. https://doi.org/10.3390/su14148942
34. W. Weibull, A statistical distribution function of wide applicability, *J. Appl. Mech.*, **18** (1951), 293–297. https://doi.org/10.1115/1.4010337
35. S. Lee, Y. Noh, Y. Chung, Inverted exponentiated Weibull distribution with applications to lifetime data, *Commun. Stat. Appl. Methods*, **24** (2017), 227–240. https://doi.org/10.5351/CSAM.2017.24.3.227
36. N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous univariate distributions, Volume 1*, John Wiley & Sons, Inc., 1994.
37. S. Sasai, M. Pararai, F. Chipepa, On the half-logistic new Weibull Pareto distribution with applications to lifetime and reliability data, *Preprints*, 2025. https://doi.org/10.20944/preprints202507.1298.v1
38. F. Ashkar, S. Mahdi, Fitting the log-logistic distribution by generalized moments, *J. Hydrol.*, **328** (2006), 694–703. https://doi.org/10.1016/j.jhydrol.2006.01.014
39. J. C. Ehiwario, J. N. Igabari, J. E. Okoh, A comparative study on the alpha power transformed family of distributions, *Abacus*, **49** (2022), 3.
40. O. A. Saudi, H. Nagy, A new three-parameter inverted exponentiated Weibull distribution: statistical inference and application, *Egypt. Stat. J.*, **68** (2024), 34–64.
41. A. M. Basheer, Alpha power inverse Weibull distribution with reliability application, *J. Taibah Univer. Sci.*, **13** (2019), 423–432. https://doi.org/10.1080/16583655.2019.1588488
42. I. Elbatal, A. S. Hassan, A. M. Gemeay, L. S. Diab, A. B. Ghorbal, M. Elgarhy, Statistical analysis of the inverse power Zeghdoudi model: estimation, simulation and modeling to engineering and environmental data, *Phys. Scr.*, **99** (2024), 065231. https://doi.org/10.1088/1402-4896/ad46d0
43. F. Jamal, A. H. Abuzaid, M. H. Tahir, M. A. Nasir, S. Khan, W. K. Mashwani, New modified Burr III distribution, properties and applications, *Math. Comput. Appl.*, **26** (2021), 82. https://doi.org/10.3390/mca26040082
44. S. Mohammad, N. Budhathoki, Modeling COVID-19 outcomes using a new extended inverse Burr distribution, *J. Probab. Stat.*, **2026** (2026), 9920669. https://doi.org/10.1155/jpas/9920669
45. A. O. Frederick, G. A. Osuji, C. K. Onyekwere, Inverted power Ishita distribution and its application to lifetime data, *Asian J. Probab. Stat.*, **18** (2022), 1–18. https://doi.org/10.9734/ajpas/2022/v18i130433
46. F. R. de Gusmao, E. M. Ortega, G. M. Cordeiro, The generalized inverse Weibull distribution, *Stat. Pap.*, **52** (2011), 591–619. https://doi.org/10.1007/s00362-009-0271-3
47. D. H. Abdelhady, Y. M. Amer, On the inverse power Gompertz distribution, *Ann. Data Sci.*, **8** (2021), 451–473. https://doi.org/10.1007/s40745-020-00246-4
48. K. V. P. Barco, J. Mazucheli, V. Janeiro, The inverse power Lindley distribution, *Commun. Stat. Simul. Comput.*, **46** (2017), 6308–6323. https://doi.org/10.1080/03610918.2016.1202274
49. A. M. A. Al-Fattah, A. A. El-Helbawy, G. R. Al-Dayian, Inverted Kumaraswamy distribution: properties and estimation, *Pak. J. Stat.*, **33** (2017), 37–61.

50. K. S. Lomax, Business failures: another example of the analysis of failure data, *J. Amer. Stat. Assoc.*, **49** (1954), 847–852. https://doi.org/10.1080/01621459.1954.10501239

51. S. J. Dugasa, A. T. Goshu, B. G. Arero, Alpha power transformation of the Lindley probability distribution, *J. Probab. Stat.*, **2024** (2024), 9068114. https://doi.org/10.1155/2024/9068114

52. M. S. Eliwa, M. El-Morshedy, M. Ibrahim, Inverse Gompertz distribution: properties and different estimation methods with application to complete and censored data, *Ann. Data Sci.*, **6** (2019), 321–339. https://doi.org/10.1007/s40745-018-0173-0

53. P. Feigl, M. Zelen, Estimation of exponential survival probabilities with concomitant information, *Biometrics*, **21** (1965), 826–838. https://doi.org/10.2307/2528247

54. F. Jamal, M. A. Nasir, M. H. Tahir, N. H. Montazeri, The odd Burr-III family of distributions, *J. Stat. Appl. Probab.*, **6** (2017), 105–122. https://doi.org/10.18576/jsap/060109

55. New York State Department of Financial Services, *Automobile insurance company complaint rankings: beginning 2009*, New York State Open Data Portal, 2026. Available from: `https://data.ny.gov/Government-Finance/Automobile-Insurance-Company-Complaint-Rankings-Be/h2wd-9xfe`.

56. S. Khan, O. S. Balogun, M. H. Tahir, W. Almutiry, A. A. Alahmadi, An alternate generalized odd generalized exponential family with applications to premium data, *Symmetry*, **13** (2021), 2064. https://doi.org/10.3390/sym13112064

Appendix

Proofs of selected propositions from Section 5

This appendix collects the longer derivations from Section 5 that follow from the root substitution Eq (3.4), but are lengthy enough to interrupt the flow of the main text; their statements and roles are given in place in Section 5.

A.1. Proof of Proposition 4 (Raw moments)

By the root substitution Eq (3.4), $f(x)\,\mathrm{d}x = \lambda e^{-\lambda t}\,\mathrm{d}t$ and $x^r = (t+1)^{2r}$, so

$$\mathbb{E}[X^r] = \int_0^\infty (t+1)^{2r}\,\lambda e^{-\lambda t}\,\mathrm{d}t.$$

Expanding $(t+1)^{2r}$ by the binomial theorem and interchanging sum and integral (valid by non-negativity),

$$\mathbb{E}[X^r] = \sum_{k=0}^{2r} \binom{2r}{k} \int_0^\infty t^k\,\lambda e^{-\lambda t}\,\mathrm{d}t.$$

Recognizing the integral as the kth moment of $\mathrm{Exp}(\lambda)$, equal to $k!/\lambda^k$, gives Eq (5.6).

A.2. Proof of Corollary 2 (mean, variance, skewness, and kurtosis)

Substituting $r = 1, 2, 3, 4$ into Eq (5.6) gives the raw moments. The central moments follow from $\mu_2 = \mu_2' - \mu_1'^2$, $\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3$, and $\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_1'^2\mu_2' - 3\mu_1'^4$. Simplification and factoring produce Eqs (5.7) and (5.8); since all numerator coefficients are positive, $\mu_3 > 0$ and $\gamma_2 > 0$.

A.3. Proof of Proposition 6 (incomplete first moment and mean deviations)

With the root substitution the upper limit x maps to $a = \sqrt{x} - 1$, and

$$J(x) = \int_0^a (t+1)^2 \, \lambda e^{-\lambda t} \, \mathrm{d}t = \int_0^a (t^2 + 2t + 1) \, \lambda e^{-\lambda t} \, \mathrm{d}t.$$

Integrating term-by-term using

$$\int_0^a t^k \, \lambda e^{-\lambda t} \, \mathrm{d}t = \gamma(k+1, \lambda a)/\lambda^k$$

for $k = 0, 1, 2$ gives Eq (5.9); Eq (5.10) then follows from standard identities.

A.4. Proof of Proposition 8 (Shannon and Rényi Entropies)

Shannon: $-\ln f(x) = \ln(2/\lambda) + \frac{1}{2}\ln x + \lambda(\sqrt{x} - 1)$. With $T = \sqrt{X} - 1 \sim \mathrm{Exp}(\lambda)$, $\mathbb{E}[\sqrt{X} - 1] = 1/\lambda$ and

$$\mathbb{E}[\ln X] = 2\,\mathbb{E}[\ln(T+1)] = 2\int_0^\infty \ln(t+1) \, \lambda e^{-\lambda t} \, \mathrm{d}t = 2e^{\lambda} E_1(\lambda),$$

which gives Eq (5.12). *Rényi:* For $\rho > 0$, $\rho \neq 1$, substituting $v = \sqrt{x}$ in $\int_1^\infty f(x)^\rho \, \mathrm{d}x$ gives $2(\lambda/2)^\rho e^{\rho\lambda} \int_1^\infty v^{1-\rho} e^{-\rho\lambda v} \, \mathrm{d}v$; the transformation $y = \rho\lambda v$ converts the integral into $(\rho\lambda)^{\rho-2}\Gamma(2-\rho, \rho\lambda)$, and applying $\frac{1}{1-\rho}\ln(\cdot)$ completes the proof.

AIMS Press