symmetry

MDPI

Article

Classical and Bayesian Estimation of the Two-Parameter Maxwell Distribution Under Random Censoring

Minghui Liu, Wenhao Gui *, Lanxi Zhang and Zihan Zhao

School of Mathematics and Statistics, Beijing Jiaotong University, Beijing 100044, China; 23271069@bjtu.edu.cn (M.L.); 23271053@bjtu.edu.cn (L.Z.); 23271112@bjtu.edu.cn (Z.Z.)
* Correspondence: whgui@bjtu.edu.cn

Abstract

This paper investigates the problem of parameter estimation and reliability analysis for the two-parameter Maxwell distribution under a random censoring mechanism. To address the limitation of the traditional single-parameter Maxwell distribution in practical applications, which lacks the threshold parameter, this paper proposes a two-parameter Maxwell distribution model. By introducing a threshold parameter, this model can more accurately characterize survival data with a minimum life or guaranteed operating time. Specifically, we construct a random censoring data model wherein both the failure time and censoring time are assumed to follow a two-parameter Maxwell distribution. The main research contents include: establishing a randomly censored data model, deriving classical inference methods based on maximum likelihood estimation. Under the general entropy loss function, Bayesian estimation is conducted using conjugate inverse Gamma priors for scale parameters and a uniform prior for the threshold parameter. A hybrid MCMC algorithm is implemented to generate posterior samples and construct highest posterior density credible intervals. We compare their performance through Monte Carlo simulations, evaluating finite-sample behavior in terms of bias, mean squared error, and interval estimation, and finally validating the practicality and superiority of the two-parameter model using real medical datasets from a colon cancer clinical trial. The results demonstrate that the two-parameter Maxwell distribution can more accurately describe survival data with threshold characteristics and outperforms the single-parameter model in terms of model fit and reliability estimation.

Keywords: random censoring; two-parameter Maxwell distribution; maximum likelihood estimation; Bayesian estimation; reliability analysis; survival analysis

Academic Editor: Hsien-Chung Wu

Received: 24 January 2026
Revised: 3 March 2026
Accepted: 9 March 2026
Published: 12 March 2026

1. Introduction

Conducting life testing experiments is inherently resource-intensive, demanding significant time and financial investment. To curtail both the economic burden and the experimental timeline, researchers routinely implement various censoring strategies in lifetime data analysis. The conventional Type I and Type II censoring mechanisms remain the most extensively utilized. More recently, increasingly flexible schemes—such as progressive, hybrid, and adaptive progressive censoring—have been thoroughly investigated in the statistical literature. A comprehensive analysis under an adaptive progressively Type-II censored framework, for instance, is provided in [1,2].

Random censoring represents another practically vital form of censoring, which arises when a test unit is lost or withdrawn from the experiment prior to its failure. To model

data under this mechanism, inference for various lifetime distributions has been developed, such as for the exponential, Rayleigh, and Weibull distributions [3]. Recent work continues to extend this line of research to other flexible models. For instance, reference [4] provided a comprehensive framework for the randomly censored Kumaraswamy distribution, employing both classical and Bayesian methods.

Various parametric models are extensively utilized in both reliability theory and survival analysis. Many other distributions, including the Maxwell or Maxwell–Boltzmann distribution, also have gained popularity as lifetime models. For an application within reliability theory, see, for instance, [5]. Reference [6] demonstrated that their proposed model achieved improved model fitting performance to failure data than popular models in certain situations. For the Maxwell distribution, subsequent scholarly work has emphasized inference methodologies under a range of advanced censoring mechanisms: the study by [7] conducted under a progressively Type-II randomly censored scheme, and the analysis presented in [2] is based on adaptive progressively Type-II censored data. Parallel to these developments, inference under random censoring has also been actively explored for other lifetime models, such as the bivariate inverse Weibull distribution [8]. Furthermore, methodological advances like E-Bayesian estimation [9] have been introduced to refine the estimation procedures. Estimation methods for the Maxwell distribution itself have also been explored in different contexts, including various point estimation techniques [10] and minimax estimation under different loss functions [11].

However, existing research on the Maxwell distribution has predominantly focused on single-parameter scale models, which assume that the lifetime begins at time zero. This assumption does not align with many practical scenarios. In reliability engineering, products often have a safe operating period or guaranteed lifetime; in medical research, there may be a latency period from treatment intervention to symptom manifestation. Ignoring this threshold characteristic and directly applying a single-parameter model may lead to biased estimates of key reliability indicators such as mean lifetime, failure rate, and reliability function. Therefore, to address this limitation, this paper extends the classical single-parameter Maxwell distribution to a two-parameter model that incorporates a threshold parameter to explicitly account for a threshold or guarantee time, alongside the scale parameter. We then focus on the statistical inference of this model under randomly censored data, assuming that both the failure time and the random censoring time follow two-parameter Maxwell distributions with a common threshold parameter μ but different scale parameters. This shared threshold assumption not only ensures model identifiability but also enhances the practical adaptability of the model in applications where an initial failure-free period is inherent to all units.

The novelty and significance of this contribution to distribution theory are threefold. First, we introduce a threshold parameter into the Maxwell distribution, fundamentally extending the distributional family to accommodate data. Second, we establish the first complete inferential framework for this two-parameter model under random censoring, addressing the substantial technical challenges posed by the resulting likelihood equations and intractable posterior distributions. Third, through rigorous real-data validation, we demonstrate that this theoretical extension yields statistically superior model fit and produces reliability estimates with clear practical interpretability.

This study aims to systematically explore classical statistical inference and Bayesian inference methods in the context of the two-parameter Maxwell distribution with randomly censored observations. We establish a random censoring data model and investigate classical methods including maximum likelihood estimation (MLE), as well as Bayesian methods. Regarding prior specification, the scale parameter is assigned a conjugate inverse Gamma prior, and the threshold parameter is given a non-informative uniform prior. The derivation

of posterior Bayesian estimates for parameters and reliability indicators is performed under the generalized entropy loss function. The posterior distribution's complexity necessitates the use of a Markov Chain Monte Carlo (MCMC) technique. Specifically, a hybrid method incorporating both Gibbs sampling and the Metropolis-Hastings algorithm is utilized to generate posterior samples. These samples then serve as the basis for constructing the highest posterior density (HPD) credible intervals for the parameters.

The finite-sample performance of various estimation methods is evaluated and compared through comprehensive Monte Carlo simulations, with assessments based on bias, mean squared error, and interval estimation. Finally, through one real lifetime dataset, we verify the practicality and effectiveness of the proposed methods. An evaluation of goodness-of-fit is conducted, specifically by comparing the proposed model against its single-parameter counterpart, thereby highlighting the pronounced advantages inherent in the two-parameter model.

The remainder of this paper is structured in the following manner. Section 2 first outlines the two-parameter Maxwell distribution model within a random censoring setting and provides derivations of essential reliability metrics. Next, Section 3 details classical estimation techniques, with emphasis on maximum likelihood estimation. Section 4 then formulates the Bayesian estimation framework. Subsequently, Section 5 assesses the performance of the proposed methods through Monte Carlo simulations. Following this, Section 6 illustrates an application of the methodology to real survival datasets. Finally, Section 7 offers concluding remarks and summarizes the main findings of the study.

2. Model Assumptions

2.1. The Two-Parameter Maxwell Distribution

The probability density function (PDF) corresponding to the two-parameter Maxwell distribution can be expressed as

$$f(x;\theta,\mu) = \frac{4}{\sqrt{\pi}}\frac{1}{\theta^{3/2}}(x-\mu)^2\,\mathrm{e}^{-\frac{(x-\mu)^2}{\theta}}, \quad x \geq \mu,\ \theta \geq 0,\ \mu \in R,$$

where μ is the threshold parameter, representing the guaranteed lifetime or latency period during which no failure event can occur, and θ is the scale parameter controlling the overall dispersion of the distribution. This specification assumes that the failure mechanism only begins to operate after time μ, and that the resulting failure times follow the Maxwell form. The corresponding cumulative distribution function (CDF) is

$$F(x;\theta,\mu) = \Gamma\left(\frac{(x-\mu)^2}{\theta}, \frac{3}{2}\right), \quad x \geq \mu,\ \theta \geq 0,\ \mu \in R,$$

where $\Gamma(x,a)$ is the incomplete gamma ratio, defined as $\Gamma(x,a) = \frac{1}{\Gamma(a)}\int_0^x e^{-w}w^{a-1}\,\mathrm{d}w$.

In reliability engineering and lifetime data analysis, the two-parameter Maxwell distribution is often used to model the lifetime of components that fail due to cumulative damage or degradation. It is particularly suitable for situations where a minimum safe life exists, represented by the threshold parameter μ.

To evaluate and predict the reliability of systems following this distribution, several core metrics are typically examined.

The mean time to system failure (MTSF) is

$$\mu' = E[X] = \mu + 2\sqrt{\frac{\theta}{\pi}}.$$

Meanwhile, the reliability function is

$$S(t) = R(t) = P[X > t] = 1 - \Gamma\left(\frac{(t-\mu)^2}{\theta}, \frac{3}{2}\right), \quad t \geq \mu,\ \theta \geq 0,\ \mu \in R,$$

and the failure rate function is

$$h(t) = \frac{f(t)}{R(t)} = \frac{\frac{4}{\sqrt{\pi}}\frac{1}{\theta^{\frac{3}{2}}}(t-\mu)^2 e^{-(t-\mu)^2/\theta}}{1 - \Gamma\left(\frac{(t-\mu)^2}{\theta}, \frac{3}{2}\right)}, \quad t \geq \mu,\ \theta > 0.$$

As shown in Figure 1, the blue solid curve represents the standard one-parameter Maxwell distribution with $\mu = 0$, while the red dashed curve shows the two-parameter version with a threshold parameter $\mu = 0.5$. The addition of the threshold parameter μ provides greater modeling flexibility, allowing the distribution to capture data with non-zero minimum values or left-censored observations.

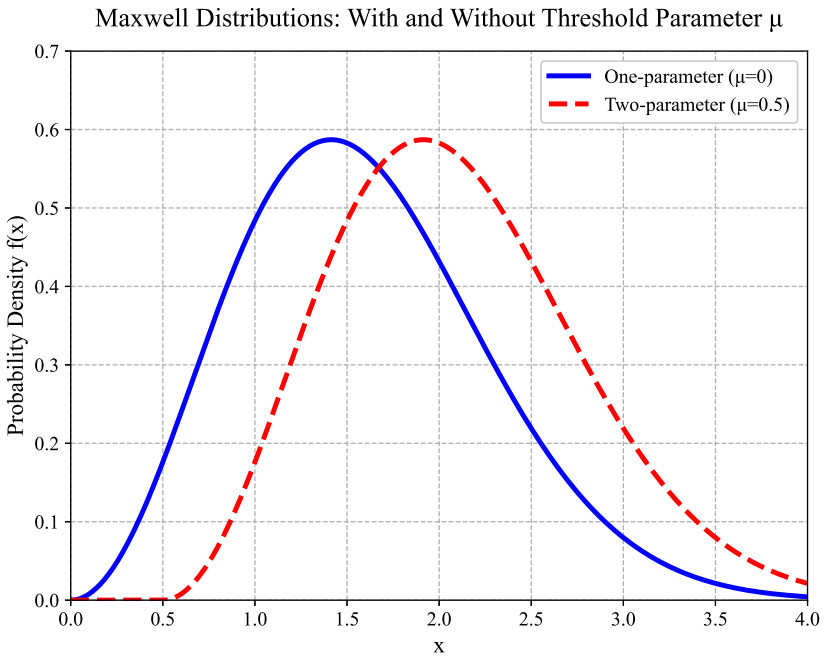

Figure 1. Comparison of single-parameter and two-parameter Maxwell distributions.

2.2. Randomly Censored Two-Parameter Maxwell Distribution

Assume that a life test is conducted on n items, and let $X_1, X_2, \ldots, X_n$ be their lifetimes, which are i.i.d. random variables following a two-parameter Maxwell distribution, hereafter abbreviated as MW(θ, μ). μ is the threshold parameter and θ is the scale parameter. The PDF and CDF of X_i are $f_X(t;\theta,\mu)$ and $F_X(t;\theta,\mu)$, respectively. $T_1, T_2, \ldots, T_n$ are the corresponding random censoring times,which are also i.i.d. random variables following a two-parameter Maxwell distribution, hereafter abbreviated as MW(λ, μ) with PDF $f_T(t;\lambda,\mu)$ and CDF $F_T(t;\lambda,\mu)$.

In actual observation, only one of X_i and T_i can be obtained. We define the actual observation time as

$$Y_i = \min(X_i, T_i), i = 1, 2, \ldots, n.$$

and the indicator variable D_i is

$$D_i = \begin{cases} 1, & \text{if } X_i \leq T_i, \\ 0, & \text{if } X_i > T_i, \end{cases} \qquad \text{for } i = 1, 2, \ldots, n.$$

Here, the indicator variable D_i follows a Bernoulli distribution, whose probability mass function is given by $P[D_i = j]$ with $j = 0, 1$.

Since X_i and T_i are independent. It can be shown that the joint probability distribution of Y and D is

$$f_{Y,D}(y, d; \theta, \mu, \lambda) = \{f_X(y; \theta, \mu)[1 - F_T(y; \lambda, \mu)]\}^d \{f_T(y; \lambda, \mu)[1 - F_X(y; \theta, \mu)]\}^{1-d},$$

$$y \geq \mu, \quad d = 0, 1.$$

$$f_{Y,D}(y, d; \theta, \mu, \lambda) = \left\{ \frac{4}{\sqrt{\pi}} \frac{1}{\theta^{\frac{3}{2}}} (y-\mu)^2 e^{-\frac{(y-\mu)^2}{\theta}} \left[1 - \Gamma\left(\frac{(y-\mu)^2}{\lambda}, \frac{3}{2}\right)\right] \right\}^d \cdot \left\{ \frac{4}{\sqrt{\pi}} \frac{1}{\lambda^{\frac{3}{2}}} (y-\mu)^2 e^{-\frac{(y-\mu)^2}{\lambda}} \left[1 - \Gamma\left(\frac{(y-\mu)^2}{\theta}, \frac{3}{2}\right)\right] \right\}^{1-d},$$

$$y \geq \mu, \quad d = 0, 1.$$

Therefore, the probability that an item fails before being censored can be computed by the integral formula above. It is worth noting that when both X and T share a threshold parameter μ,The failure probability does not depend on μ, as shown by the variable substitution $u = t - \mu$. This implies that the results for the single-parameter Maxwell distribution in Krishna et al. [12] are directly applicable to the two-parameter case.

3. Classical Statistical Inference Methods

3.1. Expected Time on Test Under Random Censoring

Consider a life-testing experiment with n independent items. Let $X_i \sim \text{MW}(\theta, \mu)$ representing the failure time for the i-th item, and $T_i \sim \text{MW}(\lambda, \mu)$ denote its associated random censoring time, with μ being a common threshold parameter. The two sequences $\{X_i\}$ and $\{T_i\}$ are independent.

For each item, we observe $Y_i = \min(X_i, T_i)$, where Y_i is the observed time. These Y_i i.i.d. with distribution $F_Y(z)$ .Because X_i and T_i are independent, the survival function of Y_i for $z \geq \mu$ is

$$F_Y(z) = P(Y_i > z) = P(X_i > z,\, T_i > z),$$

with $F_X(z) = \Gamma\left(\frac{(z-\mu)^2}{\theta}, \frac{3}{2}\right)$ and $F_T(z) = \Gamma\left(\frac{(z-\mu)^2}{\lambda}, \frac{3}{2}\right)$. Hence the CDF of Y_i is

$$F_Y(z) = 1 - \left[1 - F_X(z)\right]\left[1 - F_T(z)\right], \qquad z \geq \mu, \tag{1}$$

and $F_Y(z) = 0$ for $z < \mu$.

The experiment continues until the last item yields its observation; therefore the total duration of the test is

$$Z_n = \max\{Y_1, Y_2, \ldots, Y_n\}.$$

Its CDF follows immediately from the i.i.d. property of the Y_i

$$F_{Z_n}(z) = \left[F_Y(z)\right]^n, \tag{2}$$

The expected time on test (ETT) is defined as $\text{ETT}(n) = E[Z_n]$. The ETT has been studied for other distributions under random censoring, such as for the Pareto distribution [13]. Here we extend this concept to the two-parameter Maxwell distribution.

Using the identity $E[Z_n] = \mu + \int_{\mu}^{\infty} [1 - F_{Z_n}(z)]\, dz$ for a non-negative random variable shifted by μ, we obtain the final expression for the ETT

$$ETT(n) = \mu + \int_0^{\infty} \left\{1 - \left[1 - \left(1 - \Gamma(\frac{u^2}{\theta}, \frac{3}{2})\right)\left(1 - \Gamma(\frac{u^2}{\lambda}, \frac{3}{2})\right)\right]^n\right\} du, \tag{3}$$

where the substitution $u = z - \mu$ has been used to simplify the lower limit of integration.

The integral in Equation (3) doesn't possess a closed-form solution; it should be evaluated numerically for given n, θ, λ and μ. In the special case $\mu = 0$, Equation (3) reduces to the well-known ETT formula for the one-parameter Maxwell distribution [12].

3.2. Maximum Likelihood Estimation

Substituting the expressions for the PDF and CDF given in Section 2, from the observed data vectors $\boldsymbol{y} = (y_1, \ldots, y_n)$ and $\boldsymbol{d} = (d_1, \ldots, d_n)$, the corresponding likelihood function can be constructed.

The likelihood function for randomly censored data consists of three types of contributions

- For items that failed ($d_i = 1$), we observe the exact failure time y_i, and their contribution is $1 - F_T(y_i; \lambda, \mu)$.
- For items that were censored ($d_i = 0$), we only know that they survived until the censoring time y_i, and their contribution is $1 - F_X(y_i; \theta, \mu)$.

we obtain the likelihood function

$$\begin{aligned} L(\mu, \theta, \lambda \mid \boldsymbol{y}, \boldsymbol{d}) = C \cdot \frac{1}{\theta^{3m/2}} \cdot \frac{1}{\lambda^{3(n-m)/2}} \\ \times \exp\left[-\frac{1}{\theta}\sum_{i=1}^{n} d_i (y_i - \mu)^2 - \frac{1}{\lambda}\sum_{i=1}^{n}(1 - d_i)(y_i - \mu)^2\right] \\ \times \prod_{i=1}^{n}\left[1 - \Gamma\left(\frac{(y_i - \mu)^2}{\lambda}, \frac{3}{2}\right)\right]^{d_i} \\ \times \prod_{i=1}^{n}\left[1 - \Gamma\left(\frac{(y_i - \mu)^2}{\theta}, \frac{3}{2}\right)\right]^{1-d_i}, \end{aligned} \tag{4}$$

where

$$C = \left(\frac{4}{\sqrt{\pi}}\right)^n \prod_{i=1}^{n} (y_i - \mu)^2, \qquad m = \sum_{i=1}^{n} d_i$$

The log-likelihood function is then

$$\begin{aligned} \ell(\mu, \theta, \lambda \mid \boldsymbol{y}, \boldsymbol{d}) = \log C - \frac{3m}{2}\log\theta - \frac{3(n-m)}{2}\log\lambda \\ - \frac{1}{\theta}\sum_{i=1}^{n} d_i (y_i - \mu)^2 - \frac{1}{\lambda}\sum_{i=1}^{n}(1 - d_i)(y_i - \mu)^2 \\ + \sum_{i=1}^{n} d_i \log\left[1 - \Gamma\left(\frac{(y_i - \mu)^2}{\lambda}, \frac{3}{2}\right)\right] \\ + \sum_{i=1}^{n}(1 - d_i)\log\left[1 - \Gamma\left(\frac{(y_i - \mu)^2}{\theta}, \frac{3}{2}\right)\right]. \end{aligned} \tag{5}$$

Following the well-established approach for parameter estimation in the Maxwell distribution under censoring, as employed in recent study [2], the first-order partial derivatives of the log-likelihood function are subsequently derived for our two-parameter model under random censoring. The key step involves differentiating terms containing the incomplete gamma function $\Gamma(\cdot, \frac{3}{2})$.

By differentiating the log-likelihood function partially with respect to θ, we obtain

$$\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta} = -\frac{3m}{2\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} d_i(y_i-\mu)^2 + \sum_{i=1}^{n}(1-d_i)\frac{\partial}{\partial \theta}\log\left[1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right].$$

Using the chain rule and the derivative formula for the incomplete gamma function

$$\frac{\partial}{\partial \theta}\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right) = -\frac{3}{2\theta}\left[\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{5}{2}\right) - \Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right],$$

we obtain

$$\frac{\partial}{\partial \theta}\log\left[1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right] = \frac{-\frac{\partial}{\partial \theta}\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)} = \frac{\frac{3}{2\theta}\left[\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{5}{2}\right) - \Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)}.$$

Therefore,

$$\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta} = -\frac{3m}{2\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} d_i(y_i-\mu)^2 - \frac{3}{2\theta}\sum_{i=1}^{n}\frac{(1-d_i)\left[\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{5}{2}\right) - \Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)}.$$

Similarly, by differentiating the log-likelihood function partially with respect to λ and μ, we obtain

$$\begin{aligned}\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \lambda} = &-\frac{3(n-m)}{2\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^{n}(1-d_i)(y_i-\mu)^2 \\ &- \frac{3}{2\lambda}\sum_{i=1}^{n}\frac{d_i\left[\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{5}{2}\right) - \Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)}.\end{aligned} \tag{6}$$

$$\begin{aligned}\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \mu} = &-2\sum_{i=1}^{n}\frac{1}{y_i-\mu} + \frac{2}{\theta}\sum_{i=1}^{n} d_i(y_i-\mu) + \frac{2}{\lambda}\sum_{i=1}^{n}(1-d_i)(y_i-\mu) \\ &- \frac{3}{y_i-\mu}\sum_{i=1}^{n}\frac{(1-d_i)\left[\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{5}{2}\right) - \Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)} \\ &- \frac{3}{y_i-\mu}\sum_{i=1}^{n}\frac{d_i\left[\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{5}{2}\right) - \Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)}.\end{aligned} \tag{7}$$

To facilitate numerical optimization, we decompose the score functions into two components: a "standard" part that resembles the complete-data likelihood, and a "correction" part ψ that accounts for the censoring through the incomplete gamma function.

The standard part represents what the score equations would look like if we had observed all failure times (no censoring). The correction terms ψ adjust these equations to account for the fact that some observations are censored. When there is no censoring, the correction term ψ_θ vanishes, and the score function reduces to its complete-data form.

The first-order partial derivatives admit the compact representation.

$$\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta} = -\frac{3m}{2\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} d_i(y_i-\mu)^2 + \psi_\theta, \tag{8}$$

$$\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \lambda} = -\frac{3(n-m)}{2\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^{n}(1-d_i)(y_i-\mu)^2 + \psi_\lambda, \tag{9}$$

$$\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \mu} = -2\sum_{i=1}^{n}\frac{1}{y_i-\mu} + \frac{2}{\theta}\sum_{i=1}^{n} d_i(y_i-\mu) + \frac{2}{\lambda}\sum_{i=1}^{n}(1-d_i)(y_i-\mu) + \psi_\mu, \tag{10}$$

where

$$\psi_\theta = -\frac{3}{2\theta}\sum_{i=1}^{n}\frac{(1-d_i)\left[\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{5}{2}\right)-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)}, \tag{11}$$

$$\psi_\lambda = -\frac{3}{2\lambda}\sum_{i=1}^{n}\frac{d_i\left[\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{5}{2}\right)-\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)}, \tag{12}$$

$$\begin{aligned}\psi_\mu = &-\frac{3}{y_i-\mu}\sum_{i=1}^{n}\frac{(1-d_i)\left[\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{5}{2}\right)-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\theta},\frac{3}{2}\right)}\\ &-\frac{3}{y_i-\mu}\sum_{i=1}^{n}\frac{d_i\left[\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{5}{2}\right)-\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)\right]}{1-\Gamma\left(\frac{(y_i-\mu)^2}{\lambda},\frac{3}{2}\right)}.\end{aligned} \tag{13}$$

The second-order partial derivatives, needed for the observed Fisher information matrix, are

$$\frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta^2} = \frac{3m}{2\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} d_i(y_i-\mu)^2 + \psi'_\theta, \tag{14}$$

$$\frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \lambda^2} = \frac{3(n-m)}{2\lambda^2} - \frac{2}{\lambda^3}\sum_{i=1}^{n}(1-d_i)(y_i-\mu)^2 + \psi'_\lambda, \tag{15}$$

$$\frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \mu^2} = 2\sum_{i=1}^{n}\frac{1}{(y_i-\mu)^2} - \frac{2m}{\theta} - \frac{2(n-m)}{\lambda} + \psi'_\mu, \tag{16}$$

where ψ'_θ, ψ'_λ, and ψ'_μ are the second-order derivatives of the corresponding ψ terms, which involve higher-order incomplete gamma functions and can be derived following similar steps as in Krishna et al. [12].

The mixed derivatives are

$$\frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta \partial \lambda} = \frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \lambda \partial \theta} = 0, \tag{17}$$

$$\frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta \partial \mu} = \psi'_{\theta\mu}, \quad \frac{\partial^2 \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \lambda \partial \mu} = \psi'_{\lambda\mu}, \tag{18}$$

where $\psi'_{\theta\mu}$ and $\psi'_{\lambda\mu}$ are mixed derivatives of the corresponding ψ terms, which correspond to Equations (8)–(10). These equations are nonlinear and fail to produce closed-form solutions for (μ,θ,λ). Thus numerical iterative methods are required, for example the Newton–Raphson method and optimization algorithms available.

The maximum likelihood estimators $\hat{\mu},\hat{\theta},\hat{\lambda}$ are solutions to the score equations

$$\frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \mu} = 0, \quad \frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \theta} = 0, \quad \frac{\partial \ell(\mu,\theta,\lambda \mid \boldsymbol{y},\boldsymbol{d})}{\partial \lambda} = 0,$$

where the partial derivatives are given by Equations (8)–(10).

The parameters are naturally constrained:

- $\mu < \min\{y_i\}$, since the lifetime X_i must exceed the threshold parameter μ.
- $\theta > 0, \lambda > 0$, as they are scale parameters.

The log-likelihood function is continuous on the domain $\mu < \min y_i,\ \theta > 0,\ \lambda > 0$. Next, we explore the variation of parameters within the feasible domain through numerical simulation. Since both θ and λ are scale parameters and their effects on the likelihood function are symmetric in structure, we focus our analysis on θ without loss of generality.

Numerical results show the following behavior of the log-likelihood: when μ takes a value near its lower boundary, the log-likelihood is about -83.74; near the optimum it reaches about -34.34; and as μ approaches $\min y_i$ from below, the log-likelihood tends to $-\infty$. For θ, when it takes a very small value, the log-likelihood tends to $-\infty$; near the optimum it is about -39.79; and as θ becomes very large, the log-likelihood again tends to $-\infty$.

Since $\ell(\mu, \theta, \lambda \mid \boldsymbol{y}, \boldsymbol{d})$ is continuous on the feasible domain and tends to $-\infty$ near the lower boundaries, while remaining finite for moderate parameter values, it must attain a maximum at some interior point $(\hat{\mu}, \hat{\theta}, \hat{\lambda})$. This guarantees the existence of a solution to the score equations.

Upon obtaining the MLEs $\hat{\mu}$, $\hat{\theta}$, and $\hat{\lambda}$, the corresponding MLEs for the reliability function $R(t)$, the failure rate function $h(t)$, and the MTSF are subsequently derived. This derivation is based on the invariance property inherent to maximum likelihood estimation.

$$\hat{R}(t) = 1 - \Gamma\left(\frac{(t-\hat{\mu})^2}{\hat{\theta}}, \frac{3}{2}\right), \quad t \geq \hat{\mu}, \tag{19}$$

$$\hat{h}(t) = \frac{\hat{f}(t)}{\hat{R}(t)} = \frac{\dfrac{4}{\sqrt{\pi}} \dfrac{1}{\hat{\theta}^{3/2}} (t-\hat{\mu})^2 \exp\left(-\dfrac{(t-\hat{\mu})^2}{\hat{\theta}}\right)}{1 - \Gamma\left(\dfrac{(t-\hat{\mu})^2}{\hat{\theta}}, \dfrac{3}{2}\right)}, \quad t \geq \hat{\mu}, \tag{20}$$

$$\widehat{\mathrm{MTSF}} = \hat{\mu} + 2\sqrt{\frac{\hat{\theta}}{\pi}}. \tag{21}$$

Similarly, the MLE of the censoring distribution parameter λ, denoted as $\hat{\lambda}$, may be acquired from solving Equation.

3.3. Confidence Intervals

Once an estimate $\hat{\mu}$ of the threshold parameter has been obtained , we may treat μ as known and focus on interval estimation for the scale parameters θ and λ. This simplification is practically relevant because μ often represents a physical threshold that can be determined independently. the derivation of both asymptotic and bootstrap confidence intervals is presented for under the assumption that μ is fixed at $\hat{\mu}$.

With μ known, the log-likelihood function Equation (5) reduces to

$$\begin{aligned} \ell(\theta, \lambda \mid \mu, \boldsymbol{y}, \boldsymbol{d}) = C &- \frac{3m}{2}\log\theta - \frac{3(n-m)}{2}\log\lambda \\ &- \frac{1}{\theta}\sum_{i=1}^{n} d_i (y_i - \mu)^2 - \frac{1}{\lambda}\sum_{i=1}^{n}(1-d_i)(y_i-\mu)^2 \\ &+ \sum_{i=1}^{n} d_i \log\Big[1 - \Gamma\Big(\frac{(y_i-\mu)^2}{\lambda}, \frac{3}{2}\Big)\Big] \\ &+ \sum_{i=1}^{n}(1-d_i)\log\Big[1 - \Gamma\Big(\frac{(y_i-\mu)^2}{\theta}, \frac{3}{2}\Big)\Big], \end{aligned} \tag{22}$$

where $m = \sum_{i=1}^{n} d_i$.

Let $\boldsymbol{\eta} = (\theta, \lambda)^\top$. The Fisher information matrix can be expressed as

$$I(\boldsymbol{\eta}) = -E\left[\frac{\partial^2 \ell(\theta, \lambda \mid \mu, \boldsymbol{y}, \boldsymbol{d})}{\partial \boldsymbol{\eta}\, \partial \boldsymbol{\eta}^\top}\right]. \tag{23}$$

Based on the adopted model assumptions, The cross partial derivative of the log-likelihood function, taken with respect to θ and λ, has an expected value of zero. Hence $I(\boldsymbol{\eta})$ is diagonal

$$I(\boldsymbol{\eta}) = \begin{pmatrix} I_{\theta\theta} & 0 \\ 0 & I_{\lambda\lambda} \end{pmatrix}, \tag{24}$$

with diagonal elements

$$I_{\theta\theta} = \frac{3m}{2\theta^2} - \frac{2}{\theta^3} E\Big[\sum_{i=1}^{n} d_i (y_i - \mu)^2\Big] + \frac{\partial \psi_\theta}{\partial \theta}, \tag{25}$$

$$I_{\lambda\lambda} = \frac{3(n-m)}{2\lambda^2} - \frac{2}{\lambda^3} E\Big[\sum_{i=1}^{n} (1 - d_i)(y_i - \mu)^2\Big] + \frac{\partial \psi_\lambda}{\partial \lambda}. \tag{26}$$

The expectations can be evaluated numerically once the censoring probability $p = P(X_i < T_i)$ is estimated.

The observed Fisher information matrix evaluated at the MLEs is utilized in practice.

$$\hat{I}(\hat{\theta}, \hat{\lambda}) = -\begin{pmatrix} \dfrac{\partial^2 \ell(\theta, \lambda \mid \mu)}{\partial \theta^2} & \dfrac{\partial^2 \ell(\theta, \lambda \mid \mu)}{\partial \theta\, \partial \lambda} \\ \dfrac{\partial^2 \ell(\theta, \lambda \mid \mu)}{\partial \lambda\, \partial \theta} & \dfrac{\partial^2 \ell(\theta, \lambda \mid \mu)}{\partial \lambda^2} \end{pmatrix}_{(\theta,\lambda)=(\hat{\theta},\hat{\lambda})} = \begin{pmatrix} \hat{I}_{\theta\theta} & 0 \\ 0 & \hat{I}_{\lambda\lambda} \end{pmatrix}, \tag{27}$$

where the second derivatives are obtained from Equations (14)–(15) with μ fixed.

Under regularity conditions, the MLE $(\hat{\theta}, \hat{\lambda})$ is asymptotically normal

$$\begin{pmatrix} \hat{\theta} \\ \hat{\lambda} \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} \theta \\ \lambda \end{pmatrix},\ \hat{I}^{-1}(\hat{\theta}, \hat{\lambda})\right), \qquad \hat{I}^{-1} = \begin{pmatrix} 1/\hat{I}_{\theta\theta} & 0 \\ 0 & 1/\hat{I}_{\lambda\lambda} \end{pmatrix}. \tag{28}$$

Consequently, the asymptotic covariance between the two estimators is zero. A $100(1-\alpha)\%$ Wald confidence interval for θ is

$$\hat{\theta} \pm z_{1-\alpha/2} \sqrt{1/\hat{I}_{\theta\theta}}, \tag{29}$$

and similarly for λ.

Because $\theta, \lambda > 0$, a log-transformation often yields better normal approximation. The interval for θ based on the log-transform is

$$\Big[\hat{\theta} \exp\big(-z_{1-\alpha/2} \sqrt{1/\hat{I}_{\theta\theta}}/\hat{\theta}\big),\ \hat{\theta} \exp\big(z_{1-\alpha/2} \sqrt{1/\hat{I}_{\theta\theta}}/\hat{\theta}\big)\Big], \tag{30}$$

and an analogous expression holds for λ.

For small or moderate samples, bootstrap methods offer a robust alternative. As shown in Algorithm 1, we employ a parametric bootstrap procedure.

Algorithm 1 Percentile Bootstrap Confidence Intervals

Input: MLEs $\hat{\theta}$, $\hat{\lambda}$, $\hat{\mu}$; sample size n; bootstrap replicates B; confidence level $1-\alpha$.
Output: Bootstrap CI for θ and λ.
Fix $\mu = \hat{\mu}$.
for $b = 1$ **to** B **do**
  Generate bootstrap sample $\{(y_i^*, d_i^*)\}_{i=1}^n$ from $(\hat{\theta}, \hat{\lambda})$.
  Compute MLEs $(\hat{\theta}_b^*, \hat{\lambda}_b^*)$ with μ fixed.
end for
Sort $\{\hat{\theta}_b^*\}_{b=1}^B$ and $\{\hat{\lambda}_b^*\}_{b=1}^B$.
Compute empirical quantiles:
  $\theta_{\text{low}} = \hat{\theta}^*_{(\alpha/2)}$, $\theta_{\text{up}} = \hat{\theta}^*_{(1-\alpha/2)}$
  $\lambda_{\text{low}} = \hat{\lambda}^*_{(\alpha/2)}$, $\lambda_{\text{up}} = \hat{\lambda}^*_{(1-\alpha/2)}$
Return intervals: $[\theta_{\text{low}}, \theta_{\text{up}}]$ and $[\lambda_{\text{low}}, \lambda_{\text{up}}]$.

4. Bayesian Estimation

Bayesian estimation of the Maxwell distribution has been studied by several authors under various censoring schemes, including reference [5] for progressive censoring, and more recently, Goel et al. [7] and Kumari et al. [2] for adaptive progressively Type-II censoring and progressively Type-II, respectively. we extend this line of research through the formulation of a Bayesian inference framework tailored for the two-parameter Maxwell distribution in the presence of random censoring.

For the scale parameters θ and λ, we adopt independent inverse gamma (IG) priors. This choice is motivated by their conjugacy properties: given that the Maxwell likelihood involves terms of the form $\exp(-(\cdot)/\theta)$ and $\theta^{-3/2}$, the inverse gamma prior naturally combines with the likelihood to yield a tractable posterior distribution. The prior densities are given by

$$\pi_1(\theta) \propto \theta^{-(a_1+1)} \exp\left(-\frac{b_1}{\theta}\right), \quad \theta > 0,\ a_1 > 0,\ b_1 > 0, \tag{31}$$

$$\pi_2(\lambda) \propto \lambda^{-(a_2+1)} \exp\left(-\frac{b_2}{\lambda}\right), \quad \lambda > 0,\ a_2 > 0,\ b_2 > 0, \tag{32}$$

where a_1, b_1, a_2, b_2 are hyper-parameters that reflect prior knowledge. For the threshold parameter μ, which represents a threshold or guarantee time, prior information is often scarce. Therefore, we assign a non-informative uniform prior over a bounded interval $[L, U]$, where L and U are chosen based on the practical support of μ (e.g., $0 < L < U < \infty$) to ensure posterior propriety.

$$\pi_3(\mu) \propto 1, \quad \mu \in [L, U]. \tag{33}$$

The joint prior for (μ, θ, λ) is then

$$\pi(\mu, \theta, \lambda) = \pi_1(\theta)\, \pi_2(\lambda)\, \pi_3(\mu). \tag{34}$$

Given the randomly censored sample $\{(y_i, d_i)\}_{i=1}^n$, the likelihood function is given by Equation (4), which can be rewritten as

$$L(\theta, \lambda \mid \mu, \boldsymbol{y}, \boldsymbol{d}) = C \cdot \theta^{-3m/2}\, \lambda^{-3(n-m)/2} \times \exp\Big[-\frac{1}{\theta} S_1(\mu) - \frac{1}{\lambda} S_2(\mu)\Big] \times \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i-\mu)^2}{\lambda}, \frac{3}{2}\Big)\Big]^{d_i} \times \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i-\mu)^2}{\theta}, \frac{3}{2}\Big)\Big]^{1-d_i}, \tag{35}$$

To simplify notation, we define two summary statistics that capture the squared deviations weighted by the censoring indicators

$$S_1(\mu) = \sum_{i=1}^{n} d_i (y_i - \mu)^2, \quad S_2(\mu) = \sum_{i=1}^{n} (1 - d_i)(y_i - \mu)^2,$$

These quantities appear naturally in the exponent of the likelihood function: $S_1(\mu)$ corresponds to the contribution from items that failed , while $S_2(\mu)$ corresponds to items that were censored .

Combining the likelihood Equation (35) with the priors Equations (31)–(33), the joint posterior density is proportional to

$$\pi(\theta, \lambda, \mu \mid \boldsymbol{y}, \boldsymbol{d}) \propto L(\theta, \mu, \lambda \mid \boldsymbol{y}, \boldsymbol{d})\ \pi_1(\theta)\, \pi_2(\lambda)\, \pi_3(\mu) \propto \theta^{-\alpha_1 - 1} \exp\Big(-\frac{\beta_1}{\theta}\Big)\ \lambda^{-\alpha_2 - 1} \exp\Big(-\frac{\beta_2}{\lambda}\Big) \times \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i-\mu)^2}{\lambda}, \frac{3}{2}\Big)\Big]^{d_i} \times \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i-\mu)^2}{\theta}, \frac{3}{2}\Big)\Big]^{1-d_i}, \tag{36}$$

where the posterior parameters are obtained by combining the likelihood contributions with the prior hyperparameters

$$\alpha_1 = \frac{3m}{2} + a_1, \quad \beta_1 = S_1(\mu) + b_1, \qquad \alpha_2 = \frac{3(n-m)}{2} + a_2, \quad \beta_2 = S_2(\mu) + b_2.$$

Here, the terms $3m/2$ and $3(n-m)/2$ arise from the kernel of the Maxwell likelihood, while a_1, b_1, a_2, b_2 are the prior hyperparameters from the inverse Gamma priors.

The normalizing constant of Equation (36) involves a triple integral that lacks a closed-form representation. Therefore, we turn to Markov Chain Monte Carlo (MCMC) simulation in order to obtain samples from the posterior distribution.

For Bayesian estimation, we employ the generalized entropy loss function (GELF). The GELF maintains its relevance in modern reliability analysis, as demonstrated by recent Bayesian studies on lifetime distributions under various censoring schemes [2,4]. Let ϕ be a parameter of interest and let $\hat{\phi}$ be an estimator of ϕ. The GELF is defined as

$$L(\phi, \hat{\phi}) = \Big(\frac{\hat{\phi}}{\phi}\Big)^{\delta} - \delta \ln\Big(\frac{\hat{\phi}}{\phi}\Big) - 1, \quad \delta \neq 0 \tag{37}$$

The parameter δ controls the asymmetry of the loss:

- $\delta > 0$ penalizes overestimation more heavily;
- $\delta < 0$ penalizes underestimation more heavily;

Under GELF, the Bayes estimator of ϕ is

$$\hat{\phi}_{\text{Bayes}} = \left[E_{\text{post}}\left(\phi^{-\delta}\right)\right]^{-1/\delta}, \tag{38}$$

where E_{post} denotes expectation of the posterior distribution. A detailed derivation of this result is provided in the Appendix A.

Special cases include $\delta = -1$, which yields the posterior mean under squared error loss; $\delta = 1$, which gives the reciprocal of the posterior harmonic mean; and $\delta \to 0$, which approaches the posterior mean via limit equivalence. Since the posterior expectation $E_{\text{post}}(\phi^{-\delta})$ generally lacks a closed form, it is approximated numerically via MCMC sampling.

MCMC Implementation

Based on the Bayesian framework, the joint posterior distribution of the parameters (θ, λ, μ) is given by the product of the likelihood function and the prior distributions:

$$\begin{aligned}\pi(\theta, \lambda, \mu \mid \boldsymbol{y}, \boldsymbol{d}) &= \frac{L(\theta, \lambda, \mu \mid \boldsymbol{y}, \boldsymbol{d}) \cdot \pi_1(\theta)\pi_2(\lambda)\pi_3(\mu)}{\int_0^\infty \int_0^\infty \int_L^U L(\theta, \lambda, \mu \mid \boldsymbol{y}, \boldsymbol{d})\pi_1(\theta)\pi_2(\lambda)\pi_3(\mu)\,\mathrm{d}\mu\,\mathrm{d}\lambda\,\mathrm{d}\theta} \\ &\propto \theta^{-(\frac{3m}{2}+a_1+1)}\lambda^{-(\frac{3(n-m)}{2}+a_2+1)}\exp\left(-\frac{S_1(\mu)+b_1}{\theta} - \frac{S_2(\mu)+b_2}{\lambda}\right) \\ &\quad \times \prod_{i=1}^{n}\left[1 - \Gamma\left(\frac{(y_i-\mu)^2}{\lambda}, \frac{3}{2}\right)\right]^{d_i} \times \prod_{i=1}^{n}\left[1 - \Gamma\left(\frac{(y_i-\mu)^2}{\theta}, \frac{3}{2}\right)\right]^{1-d_i},\end{aligned} \tag{39}$$

where $S_1(\mu) = \sum_{i=1}^n d_i(y_i-\mu)^2$, $S_2(\mu) = \sum_{i=1}^n (1-d_i)(y_i-\mu)^2$, and $m = \sum_{i=1}^n d_i$.

This posterior distribution exhibits a complex analytical structure, which manifests in three main aspects

1. Normalizing constant: The denominator in Equation (39) is a triple integral over (θ, λ, μ) whose integrand contains products of incomplete gamma functions, rendering it analytically intractable.
2. Parameter coupling: The location parameter μ appears both in the quadratic form $(y_i - \mu)^2$ and inside the incomplete gamma function argument, creating a nonlinear dependence that prevents marginalization via separation or transformation.
3. Incomplete gamma complexity: As a special function defined by an integral, $\Gamma(\cdot,\cdot)$ introduces additional complexity when appearing as a multiplicative factor. For any fixed μ, the conditional posteriors of θ and λ and , despite containing inverse gamma kernels, do not belong to any standard family; similarly, the conditional posterior of μ depends intricately on all observations and admits no closed form.

The analytical complexity described above leads directly to two consequences: first, the posterior moments of the parameters cannot be obtained through analytical integration; second, it is impossible to draw independent samples directly from the posterior distribution. Therefore, it is necessary to employ Markov Chain Monte Carlo methods by constructing an ergodic Markov chain whose stationary distribution is exactly the target posterior distribution given in Equation (39), thereby obtaining posterior samples and approximating the quantities of interest.

From Equation (36), the full conditional density of θ is

$$\begin{aligned}\pi(\theta \mid \lambda, \mu, \boldsymbol{y}, \boldsymbol{d}) &\propto \theta^{-\alpha_1-1}\exp\left(-\frac{\beta_1}{\theta}\right) \\ &\quad \times \prod_{i=1}^{n}\left[1 - \Gamma\left(\frac{(y_i-\mu)^2}{\theta}, \frac{3}{2}\right)\right]^{1-d_i}.\end{aligned} \tag{40}$$

This density is not of a known form; we therefore update θ using a Metropolis–Hastings with a log-normal proposal. The log-normal proposal distribution is chosen because its support is the positive real line, which matches the domain of the scale parameter θ. This proposal distribution is symmetric, so the acceptance probability simplifies to the ratio of posterior densities $\min\big(1, \pi(\theta^* \mid \cdot)/\pi(\theta^{(t-1)} \mid \cdot)\big)$. By tuning the scale parameter of the log-normal distribution, the average acceptance rate is controlled within the range of 20–40%. This range is commonly recommended in practice to balance mixing efficiency and sampling stability for complex posterior distributions [14].

Although the conditional posterior in Equation (40) contains an inverse gamma kernel, the presence of the incomplete gamma product terms prevents it from reducing to a standard distribution. Consequently, all parameters require Metropolis-Hastings updates, and the "hybrid" designation reflects the fact that the sampler structure originates from a Gibbs framework where only the intractable components are handled via MH steps.

Similarly, the full conditional density for λ is

$$\pi(\lambda \mid \theta, \mu, \boldsymbol{y}, \boldsymbol{d}) \propto \lambda^{-\alpha_2 - 1} \exp\Big(-\frac{\beta_2}{\lambda}\Big) \times \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i - \mu)^2}{\lambda}, \frac{3}{2}\Big)\Big]^{d_i}. \tag{41}$$

Again, a MH sampler employing a log-normal proposal distribution is applied. Following the same updating strategy as for θ, a symmetric log-normal proposal distribution is used, with its scale parameter adjusted to achieve a desirable acceptance rate, targeting the same 20–40% range as for θ.

The full conditional density for μ is

$$\pi(\mu \mid \theta, \lambda, \boldsymbol{y}, \boldsymbol{d}) \propto \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i - \mu)^2}{\lambda}, \frac{3}{2}\Big)\Big]^{d_i} \times \prod_{i=1}^{n} \Big[1 - \Gamma\Big(\frac{(y_i - \mu)^2}{\theta}, \frac{3}{2}\Big)\Big]^{1-d_i}, \qquad \mu \in [L, U]. \tag{42}$$

Since μ appears inside the incomplete gamma functions, we update it via a random-walk Metropolis sampler with a normal proposal (truncated to $[L, U]$). Due to the asymmetry of the truncated normal distribution, the acceptance probability must incorporate the ratio of proposal densities. The truncated normal proposal is chosen to ensure that all proposed values lie within the valid support $[L, U]$ for μ.

The Metropolis-Hastings updates yield an ergodic chain converging to the target posterior. The log-normal proposals for θ, λ (positive support) and truncated normal for μ (bounded support $[L, U]$) ensure irreducibility and aperiodicity, as they can reach any parameter space region with positive probability. The absolutely continuous kernels guarantee convergence to the unique stationary distribution $\pi(\theta, \lambda, \mu \mid \boldsymbol{y}, \boldsymbol{d})$ in Equation (39). By the ergodic theorem, posterior estimates are consistent. Empirical support is provided in Section 5: all Gelman-Rubin $\hat{R}$ statistics are below 1.05, and trace plots confirm convergence. This foundational structure also suggests potential extensions where the threshold parameter could be interpreted via first passage times of underlying stochastic processes [15], pointing to broader applications in time-to-event modeling. The detailed procedure is provided in Algorithm 2.

Algorithm 2 MCMC for Parameter Estimation and Reliability Metrics

1: Input: Initial values $\theta^{(0)}$, $\lambda^{(0)}$, $\mu^{(0)}$; number of iterations T; burn-in size B.
2: Output: Posterior samples and estimates of reliability metrics.
3: Initialize: Set $t = 0$, assign initial values $\theta^{(0)}$, $\lambda^{(0)}$, $\mu^{(0)}$.
4: **for** $t = 1$ **to** T **do**
5: Step 1: Update $\theta^{(t)}$ using an MH sampler based on Equation (40).
6: Step 2: Update $\lambda^{(t)}$ using an MH sampler based on Equation (41).
7: Step 3: Update $\mu^{(t)}$ using an MH sampler based on Equation (42).
8: Step 4: Compute the following using current parameter values:
- $R^{(t)}(t_0)$
- $h^{(t)}(t_0)$
- $\mathrm{MTSF}^{(t)}$

9: **end for**
10: Post-processing: Discard the first B iterations as burn-in and retain the remaining samples for statistical inference.

5. Simulation Study

A Monte Carlo simulation study is conducted to assess the finite-sample performance of the proposed estimation methods for the two-parameter Maxwell distribution within a random censoring context. We compare the maximum likelihood estimation approach from Section 3.2 with the Bayesian framework from Section 4, focusing on parameter estimation accuracy and reliability function performance. The detailed simulation procedure is described in Algorithm 3.

To evaluate the performance of the proposed estimators, we consider two metrics: the mean estimate and the mean squared error (MSE). The mean estimate and the MSE are defined as

$$\text{Mean Estimate} = \frac{1}{N}\sum_{r=1}^{N}\hat{\phi}^{(r)}, \tag{43}$$

$$\text{MSE}(\hat{\phi}) = \frac{1}{N}\sum_{r=1}^{N}(\hat{\phi}^{(r)} - \phi)^2, \tag{44}$$

where ϕ denotes the true parameter value used in the data generation process. The mean estimate reflects the average behavior of the estimator across repeated sampling, while the MSE measures its overall accuracy by combining both bias and variance.

We consider a range of parameter combinations to assess the robustness of the estimation methods under different scenarios. The experimental setup is as follows:

- Sample size: $n = 50$ is chosen as a moderate sample size commonly encountered in reliability and survival studies.
- Simulation replicates: $N = 1000$ is selected to ensure the stability and reproducibility of the Monte Carlo results.
- Parameter combinations: The threshold parameter μ is set to $\mu = 0.5, 1.0, 2.0, 5.0$ to represent different threshold effects, ranging from a small guaranteed lifetime to a relatively large latency period. The scale parameters θ and λ are selected from $\{0.5, 1.0, 4.0, 9.0, 25.0, 36.0\}$ to cover low, moderate, and high failure rate regimes. These choices yield varying censoring proportions, enabling a comprehensive assessment of estimator performance under diverse censoring levels.
- Bayesian priors: $\theta, \lambda \sim \mathrm{IG}(2, 1)$ are chosen as weakly informative conjugate priors, allowing the data to dominate the posterior inference while ensuring computational stability. The hyperparameters $(2, 1)$ yield a finite prior mean and variance, providing sufficient flexibility.

- Loss function: Generalized entropy loss with $\delta = -1$ is employed, which corresponds to the squared error loss, a commonly used symmetric loss function in Bayesian analysis.

Algorithm 3 Simulation Procedure for Performance Evaluation

for each parameter combination (θ, λ, μ) **do**
  Set Bayesian priors: $\theta, \lambda \sim \text{IG}(2,1)$, $\mu \sim \mathcal{U}(0,3)$
  for each replicate $r = 1$ **to** N **do**
    1. Generate censored dataset:
      For $i = 1$ **to** n:
        Draw $U_i \sim U(0,1)$
        $X_i = F_X^{-1}(U_i; \theta, \mu)$ {Failure time from MW(θ, μ)}
        Draw $V_i \sim U(0,1)$
        $T_i = F_T^{-1}(V_i; \lambda, \mu)$ {Censoring time from MW(λ, μ)}
        $y_i = \min(X_i, T_i)$, $d_i = I(X_i \leq T_i)$
      Store $D_r = \{(y_i, d_i)\}_{i=1}^{n}$
    2. Apply estimation methods:
      (a) MLE: Solve $\frac{\partial \ell}{\partial(\theta,\lambda,\mu)} = 0$
      (b) Bayesian: Run Metropolis-Hastings with three parallel chains (5000 burn-in, 10,000 post-burn-in, thinning = 10). Convergence assessed via $\hat{R} < 1.05$ and trace plots.
    3. Compute reliability metrics at t_0:
      $\hat{R}(t_0)$, $\hat{h}(t_0)$, $\widehat{\text{MTSF}}$
  end for
  4. Summarize performance:
    For each parameter: Compute bias, MSE
    For reliability metrics: Compute bias, MSE
end for

Based on extensive computations performed in a Python 3.9.25 Jupyter notebook environment across $N = 1000$ replicates, the main findings from Tables 1–6 and Figure 2 are as follows:

- Both the MLE and Bayesian methods demonstrate satisfactory performance in estimating the parameters (μ, θ, λ) of the two-parameter Maxwell distribution under random censoring. The bias and MSE for parameter estimates decrease with increasing sample size and decreasing censoring rate, confirming the consistency of the proposed estimators.
- For the scale parameters θ and λ, the MLE exhibits slightly superior precision (lower MSE) compared to the Bayesian estimates under the specified weakly informative priors. This difference is most pronounced in small samples ($n = 20$) and diminishes as sample size increases, as expected given the asymptotic equivalence of the two approaches.
- The threshold parameter μ is estimated with high accuracy by both methods, with MSE values consistently below 0.01 across all scenarios. This stability arises because μ is bounded above by the minimum observation, which provides strong identifiability.
- Reliability characteristics, including the MTSF, hazard function $h(s)$, and reliability function $R(s)$, are accurately recovered by both estimation approaches. Notably, $R(s)$ exhibits the lowest MSE among all reliability metrics, as it is a cumulative measure less sensitive to local fluctuations in the estimated parameters.
- For the ETT, the proposed MLE estimator ($\widehat{ETT}$) consistently outperforms the OBTT, with MSE reductions of 60–80% across all experimental conditions. This improvement

is particularly valuable in reliability practice where precise test time prediction is essential for planning purposes.

- The Gelman-Rubin $\hat{R}$ statistics were below 1.02 for all parameters across both scenarios, indicating satisfactory convergence. The slightly higher values in the challenging scenario reflect slower mixing under heavy censoring, but all remain well below the conventional 1.05 threshold. Trace plots in Figure 2 confirm that chains stabilize around true values after burn-in.
- In the challenging scenario, the sampler exhibited slower mixing, reflected in slightly higher $\hat{R}$ values and occasional need for longer burn-in. Proposals for μ near the boundary also led to occasional rejections. However, all chains eventually converged successfully, and no non-convergence was detected in the final results.

Table 1. MLE estimation results for Maxwell distribution with random censoring ($n = 50$, $N = 1000$).

True Parameters			MLE Estimation Results											
μ	θ	λ	$\hat{\mu}$		$\hat{\theta}$		$\hat{\lambda}$		$\widehat{\text{MTSF}}$		$\hat{h}(s)$		$\hat{R}(s)$	
			AV	MSE	AV	MSE	AV	MSE	AV	MSE	AV	MSE	AV	MSE
0.5	1.0	1.0	0.5378	0.0088	0.9602	0.0633	0.9544	0.0547	1.6284	0.0090	0.4782	0.4592	0.9189	0.9272
1.0	1.0	1.0	1.0200	0.0056	0.9674	0.0464	0.9612	0.0490	2.1284	0.0078	0.4782	0.4833	0.9189	0.9208
2.0	1.0	1.0	2.0198	0.0065	0.9780	0.0512	0.9925	0.0618	3.1284	0.0095	0.4782	0.4796	0.9189	0.9213
5.0	1.0	4.0	5.0383	0.0102	0.9436	0.0357	4.3901	2.9278	6.1284	0.0047	0.4782	0.4560	0.9189	0.9265
0.5	4.0	9.0	0.5579	0.0276	3.8649	0.5669	8.6855	5.6437	2.7568	0.0235	0.0670	0.0578	0.9887	0.9899
1.0	4.0	9.0	1.0369	0.0280	3.9088	0.4912	9.0726	6.8929	3.2568	0.0192	0.0670	0.0624	0.9887	0.9887
2.0	9.0	4.0	2.0661	0.0354	8.7841	5.4016	3.8109	0.7738	5.3851	0.1238	0.0204	0.0176	0.9966	0.9969
5.0	25.0	9.0	5.0354	0.0713	25.3274	51.3061	9.0816	2.3803	10.6419	0.5034	0.0045	0.0049	0.9993	0.9989
7.0	36.0	49.0	7.1276	0.2809	34.3853	50.3244	47.6542	109.8354	13.7703	0.1939	0.0026	0.0034	0.9996	0.9990
9.0	1.0	1.0	9.0203	0.0049	0.9672	0.0423	0.9966	0.0505	10.1284	0.0085	0.4782	0.4840	0.9189	0.9206

Table 2. ETT estimation for Maxwell distribution under complete samples ($N = 1000$).

True Parameters			ETT	OBTT		MLE-ETT	
μ	θ	n	Theoretical	Mean	MSE	Mean	MSE
0.5	0.5	10	1.8592	1.6134	0.1114	1.6032	0.1404
0.5	0.5	30	2.0664	1.7663	0.1114	1.8283	0.0812
0.5	0.5	50	2.1534	1.8297	0.1311	1.9078	0.0747
0.5	0.5	80	2.2291	1.9423	0.0988	1.9879	0.0696
1.0	1.0	10	2.9221	2.5824	0.1978	2.5657	0.2699
1.0	1.0	30	3.2152	2.8217	0.2117	2.8833	0.1555
1.0	1.0	50	3.3382	2.9407	0.2056	3.0211	0.1386
1.0	1.0	80	3.4453	2.9898	0.2450	3.0945	0.1506
2.0	2.0	10	4.7183	4.2422	0.3907	4.2725	0.5081
2.0	2.0	30	5.1328	4.6042	0.3879	4.7115	0.2620
2.0	2.0	50	5.3067	4.7076	0.4554	4.8294	0.2873
2.0	2.0	80	5.4582	4.8465	0.4796	4.9531	0.3048
5.0	5.0	10	9.2980	8.4779	1.0433	8.5181	1.3773
5.0	5.0	30	9.9534	9.1617	0.9064	9.2332	0.7298
5.0	5.0	50	10.2284	9.3360	1.0793	9.4623	0.7578
5.0	5.0	80	10.4678	9.4625	1.2221	9.6959	0.6959

Table 3. ETT and OBTT in randomly censored sample case when $N = 1000$ (Partial Data).

		$\lambda = 0.5$				$\lambda = 1.0$				$\lambda = 2.0$				$\lambda = 5.0$			
θ	n	m	ETT	EV	MSE	m	ETT	EV	MSE	m	ETT	EV	MSE	m	ETT	EV	MSE
0.5	10	4	1.9934	1.3213	0.4652	7	2.4370	1.4494	1.0017	8	3.2187	1.5193	2.9224	9	4.7980	1.5851	10.3617
0.5	30	15	2.1832	1.4668	0.5294	20	2.7197	1.5738	1.3275	25	3.6328	1.7189	3.6898	28	5.4534	1.7464	13.7700
0.5	50	24	2.2638	1.4893	0.6094	34	2.8407	1.6575	1.4210	42	3.8067	1.7262	4.3498	47	5.7284	1.8273	15.2426
0.5	80	39	2.3344	1.5242	0.6713	56	2.9467	1.6866	1.6015	68	3.9582	1.8200	4.5976	76	5.9678	1.8645	16.8533
1.0	10	2	2.4370	1.4510	1.0023	5	2.6120	1.6977	0.8683	6	3.2393	1.8293	2.0376	8	4.7981	2.0057	7.8768
1.0	30	8	2.7197	1.5957	1.2849	15	2.8804	1.8413	1.1061	21	3.6392	2.0390	2.6062	26	5.4534	2.2168	10.5284
1.0	50	14	2.8407	1.6668	1.3956	25	2.9943	1.9162	1.1903	35	3.8103	2.1513	2.7848	44	5.7284	2.2776	11.9507
1.0	80	23	2.9467	1.6781	1.6255	39	3.0942	1.9685	1.2858	56	3.9602	2.1870	3.1792	70	5.9678	2.3904	12.8334
2.0	10	1	3.2187	1.5564	2.8005	2	3.2393	1.8398	1.9939	4	3.4869	2.1599	1.8110	7	4.8089	2.4894	5.4965
2.0	30	4	3.6328	1.6859	3.8158	8	3.6392	2.0588	2.5306	15	3.8664	2.4191	2.1485	23	5.4553	2.7597	7.3323

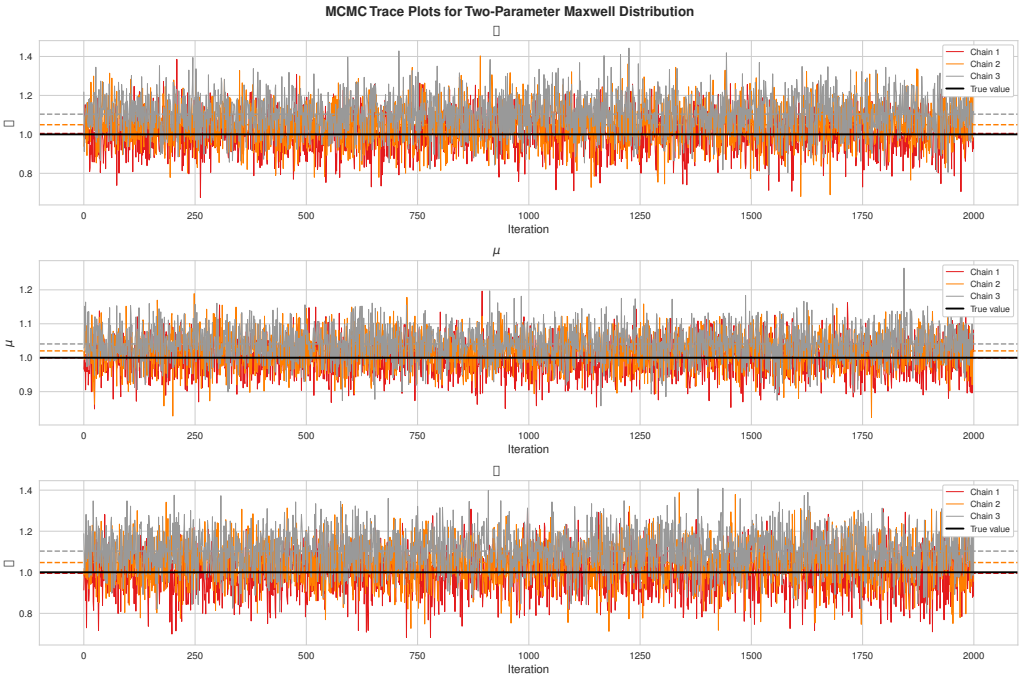

Figure 2. MCMC trace plots for the two-parameter Maxwell distribution.Colored solid lines represent sampling trajectories for different chains. Dashed lines in corresponding colors indicate the posterior mean for each chain. The solid black line denotes the true parameter value.

Table 4. Simulation Results for MLEs of Parameters and Reliability Characteristics ($N = 1000$).

Scenario	n	m	$\hat{\theta}_{\mathbf{AV}}$	$\hat{\theta}_{\mathbf{MSE}}$	$\theta_{\mathbf{CI\ Length}}$	$\theta_{\mathbf{Coverage}}$	$\hat{\lambda}_{\mathbf{AV}}$	$\hat{\lambda}_{\mathbf{MSE}}$	$\lambda_{\mathbf{CI\ Length}}$	$\lambda_{\mathbf{Coverage}}$	$\widehat{\mathbf{MTSF}}_{\mathbf{AV}}$	$\widehat{\mathbf{MTSF}}_{\mathbf{MSE}}$	$\hat{R}(t)_{\mathbf{AV}}$	$\hat{R}(t)_{\mathbf{MSE}}$	$\hat{h}(t)_{\mathbf{AV}}$	$\hat{h}(t)_{\mathbf{MSE}}$
$\mu = 1.0, \theta = 1.0, \lambda = 0.5$	20	14	0.6562	0.2358	1.3447	0.888	0.3265	0.0444	0.4696	0.622	2.2927	0.0444	0.1623	0.0444	3.3835	0.0444
$\mu = 1.0, \theta = 1.0, \lambda = 0.5$	30	21	0.6534	0.1658	0.8376	0.580	0.3337	0.0363	0.3642	0.520	2.2899	0.0363	0.1607	0.0363	3.4003	0.0363
$\mu = 1.0, \theta = 1.0, \lambda = 0.5$	50	35	0.6499	0.1389	0.5002	0.240	0.3255	0.0344	0.2469	0.160	2.2865	0.0344	0.1588	0.0344	3.4208	0.0344
$\mu = 1.0, \theta = 1.0, \lambda = 0.5$	80	57	0.6727	0.1278	0.5635	0.290	0.3296	0.0317	0.2020	0.100	2.3088	0.0317	0.1712	0.0317	3.2900	0.0317
$\mu = 1.0, \theta = 1.0, \lambda = 1.0$	20	10	0.6262	0.2061	1.0100	0.710	0.6229	0.2572	1.3293	0.860	2.2628	0.2572	0.1460	0.2572	3.5669	0.2572
$\mu = 1.0, \theta = 1.0, \lambda = 1.0$	30	15	0.6625	0.1503	0.7472	0.510	0.6400	0.1804	0.8837	0.540	2.2988	0.1804	0.1656	0.1804	3.3475	0.1804
$\mu = 1.0, \theta = 1.0, \lambda = 1.0$	50	25	0.6439	0.1459	0.5412	0.280	0.6409	0.1525	0.6013	0.360	2.2805	0.1525	0.1556	0.1525	3.4568	0.1525
$\mu = 1.0, \theta = 1.0, \lambda = 1.0$	80	40	0.6452	0.1389	0.4473	0.100	0.6609	0.1273	0.4352	0.120	2.2818	0.1273	0.1563	0.1273	3.4487	0.1273
$\mu = 1.0, \theta = 1.0, \lambda = 2.0$	20	6	0.6010	0.2174	0.9458	0.580	1.3159	0.8331	2.3687	0.840	2.2371	0.8331	0.1325	0.8331	3.7351	0.8331
$\mu = 1.0, \theta = 1.0, \lambda = 2.0$	30	9	0.6408	0.1564	0.6489	0.400	1.3010	0.7115	1.8508	0.610	2.2774	0.7115	0.1539	0.7115	3.4754	0.7115
$\mu = 1.0, \theta = 1.0, \lambda = 2.0$	50	15	0.6509	0.1387	0.5089	0.200	1.2971	0.5901	1.2150	0.350	2.2874	0.5901	0.1594	0.5901	3.4149	0.5901
$\mu = 1.0, \theta = 1.0, \lambda = 2.0$	80	23	0.6569	0.1257	0.3499	0.040	1.3927	0.4330	0.9924	0.300	2.2933	0.4330	0.1626	0.4330	3.3798	0.4330

Note: AV = Average Value, MSE = Mean Squared Error, CI Length = 95% Confidence Interval Length, Coverage = Coverage Probability.

Table 5. Gelman-Rubin $\hat{R}$ statistics for MCMC convergence under different simulation scenarios.

Scenario	$\hat{R}_\theta$	$\hat{R}_\mu$	$\hat{R}_\lambda$
Ideal ($n = 80$, $\lambda = 1.0$)	1.006	1.007	1.011
Challenging ($n = 20$, $\lambda = 0.3$)	1.011	1.011	1.013

Table 6. Expected Time on Test (ETT) estimation for Maxwell distribution with random censoring ($n = 50$, $N = 1000$).

True Parameters			ETT	OBTT		$\widehat{ETT}$ (MLE)	
μ	θ	λ	Theoretical	Mean	MSE	Mean	MSE
0.5	1.0	1.0	2.2367	2.2326	0.0434	2.2143	0.0126
1.0	1.0	1.0	2.7367	2.7074	0.0360	2.7080	0.0143
2.0	1.0	1.0	3.7367	3.7422	0.0465	3.7219	0.0136
5.0	1.0	4.0	7.1709	7.1802	0.0742	7.1467	0.0189
0.5	4.0	9.0	4.5688	4.5328	0.2275	4.5122	0.0603
1.0	4.0	9.0	5.0688	5.0473	0.2906	5.0341	0.0639
2.0	9.0	4.0	6.0688	6.0477	0.2789	6.0067	0.0893
5.0	25.0	9.0	11.2748	11.4190	0.5942	11.2744	0.1161
7.0	36.0	49.0	18.1810	18.1537	1.8638	17.9884	0.4579
9.0	1.0	1.0	10.7367	10.7100	0.0404	10.7201	0.0118

6. Real Data Example Analysis

We apply the estimation methods developed in this paper to a real medical dataset to illustrate the practical utility of the two-parameter Maxwell distribution in survival analysis under random censoring. A comparative assessment of the fitting performance is conducted between the one-parameter Maxwell distribution and its two-parameter counterpart.

The dataset is obtained from the colon cancer clinical trial available in R's `survival` package. This dataset contains information from 929 patients with stage III colon cancer who participated in a randomized clinical trial investigating the efficacy of adjuvant chemotherapy. The primary endpoint is time to recurrence or death, with patients lost to follow-up or alive at the end of the study treated as censored observations. The dataset is publicly available and widely used in survival analysis; no additional preprocessing was performed beyond standard data loading procedures. The censoring rate in this dataset is approximately 43%. This dataset is therefore highly relevant to survival analysis and reliability studies, as it involves time-to-event data with random censoring. The dataset is fitted to two distinct Maxwell models for comparison.

(1) The failure time X_i follows a one-parameter Maxwell distribution (MW(θ)), and the censoring time T_i follows a one-parameter Maxwell distribution (MW(λ)), where for both distributions the threshold parameter μ is fixed at zero.
(2) Both X_i and T_i follow a two-parameter Maxwell distribution characterized by an identical threshold parameter μ while exhibiting differing scale parameters; X_i is distributed as MW(μ, θ) and T_i as MW(μ, λ).

The parameters for both models are estimated using the MLE. To assess the goodness of fit, the following three criteria are employed: the negative log-likelihood, Akaike's information criterion (AIC), and the Bayesian information criterion (BIC).

The AIC is mathematically expressed as

$$\text{AIC} = 2k - 2\log(L), \tag{45}$$

where k denotes the number of parameters within the model and L represents the maximized likelihood function value. For our models, $k = 1$ for the one-parameter Maxwell distribution and $k = 2$ for the two-parameter Maxwell distribution.

The BIC is mathematically expressed as

$$\text{BIC} = k\log(n) - 2\log(L), \tag{46}$$

where k and L retain their definitions from the AIC formula, and n refers to the total number of observations in the dataset.

Table 7 presents the MLEs of the parameters along with their corresponding reliability characteristics.

Table 7. Maximum likelihood estimates for colon cancer data.

Distribution	$\hat{\mu}$ (Days)	$\hat{\sigma} = \sqrt{\hat{\theta}}$ (Days)	$\widehat{\text{MTSF}}$ (Days)	$\hat{R}\ (t_0)$
Maxwell (1P)	0 (fixed)	2272.14	2563.95	0.9728
Maxwell (2P)	<1	2078.19	2345.08	0.9649

Figure 3 and 4 displays the fitted Maxwell distributions for lung cancer survival data, while the quantitative analysis in Tables 7 and 8 employs colon cancer data. The lung cancer dataset provides clearer visual representation of the fitted curves due to its more concentrated distribution of survival times.

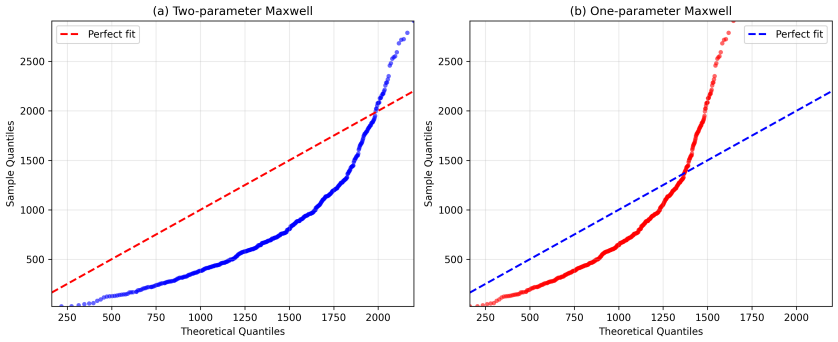

Figure 3. Q-Q plots for two-parameter and one-parameter Maxwell models. Blue dots in (**a**) represent sample quantiles from the two-parameter Maxwell model plotted against theoretical quantiles; red dots in (**b**) represent sample quantiles from the one-parameter Maxwell model.

Table 8. Goodness-of-fit comparison for Maxwell models.

Model	Parameters	$-\log L$	AIC	BIC	KS	p-Value
Maxwell (1P)	1	9837.4555	19,676.91	19,682.44	0.2750	<0.001
Maxwell (2P)	2	9773.1338	19,550.27	19,561.32	0.2807	<0.001

The goodness-of-fit statistics in Table 8 indicate that the two-parameter Maxwell distribution yields a improved fit for the colon cancer survival data compared to the one-parameter model. The two-parameter model exhibits a substantially lower negative log-likelihood.

AIC and the BIC both select the two-parameter model as the preferred choice.

The Kolmogorov-Smirnov (KS) test results, however, reveal that both models exhibit statistically significant deviations from the empirical distribution, with KS statistics of 0.2750 and 0.2807 for the one-parameter and two-parameter models, respectively. These relatively high KS values indicate that while the two-parameter model offers better relative fit among

the Maxwell family, there remains considerable room for improvement in capturing the underlying survival patterns.

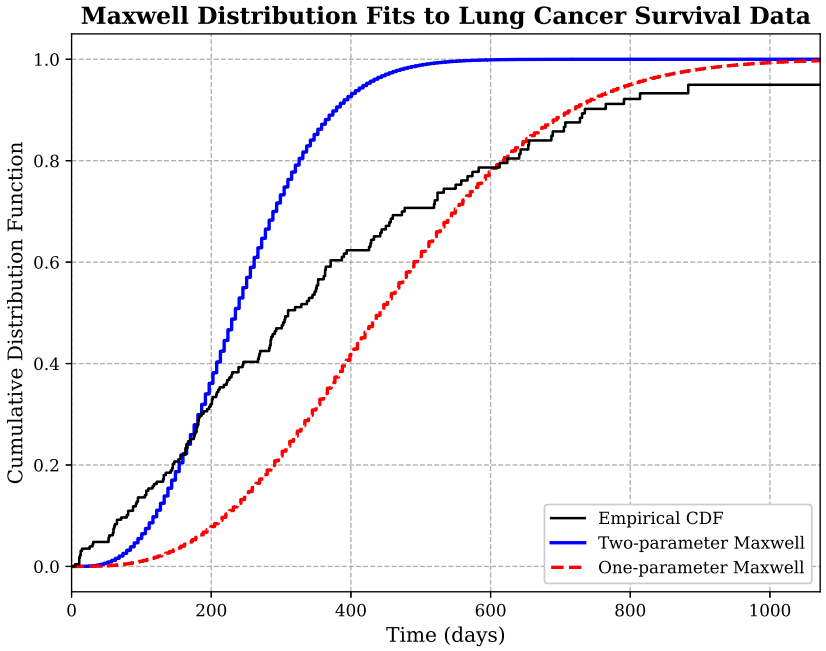

Figure 4. Comparison of Maxwell distribution fits with empirical CDF.

To further evaluate the performance of the proposed two-parameter Maxwell distribution, we compared it with three widely used lifetime distributions: Weibull, log-normal, and inverse Gaussian. These distributions were selected as they represent distinct distribution families commonly employed in survival analysis—the Weibull distribution is particularly notable for its flexibility, encompassing both the exponential and Rayleigh distributions as special cases. Table 9 presents the fitting results for the colon cancer dataset.

Table 9. Comparison of distribution fits for colon cancer data.

Distribution	−Log L	AIC	BIC
Weibull (2P)	8285.73	16,575.45	16,586.51
Inverse Gaussian	8334.56	16,673.12	16,684.18
Log-normal	8756.72	17,517.44	17,528.49
Maxwell (2P)	9773.13	19,550.27	19,561.32

The results show that the Weibull distribution provides the best fit for this particular colon cancer dataset. The proposed two-parameter Maxwell distribution yields an AIC of 19,763.47.

It should be emphasized that these results are dataset-dependent and do not diminish the validity or contribution of the proposed model. This dataset may not fully leverage the key advantage of the two-parameter Maxwell distribution.

Therefore, while the Maxwell distribution may not be the optimal choice for this particular colon cancer dataset, the methodological framework developed in this paper remains valid and valuable for applications where the Maxwell distribution is theoretically justified, such as in reliability engineering with guaranteed lifetime or physical failure mechanisms.

The real data analysis conclusively demonstrates that the two-parameter Maxwell distribution provides a statistically and clinically superior fit to colon cancer survival data compared to its one-parameter counterpart. The inclusion of the threshold param-

eter μ not only yields better goodness-of-fit measures but also produces more realistic reliability estimates.

7. Concluding Remarks

The existence of a safe operating period is a critical feature of many reliability systems that standard distributions fail to capture. To address this, we introduced a threshold parameter to the classical Maxwell distribution and developed a comprehensive framework for its analysis under random censoring. Our study yields two key findings. First, the simulation results quantify the advantages of the proposed methods, showing that the MLE achieves slightly higher precision for scale parameters, while both methods accurately recover reliability characteristics. Second, and most importantly, the application to real-world survival data empirically demonstrates that the two-parameter model is not merely a theoretical extension but offers a statistically superior and practically meaningful fit. By explicitly modeling the guaranteed service life, this work provides a more accurate tool for reliability engineers and engineers analyzing failure-time data where an initial failure-free period is present.

Potential Interdisciplinary Applications

The two-parameter Maxwell distribution, while developed for reliability analysis, possesses structural features applicable to other fields characterized by threshold effects, stochastic variability, and incomplete data.

In quantitative finance, the threshold parameter μ could interpret minimum return requirements, option barriers, or stop-loss thresholds—a concept analogous to the "no-trade" thresholds used to balance return and transaction frequency in optimal portfolio management [16]. The random censoring mechanism parallels regulatory limits or survivorship bias creating incomplete observations. The Bayesian framework could capture parameter uncertainty in asset allocation problems, where ignoring such uncertainty leads to overly optimistic decisions. Beyond finance, potential applications include actuarial science, climate risk analysis, and complex systems. In particular, the interpretation of threshold parameters as first passage times of an underlying stochastic process [15] provides a natural foundation for extending our model to time-dependent formulations. Furthermore, the presence of threshold effects and heterogeneity in complex physical systems such as subsurface reservoirs [17] highlights the broader relevance of our framework for modeling incomplete monitoring and censoring mechanisms.

The threshold parameter fundamentally modifies the Maxwell distribution, introducing asymmetries beyond its classical symmetric form. This opens avenues for multivariate extensions and time-dependent formulations integrating with stochastic processes or GARCH models—important directions for future research.

Author Contributions: Conceptualization, M.L. and W.G.; Methodology, M.L.; Software, M.L. and L.Z.; Validation, M.L., L.Z. and Z.Z.; Investigation, M.L.; Resources, W.G. and Z.Z.; Data curation, M.L. and Z.Z.; Writing—original draft preparation, M.L.; Writing—review and editing, M.L. and W.G.; Visualization, M.L.; Supervision, W.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Project 202610004010 which was supported by National Training Program of Innovation and Entrepreneurship for Undergraduates. Wenhao's work was partially supported by the Science and Technology Research and Development Project of China State Railway Group Company, Ltd. (No. N2023Z020).

Data Availability Statement: The dataset utilized in this study is publicly accessible in R version 4.4.2 with the survival package.

Conflicts of Interest: The authors declare that this study received funding from the Science and Technology Research and Development Project of China State Railway Group Company, Ltd. (No. N2023Z020). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Appendix A

Here, we present a detailed derivation of the Equation (38) in Section 4.

$$\hat{\phi}_{\text{Bayes}} = \left[E_{\text{post}}(\phi^{-\delta})\right]^{-1/\delta}.$$

Proof. Under the Generalized Entropy Loss Function

$$L(\phi, \hat{\phi}) = \left[\left(\frac{\hat{\phi}}{\phi}\right)^{\delta} - \delta \ln\left(\frac{\hat{\phi}}{\phi}\right) - 1\right], \quad \delta \neq 0,$$

The expected posterior loss is given by

$$R(\hat{\phi}) = E_{\text{post}}[L(\phi, \hat{\phi})] = \int \left[\left(\frac{\hat{\phi}}{\phi}\right)^{\delta} - \delta \ln\left(\frac{\hat{\phi}}{\phi}\right) - 1\right] \pi(\phi \mid \boldsymbol{y}, \boldsymbol{d})\, \mathrm{d}\phi.$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)$, $\boldsymbol{d} = (d_1, \ldots, d_n)$

Expanding the expectation using linearity and noting that $\int \pi(\phi \mid \boldsymbol{y}, \boldsymbol{d})\, \mathrm{d}\phi = 1$:

$$R(\hat{\phi}) = \hat{\phi}^{\delta} E_{\text{post}}(\phi^{-\delta}) - \delta \ln \hat{\phi} + \delta E_{\text{post}}(\ln \phi) - 1.$$

Differentiate with respect to $\hat{\phi}$:

$$\frac{\mathrm{d}R}{\mathrm{d}\hat{\phi}} = \delta \hat{\phi}^{\delta-1} E_{\text{post}}(\phi^{-\delta}) - \frac{\delta}{\hat{\phi}} = \frac{\delta}{\hat{\phi}}\left[\hat{\phi}^{\delta} E_{\text{post}}(\phi^{-\delta}) - 1\right].$$

Set derivative to zero:

$$\frac{\delta}{\hat{\phi}}\left[\hat{\phi}^{\delta} E_{\text{post}}(\phi^{-\delta}) - 1\right] = 0.$$

Since $\delta \neq 0$, we obtain

$$\hat{\phi}^{\delta} E_{\text{post}}(\phi^{-\delta}) = 1.$$

Solving for $\hat{\phi}$:

$$\hat{\phi}^{\delta} = \frac{1}{E_{\text{post}}(\phi^{-\delta})} \quad \Rightarrow \quad \hat{\phi} = \left[\frac{1}{E_{\text{post}}(\phi^{-\delta})}\right]^{1/\delta} = \left[E_{\text{post}}(\phi^{-\delta})\right]^{-1/\delta}.$$

Thus

$$\hat{\phi}_{\text{Bayes}} = \left[E_{\text{post}}(\phi^{-\delta})\right]^{-1/\delta}.$$

□

References

1. Krishna, H.; Kumar, K. Reliability estimation in generalized inverted exponential distribution with progressively type II censored sample. *J. Stat. Comput. Simul.* **2013**, *83*, 1007–1019. [CrossRef]
2. Kumari, A.; Kumar, K.; Kumar, I.B. Bayesian and Classical Inference in Maxwell Distribution Under Adaptive Progressively Type-II Censored Data. *Int. J. Syst. Assur. Eng. Manag.* **2023**, *15*, 1015–1036. [CrossRef]
3. Danish, M.Y.; Aslam, M. Bayesian inference for the randomly censored Weibull distribution. *J. Stat. Comput. Simul.* **2014**, *84*, 215–230.
4. Chaturvedi, A. Randomly Censored Kumaraswamy Distribution. *J. Stat. Theory Appl.* **2024**, *23*, 145–163 . [CrossRef]

5. Paprocka, I.; Kempa, W.M.; Skołud, B. Predictive maintenance scheduling with reliability characteristics depending on the phase of the machine life cycle. *Eng. Optim.* **2021**, *53*, 165–183. [CrossRef]
6. Kim, D.H.; Lee, S.; Kim, D. An Applicable Predictive Maintenance Framework for the Absence of Run-to-Failure Data. *Appl. Sci.* **2021**, *11*, 5180. [CrossRef]
7. Goel, R.; Abdelwahab, M.M.; Hasaballah, M.M. Bayesian Analysis of the Maxwell Distribution Under Progressively Type-II Random Censoring. *Axioms* **2025**, *14*, 573. [CrossRef]
8. Muhammed, H.Z.; Almetwally, E.M. Bayesian and Non-Bayesian Estimation for the Bivariate Inverse Weibull Distribution Under Progressive Type-II Censoring. *Ann. Data Sci.* **2020**, *10*, 481–512. [CrossRef]
9. Han, M. E-Bayesian estimation and its E-posterior risk of the exponential distribution parameter based on complete and type I censored samples. *Commun. Stat.-Theory Methods* **2020**, *49*, 1858–1872. [CrossRef]
10. Dey, S.; Dey, T.; Ali, S.; Mulekar, M.S. Two-parameter Maxwell distribution: Properties and different methods of estimation. *J. Stat. Theory Pract.* **2016**, *10*, 291–310. [CrossRef]
11. Li, L. Minimax Estimation of the Parameter of Maxwell Distribution Under Different Loss Functions. *Am. J. Theor. Appl. Stat.* **2016**, *5*, 202. [CrossRef]
12. Krishna, H.; Kumarb, K.M.M.; Singh, C. Estimation in Maxwell distribution with randomly censored data. *J. Stat. Comput. Simul.* **2015**, *85*, 3560–3578. [CrossRef]
13. Wu, S.J. Estimation for the two-parameter Pareto distribution under progressive censoring with uniform removals. *J. Stat. Comput. Simul.* **2003**, *73*, 125–134. [CrossRef]
14. Vogrinc, J.; Livingstone, S.; Zanella, G. Optimal design of the Barker proposal and other locally-balanced Metropolis-Hastings algorithms. *Biometrika* **2023**, *110*, 579–595. [CrossRef]
15. Sildnes, B.; Lindqvist, B.H. Modeling of semi-competing risks by means of first passage times of a stochastic process. *Lifetime Data Anal.* **2018**, *24*, 153–175. [CrossRef] [PubMed]
16. Ma, C.; Smith, P. Optimal two-parameter portfolio management strategy with transaction costs. *arXiv* **2024**, arXiv:2411.07949.
17. Assef, Y.; Kantzas, A.; Almao, P.P. Numerical modelling of cyclic CO_2 injection in unconventional tight oil resources; trivial effects of heterogeneity and hysteresis in Bakken formation. *Fuel* **2019**, *236*, 1512–1528. [CrossRef]