Pak. J. Statist.
2026 Vol. 42(3), 265-297

MAXIMUM LIKELIHOOD, BAYESIAN INFERENCE AND PREDICTION FOR THE X-GAMMA DISTRIBUTION WITH RIGHT-CENSORED SURVIVAL DATA

Saraiva, E.F.[1§] Vigas, V.P.[1] Silva, N.H.[2] and **Suzuki, A.K.[2]**
[1] Institute of Mathematics, Federal University of Mato Grosso do Sul
Caixa, Postal 549, CEP 79070-900, Camp o Grande, MS, Brazil
[2] Department of Applied Mathematics and Statistics,
University of São Paulo - São Carlos, SP, Brazil
[§] Corresponding author Email: erlandson.saraiva@ufms.br

ABSTRACT

This paper investigates the use of the XGamma distribution for modelling right-censored survival data. Parameter estimation is performed using the maximum likelihood approach, for which we establish existence and uniqueness results. In addition, Bayesian estimators are derived under squared error, LINEX, and generalized entropy loss functions. To obtain Bayesian estimates, we employ importance sampling and two Markov chain Monte Carlo schemes: independent Metropolis–Hastings and random-walk Metropolis. We further develop a procedure for predictive inference, allowing the construction of predictive distributions and intervals for future observations. The performance of the proposed estimators and predictive approach is assessed through a comprehensive simulation study under varying sample sizes and percentages of right-censored observations. The methodology is also illustrated using two real datasets. The results indicate that the proposed estimators provide accurate and stable inference across a wide range of scenarios, while the predictive procedure yields reliable and well-calibrated predictive intervals.

KEYWORDS

XGamma distribution, Right-censored data, Maximum likelihood estimation, Bayesian Inference, Importance sampling, Metropolis-Hastings algorithm.

1. INTRODUCTION

In recent years, mixture models have been successfully applied for modelling datasets arising from heterogeneous populations in various fields. In survival analysis, such models are particularly useful, as they allow the modelling of sub-populations with different risks of failure. Moreover, by combining different probability distributions, mixture models can better capture the variability in survival times compared to a single distribution. In this context, Mohammed et al. (2014) proposed a mixture of exponential, gamma, and Weibull distributions for heterogeneous survival data; Assane et al. (2018) developed a Bayesian significance test to discriminate between survival distributions via mixture models; and Li et al. (2022) introduced a nonparametric Erlang mixture model with time-dependent weights.

© 2026 Pakistan Journal of Statistics

A well-known example of a finite mixture distribution is the Lindley distribution (Lindley, 1958; Ghitany et al. 2008), which can be expressed as a mixture of an exponential distribution with parameter θ and a gamma distribution with parameter vector $(2, \theta)$. Motivated by this representation, Sen et al. (2016) introduced the XGamma distribution by replacing the gamma component with a gamma distribution having parameters $(3, \theta)$. The authors derived several distributional properties and proposed estimation procedures based on the method of moments and maximum likelihood for complete data.

However, in practical survival analysis, datasets frequently contain censored observations, i.e., observations for which the exact failure time is unknown but is known to lie within a given interval or beyond a certain point. As highlighted by Crowther and Lambert (2019), right-censored data are among the most common types encountered in survival studies.

Despite the growing interest in the XGamma distribution as a flexible lifetime model, inferential procedures for this distribution under censoring have received little attention. To the best of our knowledge, no study has developed a comprehensive inferential framework for the XGamma distribution in the presence of right-censored observations. In particular, maximum likelihood estimation, Bayesian inference under different loss functions, and predictive inference for censored XGamma data have not yet been investigated. Motivated by this gap, we study inference for the parameter θ of the XGamma distribution under right censoring and develop both classical and Bayesian inferential procedures, together with predictive methods for future observations.

We begin by studying the maximum likelihood (ML) estimator of θ. Since the associated score function does not admit a closed-form solution, numerical methods are required to obtain the estimate. However, numerical optimization alone does not guarantee the existence and uniqueness of the ML estimator. Therefore, we establish theoretical results ensuring these properties.

We also develop a Bayesian inferential framework for θ under the squared error, LINEX, and generalized entropy loss functions. The latter two loss functions are particularly appealing because, unlike the squared error loss, they are asymmetric and allow different penalties to be assigned to overestimation and underestimation. The LINEX loss function, introduced by Varian (1975) and further developed by Zellner (1986), and the generalized entropy loss function proposed by Calabria and Pulcini (1996) have been widely employed in Bayesian reliability and lifetime-data analysis, particularly in situations where estimation errors of opposite signs entail different practical consequences (Basu and Ebrahimi, 1991). Recent applications include Bayesian estimation for the generalized exponential distribution (Janeen, 2005) and Bayesian inference for the Weibull distribution under asymmetric loss functions (Singh and Singh, 2013). For inference on θ, Bayesian estimation is implemented using importance sampling (IS), an independent Metropolis–Hastings (IMH) algorithm, and a random-walk Metropolis (RWM) algorithm (Hastings, 1970, Kloek and van Dijk, 1978; Tierney, 1994; Srinivasan, 2002).

Since $\theta > 0$, we illustrate the Bayesian methodology using a Gamma prior distribution, a natural and widely adopted choice for positive parameters. Nevertheless, the proposed inferential framework is not restricted to this prior specification and can be readily adapted

to alternative prior distributions. In addition, we introduce a numerical procedure for calibrating the parameters a and b associated with the LINEX and generalized entropy loss functions, respectively. These quantities should be interpreted as tuning parameters controlling the asymmetry of the corresponding loss functions rather than as model parameters subject to Bayesian inference.

Overall, ten estimators for θ are considered: the ML estimator, three Bayesian estimators obtained via IS, three Bayesian estimators obtained via IMH, and three Bayesian estimators obtained via RWM. We also develop a predictive inference framework that enables the construction of predictive distributions and interval estimates for future observations.

To assess the performance of the proposed methods, we conduct a simulation study considering different sample sizes and proportions of right-censored observations. The comparison is based on the mean absolute error (MAE) and the root mean squared error (RMSE) with respect to the Kaplan–Meier estimator of the survival function. Additionally, two real datasets are analyzed to illustrate the practical usefulness of the XGamma distribution for modelling right-censored survival data. To evaluate predictive performance in the real-data applications, we implement a leave-one-out cross-validation procedure.

The main contributions of this paper are threefold. First, we develop a comprehensive inferential framework for the XGamma distribution under right censoring, including maximum likelihood estimation and theoretical results guaranteeing the existence and uniqueness of the ML estimator. Second, we derive Bayesian estimators under squared error, LINEX, and generalized entropy loss functions, and implement them using importance sampling, independent Metropolis–Hastings, and random-walk Metropolis algorithms. Third, we propose a predictive inference framework for constructing predictive distributions and prediction intervals for future survival times. The effectiveness of the proposed methodologies is demonstrated through extensive simulation studies and applications to two real datasets.

The remainder of the paper is organized as follows. Section 2 introduces the XGamma distribution and develops the maximum likelihood inference framework, including theoretical results guaranteeing the existence and uniqueness of the ML estimator. Section 3 presents the Bayesian methodology and the estimation procedures based on importance sampling, independent Metropolis–Hastings, and random-walk Metropolis algorithms. Section 4 develops the predictive inference framework. Section 5 reports the results of the simulation studies for parameter estimation and predictive inference, respectively. Section 6 illustrates the proposed methodology through applications to two real datasets. Finally, Section 7 concludes the paper with some final remarks.

2. THE XGAMMA DISTRIBUTION

Let T be a continuous random variable defined on $\mathbb{R}^+$, whose probability density function is denoted by $f(t|\theta)$, for $\theta \in \Theta = (0, +\infty)$. In this context, Sen et al. (2016) introduced the XGamma probability distribution as a mixture model with two components, given by

$$f(t|\theta) = w_1 f_1(t|\theta) + (1 - w_1) f_2(t|\theta),$$

where θ is the parameter of the model, and w_1 and $(1 - w_1)$ are the mixture weights. Then, the authors set up:

- the first component $f_1(t|\theta)$ as the probability density function of an exponential distribution indexed by the parameter θ, i.e., $f_1(t|\theta) = \theta e^{-\theta t}$;
- the second component as the probability density function of a Gamma distribution indexed by the vector of parameters $(3, \theta)$, i.e., $f_2(t|\theta) = \frac{\theta^3}{2} t^2 e^{-\theta t}$; and
- the weight $w_1 = \frac{\theta}{1+\theta}$, for $\theta \in \Theta$.

From this representation, Sen et al. (2016) showed that the probability density function (pdf) of the XGamma distribution can be expressed as:

$$f(t|\theta) = \frac{\theta^2}{1+\theta}\left(1 + \frac{\theta}{2}t^2\right)e^{-\theta t}, \tag{1}$$

for $t > 0$ and $\theta > 0$. Since the model is indexed by a single parameter, identifiability follows directly from the fact that distinct values of θ induce distinct probability distributions.

The survival function $S(t|\theta)$ is given by:

$$S(t|\theta) = \frac{1 + \theta + \theta t + (\theta^2 t^2/2)}{(1+\theta)}e^{-\theta t} \tag{2}$$

for $t > 0$ and $\theta > 0$. For further details see Sen et al. (2016).

2.1 Maximum Likelihood Estimation in the Presence of Right-Censored Data

Let T be a continuous random variable representing the lifetime of an individual or item, for $t > 0$. Suppose that T follows an XGamma distribution with parameter $\theta > 0$. Furthermore, consider that the lifetimes of n individuals are recorded, with some of these observations being right-censored.

Let $\boldsymbol{y} = (y_1, \dots, y_n)$ denote the n recorded observations, where $y_i = (t_i, \delta_i)$, with t_i representing the observed time and δ_i the censoring indicator defined as:

$$\delta_i = \begin{cases} 1, & \textit{if the lifetime is uncensored, i.e.,} T_i = t_i, \\ 0, & \textit{if the lifetime is censored, i.e.,} T_i > t_i, \end{cases} i = 1, \dots, n.$$

Under this setting, the likelihood function for θ is given by

$$\begin{aligned} \mathcal{L}(\theta \mid \boldsymbol{y}) &= \prod_{i=1}^{n} [f(y_i \mid \theta)]^{\delta_i} [S(t_i \mid \theta)]^{1-\delta_i} \\ &= \frac{\theta^{2r}}{(1+\theta)^n} \left[\prod_{i=1}^{n} \left(1 + \frac{\theta}{2}t_i^2\right)^{\delta_i} \left(1 + \theta + \theta t_i + \frac{\theta^2}{2}t_i^2\right)^{1-\delta_i}\right] exp\left\{-\theta \sum_{i=1}^{n} t_i\right\} \end{aligned} \tag{3}$$

where $r = \sum_{i=1}^{n} \delta_i$ denotes the number of uncensored observations. Please, see Lawless (2003) and Balakrishnan and Aggarwala (2000) for a concise explanation about the expression $\mathcal{L}(\theta|\boldsymbol{y}) = \prod_{i=1}^{n} [f(y_i|\theta)]^{\delta_i} [S(t_i|\theta)]^{1-\delta_i}$.

The log-likelihood function is given by

$$\ell(\theta|\boldsymbol{y}) = ln\big(\mathcal{L}(\theta|\boldsymbol{y})\big) = 2rln(\theta) - nln(1+\theta) + \Psi_1(\theta) + \Psi_2(\theta) - \theta\sum_{i=1}^{n} t_i,$$

where $\Psi_1(\theta) = \sum_{i=1}^{n} \delta_i ln\left(1+\frac{\theta}{2}t_i^2\right)$ and $\Psi_2(\theta) = \sum_{i=1}^{n}(1-\delta_i)ln\left(1+\theta+\theta t_i+\frac{\theta^2}{2}t_i^2\right)$.

The score function is given by:

$$U(\theta|\boldsymbol{y}) = \frac{d\ell(\theta|\boldsymbol{y})}{d\theta} = \frac{2r}{\theta} - \frac{n}{1+\theta} + {\Psi_1}'(\theta) + {\Psi_2}'(\theta) - \sum_{i=1}^{n} t_i, \tag{4}$$

where ${\Psi_1}'(\theta) = \frac{d\Psi_1(\theta)}{d\theta} = \sum_{i=1}^{n}\frac{\delta_i t_i^2}{2+\theta t_i^2}$ and ${\Psi_2}'(\theta) = \frac{d\Psi_2(\theta)}{d\theta} = \sum_{i=1}^{n}\frac{(1-\delta_i)(1+t_i+\theta t_i^2)}{1+\theta+\theta t_i+(\theta^2 t_i^2/2)}$.

The maximum likelihood estimator $\hat{\theta}_{ML}$ of θ is obtained as the solution of the score equation $U(\theta|\boldsymbol{y}) = 0$. However, since $U(\theta|\boldsymbol{y}) = 0$ is non-linear and does not admit a closed-form solution, numerical optimization methods are required. In practice, algorithms such as the Nelder-Mead algorithm (Nelder and Meal, 1965) or the Quasi-Newton method (Fletcher and Reeves, 1965) can be implemented via the optim() command in the R software (R Core Team, 2026).

2.1.1 Existence and Uniqueness of the ML Estimator

Since numerical methods used to obtain the maximum likelihood estimate do not provide information about the existence and uniqueness of the solution of the score equation $U(\theta|\boldsymbol{y}) = 0$, we establish these properties in this section.

Proposition 1: If $r \geq 1$, then the score equation $U(\theta|\boldsymbol{y}) = 0$, for $\theta \in (0,\infty)$, admits a unique solution, denoted by $\hat{\theta}_{ML}$.

Proof. From Equation (4), as $\theta \to 0^+$

$$\frac{2r}{\theta} \to \infty(\text{whenever } r \geq 1), \frac{n}{1+\theta} \to n, \Psi_1^{\prime(\theta)} \to \frac{1}{2}\sum_{i=1}^{n}\delta_i t_i^2,$$
$$\text{and } {\Psi_2}'(\theta) \to \sum_{i=1}^{n}(1-\delta_i)(1+t_i).$$

Therefore, $\lim_{\theta\to 0^+} U(\theta|\boldsymbol{y}) = \infty$. On the other hand, as $\theta \to \infty$

$$\frac{2r}{\theta} \to 0, {\Psi_1}'(\theta) \to 0, \frac{n}{1+\theta} \to 0 \text{ and } \lim_{\theta\to\infty}\psi_2{}'(\theta) = \lim_{\theta\to\infty}\sum_{i=1}^{n}\frac{(1-\delta_i)t_i^2}{1+t_i+\theta t_i^2} = 0.$$

Hence, $\lim_{\theta\to\infty} U(\theta|\boldsymbol{y}) = -\sum_{i=1}^{n} t_i < 0$. Since $U(\theta|\boldsymbol{y})$ is continuous with

$$\lim_{\theta\to 0^+} U(\theta|\boldsymbol{y}) = \infty \text{ and } \lim_{\theta\to\infty} U(\theta|\boldsymbol{y}) = -\sum_{i=1}^{n} t_i < 0,$$

the Intermediate Value Theorem guarantees the existence of at least one root $\hat{\theta}_{ML} \in (0, \infty)$ such that $U(\hat{\theta}_{ML}|\boldsymbol{y}) = 0$. To establish uniqueness, we consider the second derivative of the log-likelihood function $\frac{dU(\theta|y)}{d\theta} = \frac{d^2\ell(\theta|y)}{\partial\theta^2}$, given by:

$$\frac{d^2\ell(\theta|\boldsymbol{y})}{d\theta^2} = -\frac{2r}{\theta^2} + \frac{n}{(1+\theta)^2} - \sum_{i=1}^{n} \frac{\delta_i t_i^4}{(2+\theta t_i^2)^2} - \sum_{i=1}^{n} \frac{(1-\delta_i)(1+2t_i+\theta t_i^2+\theta t_i^3+\theta^2 t_i^4/2)}{(1+\theta+\theta t_i+\theta^2 t_i^2/2)^2} \quad (5)$$

By rewriting the first two terms from (5) as

$$-\frac{2r}{\theta^2} + \frac{n}{(1+\theta)^2} = \sum_{i=1}^{n} \delta_i \left(-\frac{2}{\theta^2} + \frac{1}{(1+\theta)^2}\right) + \sum_{i=1}^{n} (1-\delta_i) \frac{1}{(1+\theta)^2}$$

and combining them into the summations, we obtain

$$\frac{d^2\ell(\theta|\boldsymbol{y})}{d\theta^2} = \sum_{i=1}^{n} \delta_i A_i + \sum_{i=1}^{n} (1-\delta_i) B_i,$$

where

$$A_i = \left[-\frac{2}{\theta^2} + \frac{1}{(1+\theta)^2} - \frac{t_i^4}{(2+\theta t_i^2)^2}\right]$$

$$B_i = \left[\frac{1}{(1+\theta)^2} - \frac{(1+2t_i+\theta t_i^2+\theta t_i^3+\theta^2 t_i^4/2)}{(1+\theta+\theta t_i+\theta^2 t_i^2/2)^2}\right].$$

Now, A_i can be easily seen to be strictly negative for all $1 \le i \le n$, since

$2/\theta^2 > (1+\theta)^{-2}$.

For B_i, we can see it will be strictly negative if, and only if,

$$(1+\theta+\theta t_i+\theta^2 t_i^2/2)^2 - (1+\theta)^2(1+2t_i+\theta t_i^2+\theta t_i^3+\theta^2 t_i^4/2) < 0.$$

Expanding the left side of the above inequality, we obtain

$$-2[1+\theta]t_i - \theta t_i^2 - \theta[1+2\theta]t_i^3 - \theta^2\left[\frac{1}{2} + \theta + \frac{\theta^2}{4}\right]t_i^4,$$

which is always negative. Then, $B_i < 0$ for all $1 \le i \le n$ and $\frac{d^2\ell(\theta|y)}{d\theta^2} < 0$. Thus, the log-likelihood function is strictly concave, implying that the score equation $U(\theta|\boldsymbol{y}) = 0$ has a unique root whenever $r \ge 1$, denoted by $\hat{\theta}_{ML}$. This completes the proof. □

In the above proposition, we note that if $r = 0$, corresponding to the case where all observations are censored, then

$$\lim_{\theta \to 0^+} U(\theta \mid \boldsymbol{y}) = 0^-.$$

Consequently, the log-likelihood function $\ell(\theta \mid \boldsymbol{y})$ is strictly decreasing on $(0, \infty)$. Therefore, its supremum is attained at the boundary point $\theta = 0$, which does not belong to the parameter space. Hence, no maximum likelihood estimator exists in the interior of the parameter space when $r = 0$.

The observed Fisher information is given by

$$J(\hat{\theta}_{ML}) = -\frac{d^2\ell(\theta|\boldsymbol{y})}{d\theta^2}\bigg|_{\theta=\hat{\theta}_{ML}}.$$

An approximate $100(1-\alpha)\%$ Wald-type confidence interval for θ is

$$\hat{\theta}_{ML} \pm z_{1-\frac{\alpha}{2}}\sqrt{J(\hat{\theta}_{ML})^{-1}},$$

where $z_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the standard normal distribution, for $\alpha \in (0,1)$.

3. BAYESIAN APPROACH

To develop the Bayesian approach, we need to specify a prior distribution for the parameter θ. In a general setting, let $\pi(\theta|\psi)$ denote the prior density for θ, where ψ represents a set of known hyperparameters. By Bayes' theorem, the posterior distribution of θ is given by

$$\pi(\theta|\boldsymbol{y},\psi) \propto \mathcal{L}(\theta|\boldsymbol{y})\pi(\theta|\psi),$$

where $\mathcal{L}(\theta|\boldsymbol{y})$ is the likelihood function given in expression (3).

Let $L(\theta, d)$ be a function that determines the loss resulting of making a decision d when the true value is θ. Then, the posterior risk is given by

$$R(\theta, d|\boldsymbol{y}) = \mathbb{E}(L(\theta, d)|\boldsymbol{y}) = \int_0^\infty L(\theta, d)\pi(\theta|\boldsymbol{y},\psi)d\theta.$$

The Bayes estimator is given by $\hat{\theta}_{BE} = \underset{d}{argmin} R(\theta, d|\boldsymbol{y})$. That is, the choice of loss function plays a crucial role in Bayesian inference.

In this article, we consider the following three types of loss functions:

1) The squared error loss is defined by: $L(\theta, d|\boldsymbol{y}) = \left(\hat{\theta} - \theta\right)^2$. This loss function is symmetric, assigning equal weight to overestimation and underestimation. The corresponding Bayes estimator is $\hat{\theta}_{SE} = \mathbb{E}(\theta|\boldsymbol{y})$, i.e., the posterior expected value.
2) The Linex loss function (LI) is defined as: $L(\theta, d|\boldsymbol{y}) = e^{a(\hat{\theta}-\theta)} - a\left(\hat{\theta} - \theta\right)$, for $a \neq 0$. This loss is asymmetric and allows different penalties for overestimation and underestimation. The Bayes estimator is $\hat{\theta}_{LI} = -\frac{1}{a} ln\left(\mathbb{E}\left(e^{-a\theta}|\boldsymbol{y}\right)\right)$. This loss function was originally proposed by Varian (1975) and popularized by Zellner (1986) as an alternative to the squared error loss function.

3) The Generalized Entropy loss function is defined by: $L(\theta, d|\boldsymbol{y}) = \left(\frac{\hat{\theta}}{\theta}\right)^{-b} - bln\left(\frac{\hat{\theta}}{\theta}\right) - 1$, for $b \neq 0$. This loss function is also asymmetric. The corresponding Bayes estimator is $\hat{\theta}_{GE} = \left(\mathbb{E}(\theta^{-b}|\boldsymbol{y})\right)^{-\frac{1}{b}}$, for $b \neq 0$. For further details on GE loss function, see Dey et al. (1987) and Day and Liu (1992).

The derivations of the Bayes estimators and the corresponding posterior risks under the LINEX and generalized entropy loss functions are provided in Appendix D of the Supplementary Material. The choice of the tuning parameters a and b in the LINEX and generalized entropy loss functions, respectively, can substantially affect the resulting Bayesian estimators $\hat{\theta}_{LI}$ and $\tilde{\theta}_{GE}$. In practice, the specification of these parameters depends on the application at hand and on the relative costs associated with overestimation and underestimation. At this stage of the analysis, we leave their values unspecified and, in the next section, introduce a data-driven numerical procedure for calibrating a and b.

3.1 Bayesian Estimation

We now consider a particular case by specifying a prior distribution for θ. Since $\theta \in (0, +\infty)$, a natural choice for the prior distribution is the Gamma distribution. Throughout this paper, we use the shape–rate parametrization and write $\theta \mid \gamma, \beta \sim \Gamma(\gamma, \beta)$, where $\gamma > 0$ and $\beta > 0$ denote the shape and rate parameters, respectively. That is, we adopt the parametrization for which $\mathbb{E}(\theta) = \frac{\gamma}{\beta}$ and $\mathbb{V}ar(\theta) = \frac{\gamma}{\beta^2}$.

Applying Bayes' theorem, the posterior density of θ is given by

$$\pi(\theta \mid y, \gamma, \beta) \propto \mathbb{A}(\theta) \cdot \theta^{\gamma+2r-1} exp\left\{-\theta\left(\beta + \sum_{i=1}^{n} t_i\right)\right\}, \tag{6}$$

where

$$\mathbb{A}(\theta) = \frac{1}{(1+\theta)^n} \prod_{i=1}^{n} \left[\left(1 + \frac{\theta t_i^2}{2}\right)^{\delta_i} \left(1 + \theta + \theta t_i + \frac{\theta^2 t_i^2}{2}\right)^{1-\delta_i}\right].$$

Hence, the posterior density does not belong to a standard parametric family because of the correction factor $\mathbb{A}(\theta)$. Nevertheless, the kernel

$$\theta^{\gamma+2r-1} exp\left\{-\theta\left(\beta + \sum_{i=1}^{n} t_i\right)\right\}$$

coincides with the kernel of a Gamma distribution with parameters $\gamma^* = \gamma + 2r$ and $\beta^* = \beta + \sum_{i=1}^{n} t_i$. This representation will be exploited in Sections 3.1.1 and 3.1.2 to construct a Gamma distribution with parameters γ^* and β^*, which is used as the proposal distribution in the importance sampling procedure and as the candidate generating distribution in the independent Metropolis–Hastings algorithm.

Since $\gamma^* > 0$ and $\beta^* > 0$, for any measurable function $\kappa(\theta): \Theta \to \mathbb{R}$ such that the expectation exists, the posterior expectation is given by

$$\mathbb{E}(\kappa(\theta)|\boldsymbol{y}) = \int_{\Theta} \kappa(\theta)\pi(\theta|\boldsymbol{y},\gamma,\beta)d\theta = \frac{\int_{\Theta} \kappa(\theta)\mathbb{A}(\theta)g(\theta|\boldsymbol{y},\gamma,\beta)d\theta}{\int_{\Theta} \mathbb{A}(\theta)g(\theta|\boldsymbol{y},\gamma,\beta)d\theta}. \quad (7)$$

However, the integrals in both the numerator and denominator do not admit closed-form expressions. Consequently, posterior quantities such as $\mathbb{E}[\kappa(\theta) \mid \boldsymbol{y}]$ cannot be computed analytically. Moreover, direct sampling from $\pi(\theta \mid \boldsymbol{y},\gamma,\beta)$ is not feasible. To address this issue, we employ simulation-based methods to approximate posterior expectations. In particular, we consider the use of Importance Sampling and the Metropolis-Hastings algorithm, which allow us to generate samples from the posterior distribution indirectly and obtain Monte Carlo estimates of the quantities of interest.

3.1.1 Importance Sampling

The importance sampling (IS) algorithm is a Monte Carlo method used to approximate expectations with respect to a target distribution by sampling from an alternative distribution that is easier to generate values. Thus, to obtain an approximation to the expected value given in Equation (7) using the IS algorithm, let $q(\theta)$ be a probability density function defined on the same support set as the posterior distribution $\pi(\theta|\boldsymbol{y},\gamma,\beta)$and from which sampling is straightforward. The density $q(\theta)$ is named as the importance function or proposal density. Then, the posterior expectation in (7) can be rewritten as:

$$\mathbb{E}(\kappa(\theta)|\boldsymbol{y}) = \frac{1}{Z_{\boldsymbol{y}}}\int_{\Theta} \frac{\kappa(\theta)\mathbb{A}(\theta)g(\theta|\boldsymbol{y},\gamma,\beta)}{q(\theta)}q(\theta)d\theta = \frac{1}{Z_{\boldsymbol{y}}}\mathbb{E}_q[\mathbb{W}(\theta)],$$

where, $\mathbb{E}_q[\cdot]$ denotes expectation with respect to $q(\theta)$, and

$$\mathbb{W}(\theta) = \frac{\kappa(\theta)\mathbb{A}(\theta)g(\theta|\boldsymbol{y},\gamma,\beta)}{q(\theta)}$$

is the importance weight function. The normalizing constant $Z_{\boldsymbol{y}}$ is given by:

$$Z_{\boldsymbol{y}} = \int_{\Theta} \frac{\mathbb{A}(\theta)g(\theta|\boldsymbol{y},\gamma,\beta)}{q(\theta)}q(\theta)d\theta = \mathbb{E}_q\left[\frac{\mathbb{A}(\theta)g(\theta|\boldsymbol{y},\gamma,\beta)}{q(\theta)}\right].$$

At this point, it is easy to see that a natural and convenient choice for the proposal density is $q(\theta) = g(\theta|\boldsymbol{y},\gamma,\beta)$, i.e., the Gamma distribution appearing in the posterior kernel. Under this choice, the posterior expectation simplifies to

$$\mathbb{E}(\kappa(\theta)|\boldsymbol{y}) = \frac{\mathbb{E}_q[\kappa(\theta)\mathbb{A}(\theta)]}{\mathbb{E}_q[\mathbb{A}(\theta)]}.$$

Therefore, an estimate of $\mathbb{E}[\kappa(\theta) \mid \boldsymbol{y}]$ can be obtained by generating a sample of size N from the proposal distribution, i.e., $\theta_1, \ldots, \theta_N \overset{iid}{\sim} Gamma(\gamma^*, \beta^*)$, and computing

$$\hat{\kappa}(\theta) = \frac{\sum_{l=1}^{N} \kappa(\theta_l)\mathbb{A}(\theta_l)}{\sum_{l=1}^{N} \mathbb{A}(\theta_l)}. \quad (8)$$

This corresponds to a self-normalized importance sampling estimator.

Using (8), the Bayesian estimators $\hat{\theta}_{SE}$, $\hat{\theta}_{LI}$ and $\hat{\theta}_{GE}$ are given by:

$$\hat{\theta}_{SE} = \frac{\sum_{l=1}^{N} \theta_l \mathbb{A}(\theta_l)}{\sum_{l=1}^{N} \mathbb{A}(\theta_l)}, \hat{\theta}_{LI} = -\frac{1}{a} ln\left(\frac{\sum_{l=1}^{N} e^{-a\theta_l} \mathbb{A}(\theta_l)}{\sum_{l=1}^{N} \mathbb{A}(\theta_l)}\right),$$

$$\text{and } \hat{\theta}_{\mathrm{GE}} = \left(\frac{\sum_{\mathrm{l}=1}^{\mathrm{N}} (\theta_{\mathrm{l}})^{-\mathrm{b}} \mathbb{A}(\theta_{\mathrm{l}})}{\sum_{\mathrm{l}=1}^{\mathrm{N}} \mathbb{A}(\theta_{\mathrm{l}})}\right)^{-\frac{1}{\mathrm{b}}}. \tag{9}$$

for $a \neq 0$ and $b \neq 0$.

To construct credible intervals for θ, we adopt the method proposed by Chen and Shao (1999). For this, let $\Pi(\theta|\boldsymbol{y}, \gamma, \beta)$ be the posterior cumulative distribution function, and let $\theta^{(\alpha)}$ denote the α-quantile, defined by

$$\theta^{(\alpha)} = inf\{\theta; \Pi(\theta|\boldsymbol{y}, \gamma, \beta) \geq \alpha\}, \text{ for } \alpha \in [0,1].$$

For a given θ^*, the posterior cumulative distribution can be written as

$$\Pi(\theta^*|\boldsymbol{y}, \gamma, \beta) = \mathbb{E}(\mathbb{I}_{\theta \leq \theta^*}|\boldsymbol{y}) = \int_{\Theta} \mathbb{I}_{\theta \leq \theta^*} \pi(\theta|\boldsymbol{y}, \gamma, \beta) d\theta = \frac{\mathbb{E}_q[\mathbb{A}(\theta)\mathbb{I}_{\theta \leq \theta^*}]}{\mathbb{E}_q[\mathbb{A}(\theta)]},$$

where $\mathbb{I}_{\theta \leq \theta^*}$ is the indicator function.

Therefore, similarly to Equation (8), an approximation for $\Pi(\theta^*|\boldsymbol{y}, \alpha, \beta)$ is given by:

$$\widehat{\Pi}(\theta^*|\boldsymbol{y}, \gamma, \beta) = \frac{\sum_{l=1}^{N} \mathbb{A}(\theta_l)\mathbb{I}_{\theta_l \leq \theta^*}}{\sum_{l=1}^{N} \mathbb{A}(\theta_l)}.$$

Denoting by $\theta_{(1)}, \ldots, \theta_{(N)}$ the ordered sample, the normalized weights are defined as

$$w_l = \frac{\mathbb{A}(\theta_{(l)})}{\sum_{l=1}^{L} \mathbb{A}(\theta_{(l)})}. \tag{10}$$

According to Chen and Shao (1999), the estimated cumulative distribution function can be expressed as

$$\widehat{\Pi}(\theta^*|\boldsymbol{y}, \gamma, \beta) = \begin{cases} 0, & if\, \theta^* < \theta_{(1)}; \\ \sum_{j=1}^{l} w_j, & if\, \theta_{(l)} \leq \theta^* < \theta_{(l+1)}; \\ 1, & if\, \theta^* \geq \theta_{(N)}. \end{cases}$$

Consequently, the α-quantile can be estimated by

$$\hat{\theta}^{(\alpha)} = \begin{cases} \theta_{(1)}, & if \alpha = 0; \\ \theta_{(l)}, & if \sum_{j=1}^{l-1} w_j < \alpha \leq \sum_{j=1}^{l} w_j \end{cases}.$$

To construct a $100(1-\alpha)\%$ highest posterior density (HPD) credible interval, consider the intervals

$$R_j(N) = \left(\hat{\theta}^{(j/N)}, \hat{\theta}^{((j+(1-\alpha)N)/N)}\right),$$

for $j = 1, \ldots, J$, where $J = N - \lfloor (1-\alpha)N \rfloor$. The HPD interval is then selected as

$$R_{j^*}(N) = \min_j \left| R_j(N) \right|,$$

that is, the interval with the shortest length. Under regularity conditions, Chen and Shao (1999) showed that $R_{j^*}(N)$ converges to the true $100(1-\alpha)\%$ HPD interval as $N \to \infty$.

In practical terms, the IS algorithm is implemented as described in Algorithm 1

Algorithm 1. Importance Sampling algorithm

1. Input: Sample size N, credibility level $(1-\alpha)$.
2. Generate $\theta_1, \ldots, \theta_N \overset{iid}{\sim} Gamma(\gamma^*, \beta^*)$.
3. Obtain $\theta_{(1)} \leq \cdots \leq \theta_{(N)}$ and reorder weights accordingly
4. Compute unnormalized weights: $w_l^* = \mathbb{A}\left(\theta_{(l)}\right)$, for $l = 1, \ldots, N$.
5. Normalize the weights: $w_l = \frac{w_l^*}{\sum_{j=1}^{N} w_j^*}$.
6. Compute the estimates $\hat{\theta}_{SE}$, $\hat{\theta}_{LI}$ and $\hat{\theta}_{GE}$ using the estimators given in (9).
7. Compute cumulative weights: $S_l = \sum_{j=1}^{l} w_{(j)}$, for $l = 1, \ldots, N$.
8. Compute equal-tail credible interval: $\hat{\theta}_L = \theta_{(l_1)}$ and $\hat{\theta}_U = \theta_{(l_2)}$, where $S_{l_1-1} < \alpha/2 \leq S_{l_1}$ and $S_{l_2-1} < 1-\alpha/2 \leq S_{l_2}$.
9. Output: $\hat{\theta}_{SE}$, $\hat{\theta}_{LI}$, $\hat{\theta}_{GE}$ and interval $\left(\hat{\theta}_L, \hat{\theta}_U\right)$

3.1.2. Independent Metropolis-Hastings Algorithm

An alternative to the IS algorithm for computing Bayesian estimates is the Metropolis-Hastings (MH) algorithm. Similar to the acceptance-rejection method, the MH algorithm generates, at each iteration, a candidate value from a proposal distribution and accepts or rejects it according to a specified acceptance probability. This mechanism ensures that the resulting Markov chain converges to the target posterior distribution. For further details on MH algorithm, please see Hasting (1970), Chib and Greenberg (1995), Gelman et al. (2014), Gilks et al. (1996) and Saraiva and Suzuki (2017).

To generate a sample from posterior distribution in (6), denotes θ be the current state of a Markov chain and let θ^* be a candidate value generated from a candidate-generating density $q[\theta^*|\theta]$. The candidate θ^* is accepted with probability $\Psi(\theta^*|\theta) = min(1, A_\theta)$, where

$$A_\theta = \frac{\mathcal{L}(\theta^*|\boldsymbol{y})\pi(\theta^*|\gamma,\beta)}{\mathcal{L}(\theta|\boldsymbol{y})\pi(\theta|\gamma,\beta)}\frac{q[\theta|\theta^*]}{q[\theta^*|\theta]} = \frac{\mathbb{A}(\theta^*|\boldsymbol{y})g(\theta^*|\boldsymbol{y},\gamma,\beta)}{\mathbb{A}(\theta|\boldsymbol{y})g(\theta|\boldsymbol{y},\gamma,\beta)}\frac{q[\theta|\theta^*]}{q[\theta^*|\theta]}.$$

Similarly to the IS algorithm, a natural and convenient choice for the candidate-generating density is to take it independent of the current state and equal to the Gamma density appearing in the posterior kernel, that is,

$$q(\theta^* \mid \theta) = q(\theta^*) = g\left(\theta^* \mid \gamma^*, \beta^*\right).$$

Under this choice, the proposal terms cancel out, and the acceptance probability simplifies to

$$A_\theta = \frac{\mathbb{A}(\theta^*|\boldsymbol{y})}{\mathbb{A}(\theta|\boldsymbol{y})}. \tag{11}$$

Thus, the MH algorithm used in this article corresponds to an independent Metropolis-Hastings algorithm. In practical terms, the algorithm is implemented as described in Algorithm 2.

Algorithm 2. Independent Metropolis-Hastings algorithm

1. Initialize $\theta^{(0)}$;
2. For $l = 1, \ldots, L$ do
3. Generate a candidate $\theta^* \sim Gamma\left(\gamma^*, \beta^*\right)$
4. Compute the acceptance probability $\Psi\left(\theta^{(l-1)}, \theta^*\right) = min\left\{1, \frac{\mathbb{A}(\theta^*|\boldsymbol{y})}{\mathbb{A}(\theta^{(l-1)}|\boldsymbol{y})}\right\}$.
5. Generate $u \sim \mathcal{U}(0,1)$. If $u \leq \Psi\left(\theta^{(l-1)}, \theta^*\right)$, set $\theta^{(l)} = \theta^*$. Else, set $\theta^{(l)} = \theta^{(l-1)}$

The independent MH algorithm adopted in this work uses the Gamma density $g(\theta \mid \boldsymbol{y}, \gamma^*, \beta^*)$ as the proposal distribution. Although this choice leads to a particularly simple acceptance probability, its efficiency depends on the behaviour of the correction factor $\mathbb{A}(\theta)$. Since $\mathbb{A}(\theta)$ involves products over all observations, it may become highly concentrated around the posterior mode, especially for moderate or large sample sizes. Consequently, the acceptance probability may become small, leading to increased autocorrelation, reduced effective sample sizes, and slower mixing of the generated Markov chain. Nevertheless, the proposal remains valid and produces samples from the target posterior distribution.

3.1.3. Random-walk Metropolis Algorithm

To mitigate potential mixing difficulties associated with the independent MH sampler, we also implemented a random-walk Metropolis (RWM) algorithm. Unlike the

independent MH algorithm, where candidate values are generated from a fixed proposal distribution, the RWM algorithm generates candidates in a neighborhood of the current state of the Markov chain. This local proposal mechanism often leads to improved mixing properties, higher acceptance rates, and larger effective sample sizes.

To generate a sample from the posterior distribution in (6), let θ denote the current state of the Markov chain and let θ^* denote a candidate value generated from the proposal density $q(\theta^*|\theta) := N(\theta, \sigma^2)$, where $\sigma > 0$ is a tuning parameter. Since the parameter space is restricted to positive values, the candidate is generated according to

$$\theta^* = |\theta + \varepsilon|, \varepsilon \sim N(0, \sigma^2).$$

The candidate value θ^* is accepted with probability $\Psi(\theta, \theta^*) = (min1, A_\theta)$ where

$$A_\theta = \frac{\pi(\theta^*|\boldsymbol{y}, \gamma, \beta)}{\pi(\theta|\boldsymbol{y}, \gamma, \beta)} \frac{q(\theta|\theta^*)}{q(\theta^*|\theta)}.$$

Since the random-walk proposal is symmetric, $q(\theta|\theta^*) = q(\theta^*|\theta)$, the proposal terms cancel out and the acceptance probability reduces to

$$A_\theta = \frac{\pi(\theta^*|\boldsymbol{y}, \gamma, \beta)}{\pi(\theta|\boldsymbol{y}, \gamma, \beta)}. \tag{12}$$

To improve the efficiency of the algorithm, the proposal variance is adaptively adjusted during the burn-in period. Let $\sigma_l > 0$ denote the proposal standard deviation at iteration l. Initially, a value σ_0 is specified, and during the burn-in period the sequence σ_l is updated according to the observed acceptance rate (AR). After burn-in, the final value of σ_l is kept fixed for the remainder of the simulation. In the numerical studies presented in this paper, we used $\sigma_0 = 0.1$ and updated σ_l every $m = 500$ iterations. Thus, the RWM algorithm used in this article is implemented as described in Algorithm 3.

Algorithm 3. Adaptative random-walk Metropolis Algorithm-1

1. Initialize $\theta^{(0)}$ and choose an initial proposal standard deviation $\sigma_0 > 0$.
2. For $l = 1, \ldots, L$ do
3. Generate a candidate value $\theta^* = \left|\theta^{(l-1)} + \varepsilon\right|, \varepsilon \sim N(0, \sigma_l^2)$.
4. Compute the acceptance probability $\Psi\left(\theta^{(l-1)}, \theta^*\right) = min\{1, A_\theta\}$ where A_θ is given in Equation (12).
5. Generate $u \sim U(0,1)$. If $u \leq \Psi\left(\theta^{(l-1)}, \theta^*\right)$, set $\theta^{(l)} = \theta^*$; otherwise set $\theta^{(l)} = \theta^{(l-1)}$
6. During the burn-in period, update σ_l every m iterations according to the observed acceptance rate $\widehat{AR}$ computed over the previous m iterations:
$$\sigma_{l+1} = \begin{cases} 0.9, \sigma_l, & if \widehat{AR} < 0.20, \\ 1.1, \sigma_l, & if \widehat{AR} > 0.50, \\ \sigma_l, & otherwise. \end{cases}$$
7. After the burn-in period, keep the final value of σ_l fixed and continue the simulation without further adaptation.

After running either the IMH or the RWM algorithm for L iterations, the first B iterations are discarded as burn-in. To reduce the dependence among successive sampled

values, a thinning procedure with lag J is applied, that is, only every J-th draw is retained. This yields a posterior sample $\mathbb{S} = \{\theta^{(1)}, \dots, \theta^{(N)}\}$, of size $N = \left\lfloor \frac{L-B}{J} \right\rfloor$, which is used for posterior inference.

The Bayesian estimates under the squared-error, LINEX, and generalized entropy loss functions are computed from the sampled values in $\mathbb{S}$ as

$$\tilde{\theta}_{SE} = \frac{1}{N}\sum_{l=1}^{N}\theta^{(l)}, \tilde{\theta}_{LI} = -\frac{1}{a}ln\left(\frac{1}{N}\sum_{l=1}^{N}exp\left(-a\theta^{(l)}\right)\right), \tilde{\theta}_{GE} = \left(\frac{1}{N}\sum_{l=1}^{N}\left(\theta^{(l)}\right)^{-b}\right)^{-1/b}, \tag{13}$$

for $a \neq 0$ and $b \neq 0$. A $100(1-\alpha)\%$ credible interval for θ is obtained from the empirical quantiles of $\mathbb{S}$, namely the $(\alpha/2)$- and $(1-\alpha/2)$-quantiles, where $\alpha \in (0,1)$. When the posterior sample is generated by the RWM algorithm, the corresponding Bayesian estimators are denoted by $\tilde{\theta}^*_{SE}$, $\tilde{\theta}^*_{LI}$, and $\tilde{\theta}^*_{GE}$.

4. BAYESIAN POSTERIOR PREDICTIVE INFERENCE AND FORECASTING

In addition to parameter estimation, an important advantage of the Bayesian framework is the possibility of making probabilistic forecasts for future lifetimes while fully accounting for parameter uncertainty. In the context of survival analysis, this is achieved through the posterior predictive distribution, which provides a natural basis for forecasting survival probabilities and future event times.

Let T^{new} denote the lifetime of a future individual, assumed to follow the XGamma distribution indexed by the same parameter θ as the observed data. The future lifetime T^{new} is also assumed to be independent of the observed data $\boldsymbol{y} = (y_1, \dots, y_n)$, where $y_i = (t_i, \delta_i)$ for $i = 1, \dots, n$.

Under the Bayesian framework, the posterior predictive density of T^{new} given the observed data $\boldsymbol{y}$ is defined as

$$p(t^{new} \mid \boldsymbol{y}) = \int_0^\infty f\,(t^{new} \mid \theta)\pi(\theta \mid \boldsymbol{y}, \gamma, \beta)d\theta, \tag{14}$$

where $\pi(\theta \mid \boldsymbol{y}, \gamma, \beta)$ denotes the posterior density of θ given in Equation (6).

4.1. Monte Carlo Approximation

Due to the complexity of the posterior distribution, the integral in (14) does not admit a closed-form expression. However, it can be accurately approximated using Monte Carlo methods.

For the IMH and RWM algorithms, let $\theta^{(1)}, \dots, \theta^{(N)}$ denote a sample drawn from the posterior distribution of θ. For the IS method, let $\theta^{(1)}, \dots, \theta^{(N)}$ denote a sample generated from the proposal distribution $q(\theta)$, with associated normalized importance weights $w_1, \dots, w_N$ and $\sum_{l=1}^{N} w_l = 1$. A sample from the posterior predictive distribution is approximated by generating

$$T^{new(l)} \sim XGamma\left(\theta^{(l)}\right), l = 1, \ldots, N.$$

For the IS approach, posterior predictive expectations are estimated using weighted averages, whereas for the IMH and RWM approaches they are estimated using ordinary Monte Carlo averages. In particular, the posterior predictive mean is approximated by

$$\hat{T}_{IS}^{new} = \sum_{l=1}^{N} w_l \, T^{new(l)}, \hat{T}_{MH}^{new} = \frac{1}{N} \sum_{l=1}^{N} T^{new(l)},$$

where $\hat{T}_{MH}^{new}$ denotes the estimator obtained from either the IMH or RWM sample.

Moreover, for the IS approach, the posterior predictive cumulative distribution function can be approximated by

$$\hat{F}_{IS}(t) = \sum_{l=1}^{N} w_l \, 1\{T^{new(l)} \leq t\}.$$

Hence, a $100(1-\alpha)\%$ predictive interval is obtained from the weighted quantiles of the predictive sample, namely

$$\left(\hat{F}_{IS}^{-1}(\alpha/2), \hat{F}_{IS}^{-1}(1-\alpha/2)\right),$$

where $\hat{F}_{IS}^{-1}(\cdot)$ denotes the generalized inverse of the estimated predictive distribution function.

For the IMH and RWM approaches, the predictive interval is obtained from the empirical quantiles of the predictive sample. Let

$$T_{(1)}^{new} \leq \cdots \leq T_{(N)}^{new}$$

denote the ordered predictive sample. Then a $100(1-\alpha)\%$ predictive interval is given by

$$\left(T_{([N\alpha/2])}^{new}, T_{([N(1-\alpha/2)])}^{new}\right).$$

The predictive distribution under the MLE approach is obtained using the plug-in principle, whereby the unknown parameter is replaced by its maximum likelihood estimate. Under this approach, an independent predictive sample is generated according to

$$T^{new(l)} \sim XGamma\left(\hat{\theta}_{ML}\right), l = 1, \ldots, N.$$

The corresponding predictive mean is estimated by

$$\hat{T}_{MLE}^{new} = \frac{1}{N} \sum_{l=1}^{N} T^{new(l)}.$$

A $100(1-\alpha)\%$ predictive interval is obtained from the empirical quantiles of the predictive sample, namely

$$\left(\hat{T}^{new} *_{\alpha/2}, \hat{T}^{new} *_{1-\alpha/2}\right),$$

where $\hat{T}_p^{new}$ denotes the empirical p-quantile of the generated predictive sample.

5. SIMULATION STUDY ON THE FIT PERFORMANCE

In this section, we illustrate and compare the performance of the previously described estimation procedures using simulated data sets. To generate the artificial datasets, we adopted the following procedure:

1. Fix the value of θ, the sample size n, and the percentage p of right-censored observations.
2. For $i = 1, \dots, n$, perform the following steps:
   a) Generate $V_i \sim \mathcal{E}xp(\theta)$ and $W_i \sim \mathcal{G}amma(3, \theta)$.
   b) Generate $U_i \sim \mathcal{U}(0,1)$. Define the lifetime T_i as $T_i = \begin{cases} V_i, & if U_i \leq \frac{\theta}{1+\theta} \\ W_i, & otherwise. \end{cases}$
   c) Generate a censoring time $C_i \sim \mathcal{U}(e_1, e_2)$, where $\mathcal{U}(e_1, e_2)$ denotes a continuous uniform distribution with parameters e_1 and e_2. The values of e_1 and e_2 are chosen such that approximately $p\%$ of the observations are right-censored.
   d) Define the observed data $y_i = (t_i, \delta_i)$, where $t_i = min(T_i, C_i)$ and $\delta_i = \begin{cases} 1, & if T_i \leq C_i \\ 0, & if T_i > C_i. \end{cases}$

After generating a dataset, we obtain the ML estimate $\hat{\theta}_{ML}$ and the Bayesian estimates (BE):

(i) BE-IS: $\hat{\theta}_{SE}$, $\hat{\theta}_{LI}$ and $\hat{\theta}_{GE}$ using the IS algorithm, given by Equations in (9);
(ii) BE-MH: $\tilde{\theta}_{SE}$, $\tilde{\theta}_{LI}$ and $\tilde{\theta}_{GE}$ using the IMH algorithm, given by Equations in (13);
(iii) BE-MH: θ^*_{SI}, θ^*_{LI} and θ^*_{GE} using the RWM algorithm, given by Equations in (13).

To obtain the Bayesian estimators, we fix the hyperparameter values at $\gamma = \beta = 0.01$ in order to specify a weakly informative prior distribution. This specification is commonly adopted for positive parameters because it yields a highly dispersed prior distribution while preserving support on the positive real line (see, e.g., Banerjee and Carlin, 2014; Bernardo and Smith, 1994, Saraiva and Suzuki, 2017). We then apply the IS algorithm with $N = 20{,}000$ iterations, the independent Metropolis–Hastings (IMH) algorithm with $L = 120{,}000$, $B = 20{,}000$, and $J = 5$, and the random-walk Metropolis (RWM) algorithm with the same values of L, B, and J. Consequently, the Bayesian estimators obtained via IS, IMH, and RWM are all based on $20{,}000$ posterior draws, allowing a direct comparison among the three computational approaches.

Furthermore, to obtain the Bayesian estimators under the LINEX and generalized entropy loss functions, it is necessary to specify values for the asymmetry parameters a and b, respectively. Since a and b are not parameters of the XGamma distribution, they should be interpreted as tuning parameters whose values must be selected prior to the implementation of the corresponding estimation procedures. Consequently, our objective is not to infer a and b within the Bayesian framework, but rather to calibrate them. To this end, we adopt an adaptive grid-search procedure that selects values of a and b yielding estimators whose fitted survival functions exhibit satisfactory agreement with the Kaplan–Meier estimator. Therefore, the procedure described below should be interpreted as a calibration mechanism for selecting suitable values of a and b, rather than as a Bayesian estimation procedure for these quantities.

To select suitable values of a and b, we implement the following numerical calibration procedure:

i) Let $\mathbb{G}^a = \{g_1^a, \ldots, g_m^a\}$ be a grid of values starting from g_1^a and ending at g_m^a with increments of size c, and let $\mathbb{G}^b = \{g_1^b, \ldots, g_m^b\}$ be defined analogously, for $c > 0$. In the remainder of the paper, we fix $g_1^a = g_1^b = -10$, $g_m^a = g_m^b = 10$, and $c = 0.1$.

ii) For each value $a \in \mathbb{G}^a$ and $b \in \mathbb{G}^b$, compute the estimators $\hat{\theta}_{LI}$ and $\hat{\theta}_{GE}$ using the IS algorithm, the estimators $\tilde{\theta}_{LI}$ and $\tilde{\theta}_{GE}$ using the IMH algorithm, and $\tilde{\theta}_{LI}^*$ and $\tilde{\theta}_{GE}^*$ estimators using the RWM algorithm.

iii) Measure the agreement between the fitted survival function and the Kaplan–Meier estimator using the RMSE criterion:

$$RMSE(\theta_{BE-LI}, a) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{S}(t_i \mid \theta = \theta_{BE-LI}, a) - S_{KM}(t_i)\right)^2},$$

$$RMSE(\theta_{BE-GE}, b) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{S}(t_i \mid \theta = \theta_{BE-GE}, b) - S_{KM}(t_i)\right)^2},$$

where $\theta_{BE-LI} \in \{\hat{\theta}_{LI}, \tilde{\theta}_{LI}, \tilde{\theta}_{LI}^*\}$ and $\theta_{BE-GE} \in \{\hat{\theta}_{GE}, \tilde{\theta}_{GE}, \tilde{\theta}_{GE}^*\}$.

iv) Select the values

$$a_{IS}^* = \underset{a \in \mathbb{G}^a}{argmin} RMSE\left(\hat{\theta}_{LI}, a\right) \text{ and } b_{IS}^* = \underset{b \in \mathbb{G}^b}{argmin} RMSE\left(\hat{\theta}_{GE}, b\right).$$

If the minimum RMSE is attained at an interior point of the corresponding grid, then the selected value is retained. Otherwise, the grid is expanded and the procedure is repeated. Specifically, if the selected value is non-positive, the grid is updated to span from the selected value minus 10 to the selected value plus 0.1, with increments of size $c = 0.1$. If the selected value is positive, the grid is updated to span from the selected value minus 0.1 to the selected value plus 10, again using increments of size $c = 0.1$. The procedure is repeated until the minimum RMSE is attained at an interior point of the grid.

v) Repeat Step (iv) using the MH estimators $\tilde{\theta}_{LI}$ and $\tilde{\theta}_{GE}$ to obtain the selected values a_{MH}^* and b_{MH}^*.

Since the optimum may not always be enclosed by the current search region, repeated grid expansions could continue indefinitely. Moreover, extreme values of the calibration parameters may result in numerical instability, as the LINEX and generalized entropy estimators involve exponential and power transformations that can produce overflow, underflow, or non-finite values. Therefore, the search is terminated whenever a non-finite estimator value is produced or when one of the grid limits reaches the prespecified bounds $[-500,500]$. In either case, the last finite value obtained before termination is retained as the selected calibration parameter.

Once the selected values a_{IS}^* and b_{IS}^* have been obtained, the corresponding estimators $\hat{\theta}_{LI}$ and $\hat{\theta}_{GE}$ are adopted as the importance sampling (IS) Bayesian estimators under the LINEX and generalized entropy loss functions, respectively. Similarly, after obtaining the selected values a_{IMH}^* and b_{IMH}^*, the corresponding estimators $\tilde{\theta}_{LI}$ and $\tilde{\theta}_{GE}$ are adopted as the independent Metropolis–Hastings (IMH) Bayesian estimators. Finally, after obtaining the selected values a_{RWM}^* and b_{RWM}^*, the corresponding estimators $\tilde{\theta}_{LI}^*$ and $\tilde{\theta}_{GE}^*$ are adopted as the random-walk Metropolis (RWM) Bayesian estimators.

Although the present work treats a and b as tuning parameters selected through a calibration procedure, alternative approaches that incorporate these quantities directly into the inferential framework and estimate them jointly with θ constitute an interesting direction for future research.

5.1. Simulation Study 1: An Illustrative Example Based on a Single Simulation

To illustrate the proposed methodology, we generated a right-censored sample of size $n = 100$ from the XGamma distribution with parameter $\theta = 0.6$, with 50% of the observations being right-censored. The sample size, parameter value, and censoring proportion were selected to resemble the main characteristics of the first real data set analysed in the next section, providing a realistic setting for evaluating the proposed inferential procedures.

The data-generation procedure is described in Appendix A.1 of the Supplementary Material. In addition, this appendix presents graphical summaries of the simulated data, including the corresponding Kaplan–Meier curve and its 95% confidence band.

After generating the data set, the methods described in the previous sections were applied to estimate θ. Table 1 presents the resulting point estimates together with their absolute errors (AE) relative to the true parameter value $\theta = 0.6$. Additional illustrations of the log-likelihood and score functions, details of the calibration procedure used to select the tuning parameters a and b, and supplementary convergence diagnostics for the IMH and RWM algorithms are reported in Appendices A.2, A.3, and A.4 of the Supplementary Material, respectively.

Using the RMSE-based calibration procedure described in Section 3.1.1, the selected tuning parameters were $a_{IS}^{*} = -340.8$ and $b_{IS}^{*} = -214.4$ for the IS algorithm, $a_{IMH}^{*} = -29.1$ and $b_{IMH}^{*} = -19.1$ for the IMH algorithm, and $a_{RWM}^{*} = -4.6$ and $b_{RWM}^{*} = -3.9$ for the RWM algorithm. These results indicate that the optimal values of a and b depend on the posterior sample used in the calibration procedure. Since the calibration is based on minimizing the discrepancy between the estimated survival function and the Kaplan–Meier estimator, different posterior sampling schemes may lead to different optimal values of the tuning parameters, even when targeting the same posterior distribution. Despite these differences, the three estimation procedures produced fitted survival functions with very similar RMSE values relative to the Kaplan–Meier estimator.

As shown in Table 1, all estimators provided accurate estimates of the parameter, with absolute errors smaller than 0.02. Among the methods considered, the Bayesian estimator $\tilde{\theta}_{SE}^{*}$ obtained through the RWM algorithm yielded the smallest absolute error ($AE = 0.0090$), followed closely by the importance sampling estimator $\hat{\theta}_{SE}$ ($AE = 0.0094$) and the IMH estimator $\tilde{\theta}_{SE}$ ($AE = 0.0099$). The estimators derived under the LINEX and generalized entropy loss functions produced nearly identical results across the IS, IMH, and RWM approaches, suggesting that the choice of posterior sampling scheme had only a minor effect on the resulting Bayesian point estimates. Overall, all methods exhibited satisfactory estimation accuracy, with estimates remaining very close to the true parameter value $\theta = 0.60$.

Table 1
Point Estimates and Absolute Errors for the Simulated Data Set-1

Results	Method									
	ML	IS			IMH			RWM		
	$\hat{\theta}_{ML}$	$\hat{\theta}_{SE}$	$\hat{\theta}_{LI}$	$\hat{\theta}_{GE}$	$\tilde{\theta}_{SE}$	$\tilde{\theta}_{LI}$	$\tilde{\theta}^*_{GE}$	$\tilde{\theta}^*_{SE}$	$\tilde{\theta}^*_{LI}$	$\tilde{\theta}^*_{GE}$
Estimate	0.6093	0.5906	0.616	0.616	0.5901	0.6160	0.6160	0.6091	0.6159	0.6160
AE	0.0093	0.0094	0.016	0.016	0.0099	0.0160	0.0160	0.0091	0.0159	0.0160

Table 2 presents the 95% confidence and credible intervals obtained using the different estimation procedures. For the Bayesian approaches, both equal-tail and highest posterior density (HPD) credible intervals were considered. The equal-tail intervals were constructed from the posterior quantiles, while the HPD intervals were obtained using the HPD interval() function available in the coda package in R. As shown in the table, all intervals contain the true parameter value $\theta = 0.6$, indicating satisfactory interval estimation performance for all methods considered. Among the Bayesian procedures, the HPD intervals were uniformly shorter than their corresponding equal-tail intervals. In particular, the HPD and equal-tail intervals obtained through importance sampling (IS) yielded the two shortest interval lengths, namely 0.1253 and 0.1384, respectively, while still covering the true parameter value. Overall, the Bayesian procedures produced substantially shorter intervals than the Wald confidence interval associated with the maximum likelihood estimator, suggesting improved estimation precision.

Table 2
Comparison of Interval Estimators for the XGamma Distribution

Method	Interval	Lower	Upper	Length
MLE	Wald	0.5028	0.7158	0.2130
IS	Equal-tail	0.5032	0.6416	0.1384
	HPD	0.5187	0.6440	0.1253
IMH	Equal-tail	0.5021	0.6794	0.1773
	HPD	0.5163	0.6794	0.1631
RWM	Equal-tail	0.5071	0.7198	0.2127
	HPD	0.5069	0.7196	0.2126

The convergence diagnostics presented in Appendix A.4 indicate that the IMH algorithm generated considerably smaller effective sample sizes than the RWM algorithm, suggesting stronger autocorrelation among successive posterior draws and consequently lower sampling efficiency. Despite this difference in mixing behaviour, the two algorithms produced very similar point estimates (see Table 1) and credible intervals (see Table 2). In particular, the IMH estimator $\tilde{\theta}_{SE}$ yielded one of the smallest absolute errors among the methods considered, while the corresponding equal-tail and HPD credible intervals

exhibited lengths comparable to those obtained by the RWM algorithm. Moreover, all credible intervals successfully covered the true parameter value $\theta = 0.6$. These findings indicate that, although the IMH sampler mixes less efficiently than the RWM sampler, it still provides reliable posterior inference for the present data set. Consequently, the primary drawback of the poorer mixing behaviour is computational rather than inferential, since a larger number of iterations is required to obtain a given number of effectively independent posterior samples.

Figure 1(a) compares the true survival function (black line, corresponding to $\theta_0 = 0.6$) with the fitted survival functions obtained from the different estimation procedures (coloured lines). Since the resulting estimates are very close to one another, only the fitted survival curves corresponding to $\hat{\theta}_{ML}$, $\hat{\theta}_{SE}$, $\tilde{\theta}_{SE}$ and $\tilde{\theta}^{*}_{SE}$ are displayed to improve visual clarity. Figure 1(b) compares the Kaplan–Meier estimator with the corresponding fitted survival functions. In both panels, the estimated curves closely follow the reference curves, indicating that all methods provide satisfactory fits to the simulated data.

To further assess the quality of the fitted models, we computed the RMSE between the estimated survival functions and the Kaplan–Meier estimator,

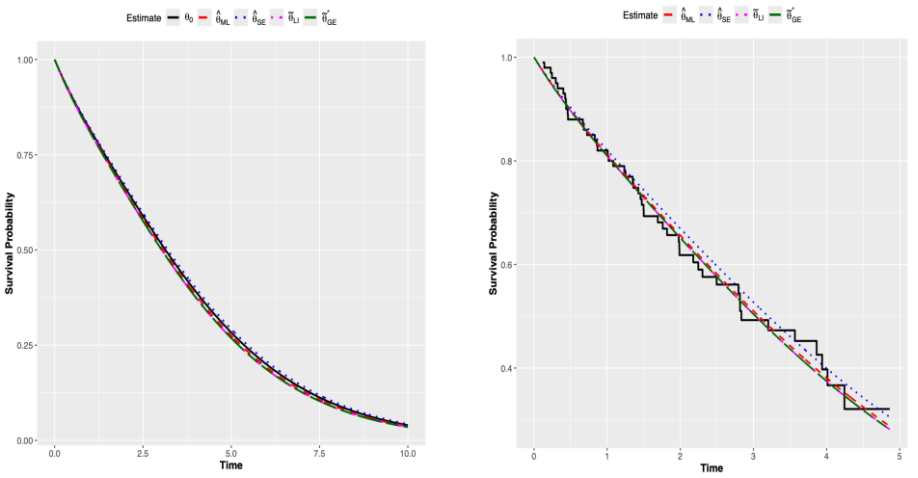
$$RMSE_{KM}\left(\hat{\theta}_{method}\right) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\hat{S}\left(t_i|\hat{\theta}_{method}\right) - \hat{S}_{KM}(t_i)\right]^2},$$

for $\hat{\theta}_{method} \in \{\hat{\theta}_{ML}, \hat{\theta}_{SE}, \hat{\theta}_{LI}, \hat{\theta}_{GE}, \tilde{\theta}_{SE}, \tilde{\theta}_{LI}, \tilde{\theta}_{GE}, \tilde{\theta}^{*}_{SE}, \tilde{\theta}^{*}_{LI}, \tilde{\theta}^{*}_{GE}, \}$.

The corresponding RMSE values are reported in Table 3. The results indicate that all fitted survival functions are very close to the Kaplan–Meier estimator, with RMSE values ranging from 0.0205 to 0.0283. As expected, the estimators obtained under the LINEX and generalized entropy loss functions yielded the smallest RMSE values, since the tuning parameters a and b were selected through a calibration procedure that explicitly minimizes this criterion. Notably, the LINEX- and GE-based estimators achieved identical RMSE values, suggesting that the calibration procedure led to fitted survival functions with essentially the same agreement with the Kaplan–Meier estimator. This finding is consistent with the results reported in Table 1, where the corresponding point estimates were nearly identical. Although the LINEX- and GE-based estimators provided a slightly better agreement with the Kaplan–Meier curve, the differences among the competing methods were small. Overall, both the maximum likelihood and Bayesian approaches produced satisfactory fits to the simulated data.

Table 3
RMSE Values with respect to the Kaplan–Meier Estimator

Estimate	$\hat{\theta}_{ML}$	$\hat{\theta}_{SE}$	$\hat{\theta}_{LI}$	$\hat{\theta}_{GE}$	$\tilde{\theta}_{SE}$	$\tilde{\theta}_{LI}$	$\tilde{\theta}^{*}_{GE}$	$\tilde{\theta}^{*}_{SE}$	$\tilde{\theta}^{*}_{LI}$	$\tilde{\theta}^{*}_{GE}$
$RMSE_{KM}$	0.0211	0.0275	0.0205	0.0205	0.0277	0.0205	0.0205	0.0211	0.0205	0.0205



(a) True and estimated survival curves **(b) Kaplan–Meier and estimated survival curves**

Figure 1: Survival Estimates

5.2 Simulation Study 2: Comparative Performance Evaluation Across Multiple Scenarios

To complement the results obtained from a single simulated data set, we conducted a broader simulation study involving different sample sizes and censoring proportions. Specifically, random samples were generated from the XGamma distribution with parameter $\theta = 0.6$, considering sample sizes $n \in \{25,50,100\}$ and censoring percentages $p \in \{0\%, 15\%, 30\%, 50\%\}$.

For each combination of n and p, we generated $M = 1{,}000$ independent data sets and compared the competing estimators using the following two criteria:

1) the mean absolute error (MAE), which assesses the accuracy of parameter estimation;
2) the mean root mean squared error relative to the Kaplan–Meier estimator, denoted by $\overline{RMSE}_{KM}$, which assesses the accuracy of the estimated survival function.

Similarly to Simulation Study 1, the IS algorithm was implemented with a posterior sample size of $N = 20{,}000$. For the IMH and RWM algorithms, $L = 120{,}000$ MCMC iterations were generated, with a burn-in period of size $B = 5{,}000$ and thinning interval $J = 5$. This configuration yields $20{,}000$ retained posterior draws for each algorithm, thereby ensuring a fair comparison among the competing Bayesian estimation procedures.

Table 4 summarizes the mean estimates (ME) and mean absolute errors (MAE, reported in parentheses) of the MLE, IS, IMH, and RWM estimators based on $1{,}000$ Monte Carlo replications under different sample sizes and censoring percentages. Overall, all estimators performed satisfactorily, with mean estimates remaining close to the true value $\theta = 0.6$ across all scenarios. Estimation accuracy improved with increasing sample size, as reflected by decreasing MAE values, whereas higher censoring proportions led to larger

errors due to information loss. Nevertheless, all methods remained reasonably accurate even under 50% censoring.

The MLE and the Bayesian estimator under squared-error loss generally produced the smallest MAE values. In contrast, the LINEX and generalized entropy estimators yielded slightly larger errors, which is expected since they are optimized under alternative loss functions. For each loss function, the IS, IMH, and RWM implementations produced nearly identical estimates and MAE values, indicating that the choice of posterior sampling scheme has little impact on the resulting Bayesian point estimates.

Table 3 in Appendix B of the Supplementary Material reports the average RMSE between the fitted survival functions and the Kaplan–Meier estimator over 1,000 Monte Carlo replications. As expected, RMSE values decreased with increasing sample size and increased with the censoring proportion. The Bayesian estimators based on the LINEX and generalized entropy loss functions consistently achieved smaller RMSE values than those obtained under squared-error loss, particularly for small and moderate sample sizes. Moreover, the IS, IMH, and RWM algorithms yielded very similar RMSE values, suggesting that all three methods provide comparable approximations to the posterior distribution. These differences became negligible for larger samples, where all estimators exhibited good agreement with the Kaplan–Meier estimator.

Table 4
Mean Estimates and Mean Absolute Error Over 1,000 Monte Carlo Replications

Method	Est	Sample Size and Percentage of Censures											
		25				50				100			
		0	15%	30%	50%	0%	15%	30%	50%	0%	15%	30%	50%
ML	$\widehat{\theta}_{ML}$	0.6132 (0.0657)	0.6107 (0.0686)	0.6087 (0.0743)	0.5974 (0.0891)	0.6030 (0.0468)	0.6047 (0.0491)	0.6045 (0.0520)	0.6052 (0.0609)	0.6030 (0.0309)	0.6035 (0.0330)	0.6024 (0.0358)	0.6013 (0.0438)
BE-IS	$\widehat{\theta}_{SE}$	0.6133 (0.0657)	0.6107 (0.0686)	0.6086 (0.0742)	0.5962 (0.0897)	0.6030 (0.0468)	0.6046 (0.0493)	0.6043 (0.0521)	0.6041 (0.0619)	0.6028 (0.0311)	0.6031 (0.0335)	0.6015 (0.0371)	0.5969 (0.0482)
	$\widehat{\theta}_{LI}$	0.6396 (0.0804)	0.6366 (0.0814)	0.6310 (0.0849)	0.6134 (0.0950)	0.6145 (0.0542)	0.6160 (0.0560)	0.6149 (0.0566)	0.6134 (0.0640)	0.6092 (0.0354)	0.6090 (0.0372)	0.6055 (0.0399)	0.5989 (0.0507)
	$\widehat{\theta}_{GE}$	0.6396 (0.0805)	0.6367 (0.0814)	0.6310 (0.0849)	0.6134 (0.0950)	0.6145 (0.0542)	0.6160 (0.0590)	0.6150 (0.0566)	0.6135 (0.0639)	0.6095 (0.0353)	0.6098 (0.0371)	0.6076 (0.0386)	0.6041 (0.0461)
BE-IMH	$\widetilde{\theta}_{SE}$	0.6133 (0.0657)	0.6107 (0.0686)	0.6084 (0.0744)	0.5965 (0.0895)	0.6030 (0.0469)	0.6047 (0.0491)	0.6043 (0.0521)	0.6050 (0.0609)	0.6030 (0.0311)	0.6034 (0.0331)	0.6034 (0.0360)	0.5991 (0.0465)
	$\widetilde{\theta}_{LI}$	0.6396 (0.0804)	0.6367 (0.0814)	0.6310 (0.0849)	0.6134 (0.0950)	0.6145 (0.0542)	0.6160 (0.0561)	0.6150 (0.0566)	0.6135 (0.0639)	0.6095 (0.0353)	0.6098 (0.0370)	0.6079 (0.0385)	0.6049 (0.0452)
	$\widetilde{\theta}_{GE}$	0.6396 (0.0804)	0.6367 (0.0814)	0.6310 (0.0849)	0.6134 (0.0950)	0.6145 (0.0542)	0.6160 (0.0561)	0.6150 (0.0566)	0.6135 (0.0639)	0.6095 (0.0353)	0.6098 (0.0370)	0.6079 (0.0385)	0.6050 (0.0451)
BE-RWM	$\widetilde{\theta}_{SE}$	0.6133 (0.0657)	0.6107 (0.0686)	0.6084 (0.0745)	0.5964 (0.0895)	0.6030 (0.0469)	0.6047 (0.0491)	0.6044 (0.0521)	0.6047 (0.0610)	0.6030 (0.0309)	0.6035 (0.0330)	0.6023 (0.0358)	0.6010 (0.0438)
	$\widetilde{\theta}_{LI}$	0.6396 (0.0804)	0.6367 (0.0814)	0.6310 (0.0849)	0.6134 (0.0950)	0.6145 (0.0542)	0.6161 (0.0561)	0.6150 (0.0566)	0.6135 (0.0639)	0.6095 (0.0353)	0.6098 (0.0370)	0.6079 (0.0385)	0.6050 (0.0451)
	$\widetilde{\theta}_{GE}$	0.6396 (0.0804)	0.6367 (0.0814)	0.6310 (0.0849)	0.6134 (0.0950)	0.6146 (0.0542)	0.6160 (0.0561)	0.6150 (0.0566)	0.6135 (0.0639)	0.6095 (0.0353)	0.6098 (0.0370)	0.6079 (0.0385)	0.6050 (0.0451)

Table 5 reports the empirical coverage probabilities and average interval lengths (reported in parentheses) obtained from 1,000 Monte Carlo replications. Overall, all interval estimators achieved coverage probabilities close to the nominal 95% level. As expected, interval lengths decreased with increasing sample size and increased with the censoring proportion. The Bayesian credible intervals exhibited performance comparable

to that of the maximum likelihood confidence intervals in terms of both coverage and average length

Table 5
Coverage Probabilities and Average Interval Lengths obtained from the Monte Carlo Simulation Study

Method	Est	Sample Size and Percentage of Censures											
		25				50				100			
		0	15%	30%	50%	0%	15%	30%	50%	0%	15%	30%	50%
ML	Wald	0.9510 (0.3143)	0.9490 (0.3365)	0.9480 (0.3660)	0.9380 (0.4216)	0.9300 (0.2181)	0.9410 (0.2354)	0.9560 (0.2565)	0.9530 (0.3006)	0.9500 (0.1542)	0.9550 (0.1662)	0.9580 (0.1804)	0.9460 (0.2109)
BE-IS	Equal Tail	0.9460 (0.3134)	0.9480 (0.3358)	0.9490 (0.3599)	0.9280 (0.4154)	0.9230 0.2174	0.9320 (0.2340)	0.9360 (0.2519)	0.8950 (0.2853)	0.9350 (0.1516)	0.9140 (0.1585)	0.8720 (0.1613)	0.7610 (0.1657)
	HPD	0.9450 (0.3104)	0.9470 (0.3321)	0.9520 (0.3599)	0.9320 (0.4091)	0.9270 0.2151	0.9330 (0.2310)	0.9400 (0.2476)	0.9090 (0.2791)	0.9340 (0.1491)	0.9270 (0.1545)	0.8720 (0.1547)	0.7580 (0.1547)
BE-IMH	Equal Tail	0.9490 (0.3139)	0.9440 (0.3360	0.9530 (0.3658)	0.9380 (0.4211)	0.9250 0.2180	0.9410 (0.2360)	0.9480 (0.2573)	0.9230 (0.2970)	0.9430 (0.1545)	0.9320 (0.1653)	0.9150 (0.1731)	0.8350 (0.1826)
	HPD	0.9500 (0.3117)	0.9490 (0.3332)	0.9500 (0.3616)	0.9330 (0.4134)	0.9280 0.2162	0.9400 (0,2331)	0.9470 (0.2518)	0.9220 (0.2876)	0.9440 (0.1599)	0.9350 (0.1647)	0.9120 (0.1704)	0.8180 (0.1820)
BE-RWM	HPD	0.9470 (0.3138)	0.9470 (0.3360)	0.9500 (0.3653)	0.9470 (0.4201)	0.9290 0.2180	0.9460 (0.2352)	0.9610 (0.2564)	0.9530 (0.3001)	0.9540 (0.1541)	0.9540 (0.1661)	0.9580 (0.1804)	0.9500 (0.2107)
	Equal Tail	0.9490 (0.3120)	0.9470 (0.3340)	0.9470 (0.3627)	0.9370 (0.4165)	0.9290 0.2172	0.9390 (0.2342)	0.9560 (0.2552)	0.9510 (0.2985)	0.9490 (0.1537)	0.9530 (0.1656)	0.9550 (0.1798)	0.9490 (0.2099)

Table 4 in Appendix B of the Supplementary material summarizes the MCMC convergence diagnostics over 1,000 Monte Carlo replications for the IMH and RWM algorithms. Overall, the RWM algorithm exhibited higher ESS and RESS values, together with more stable acceptance rates across all scenarios. Although the performance of the IMH algorithm deteriorated as the sample size and censoring proportion increased, the Gelman–Rubin statistics remained close to one for both algorithms, indicating satisfactory convergence of the generated chains. Thus, while the IMH sampler mixes less efficiently than the RWM sampler, it still provides reliable posterior inference. Its main limitation is computational rather than inferential, as a larger number of iterations is required to obtain a given number of effectively independent posterior samples.

In summary, the Monte Carlo study confirms that the proposed Bayesian procedures yield accurate and robust inference for the XGamma distribution. Estimation accuracy improves with increasing sample size and decreases with heavier censoring, while empirical coverages remain close to the nominal level. The IS, IMH, and RWM implementations produce virtually identical inferential results, with differences arising primarily in computational efficiency. Overall, the results support the effectiveness and practical applicability of the proposed methodology across a broad range of scenarios.

5.3 Simulation Study on Predictive Performance

To evaluate the predictive performance of the proposed methods, we considered the same simulation settings adopted in Simulation Study 2. For each combination of sample size $n \in \{25,50,100\}$ and censoring proportion $p \in \{0\%, 15\%, 30\%, 50\%\}$, we generated $M = 1{,}000$ independent datasets from the XGamma distribution with parameter $\theta = 0.6$. For each simulated dataset, a set of $K = 30$ future lifetimes,

$$T_{m,1}^{new}, \ldots, T_{m,K}^{new} \sim XGamma(\theta),$$

was generated independently from the true model, representing future observations from new individuals. These values were used to assess the predictive performance of the maximum likelihood and Bayesian predictive procedures.

For each generated dataset, the predicted value of the future lifetime is obtained using the predictive mean defined in Section 4, denoted by $\hat{T}_m$, for $m = 1, \dots, M$. Let $T_{m,1}^{new}, \dots, T_{m,K}^{new}$ denote the K future observations generated under the true model for the mth replication. Prediction accuracy is evaluated through the root mean squared prediction error (RMSPE), defined as

$$\text{RMSPE} = \sqrt{\frac{1}{\text{MK}} \sum_{\text{m=1}}^{\text{M}} \sum_{\text{k=1}}^{\text{K}} \left(\text{T}_{\text{m,k}}^{\text{new}} - \hat{T}_{\text{m}}\right)^2}.$$

In addition, predictive interval performance is assessed through the empirical coverage probability (CP) and the average interval length (AL). Together, these measures provide a comprehensive evaluation of both point prediction accuracy and uncertainty quantification.

Table 6 summarizes the predictive performance of the competing methods in terms of RMSPE, coverage probability (CP), and average interval length (AL). All methods achieve CP values close to the nominal 95% level, indicating reliable predictive inference. As expected, RMSPE decreases with increasing sample size and tends to increase with heavier censoring. The Bayesian and ML approaches exhibit very similar predictive performance, while the IS, IMH, and RWM algorithms yield nearly identical results across all scenarios.

Table 6
Coverage Probabilities and Average Lengths of Predictive Intervals

Method	Sample Size and Percentage of Censures											
	25				50				100			
	0	15%	30%	50%	0%	15%	30%	50%	0%	15%	30%	50%
ML	3.0362	2.9976	3.0531	3.1402	3.0366	3.0387	3.0114	3.0361	2.9818	2.9671	3.0178	3.0076
	0.9412	0.9445	0.9413	0.94134	0.9455	0.9455	0.9471	0.9461	0.9478	0.9493	0.9475	0.9468
	10.8835	10.9558	11.0292	11.3967	10.9847	10.9576	10.9747	11.0193	10.9150	10.9194	10.9453	10.9978
BE-IS	3.1340	3.1883	3.4004	4.0595	3.1123	3.1585	3.2261	3.5044	3.0344	3.0588	3.1883	3.4432
	0.9537	0.9566	0.9547	0.9520	0.9549	0.9482	0.9585	0.9557	0.9571	0.9580	0.9537	0.9436
	12.9677	13.5532	14.5960	17.2648	12.6804	11.2163	13.8553	15.2934	12.4449	12.8657	13.3938	14.2438
BE-IMH	3.0387	3.0037	3.0631	3.1721	3.0372	3.0403	3.0146	3.0411	2.9823	2.9676	3.0181	3.0160
	0.9460	0.94943	0.9473	0.9488	0.9483	0.9482	0.9500	0.9557	0.9494	0.9508	0.9489	0.9493
	11.3300	11.4820	11.6882	12.4053	11.2028	11.2163	13.8553	15.2934	11.0323	11.0502	11.0958	11.2440
BE-IMH	3.0390	3.0039	3.0630	3.1739	3.0372	3.0395	3.0140	3.0318	2.9818	2.9674	3.0177	3.00887
	0.9463	0.9496	0.9475	0.9493	0.9484	0.9486	0.9504	0.9503	0.9496	0.9506	0.9490	0.9495
	11.3318	11.4800	11.6936	12.4194	11.2086	11.2198	11.2912	11.4783	11.0329	11.0419	11.0982	11.2201

In summary, the Monte Carlo study confirms the effectiveness of the proposed predictive methodology. Both the frequentist and Bayesian approaches provide reliable predictive inference, while the Bayesian procedures exhibit stable performance regardless of the posterior sampling algorithm employed. These results support the practical applicability of the proposed prediction methods for right-censored survival data modeled by the XGamma distribution.

6. APPLICATION

In this section, we examine the performance of the ML, BE-IS, BE-IMH and BE-RWM estimators using two publicly available datasets. The first dataset, obtained from Colosimo and Giolo (2006), concerns the remission times of leukemia in $n = 103$ children. The second dataset contains survival times from patients diagnosed with laryngeal cancer, also reported in Colosimo and Giolo (2006). Both datasets are available in the R packages survival and KMsurv.

To assess predictive performance, we employ a leave-one-out cross-validation procedure restricted to uncensored observations, since the true event times are required for evaluating prediction accuracy. Specifically, for each uncensored observation $y_i = (t_i, \delta_i = 1)$, we remove y_i from the dataset, estimate the parameter θ using the remaining $n - 1$ observations, and compute a prediction $\hat{t}_i$ for the omitted lifetime t_i. Let $\hat{\theta}_{ML}^{(-i)}$, $\hat{\theta}_{SE}^{(-i)}$, $\hat{\theta}_{LI}^{(-i)}$, $\hat{\theta}_{GE}^{(-i)}$, $\tilde{\theta}_{SE}^{(-i)}$, $\tilde{\theta}_{LI}^{(-i)}$, $\tilde{\theta}_{GE}^{*(-i)}$, $\tilde{\theta}_{SE}^{*(-i)}$, $\tilde{\theta}_{LI}^{*(-i)}$, and $\tilde{\theta}_{GE}^{*(-i)}$ denote the corresponding estimators obtained from the reduced dataset.

For each i, the predictive quantity $\hat{t}_i$ is computed as the predictive mean described in Section 4.1. Out-of-sample predictive performance is then evaluated using the root mean squared prediction error (RMSPE),

$$RMSPE = \sqrt{\frac{1}{r}\sum_{i \in D}(t_i - \hat{t}_i)^2},$$

where D denotes the set of uncensored observations, $r = |D|$ is the number of observations in D, and $\hat{t}_i$ is the predicted value of t_i. Smaller RMSPE values indicate better predictive performance.

6.1 Data Set 1: Leukemia Data Set

This dataset contains remission times for $n = 103$ children diagnosed with leukemia, of which 64 (62.14%) observations are right-censored. We assume that the remission times follow an XGamma distribution with parameter θ, where $\theta \in \Theta = (0, \infty)$.

To estimate θ, we compute the maximum likelihood estimator and the Bayesian estimators obtained via the IS, IMH, and RWM algorithms. As in the simulation study, the IS algorithm is implemented using $N = 20{,}000$ samples. For the IMH and RWM algorithms, we generate $L = 120{,}000$ iterations, discard the first $B = 20{,}000$ iterations as burn-in, and retain every $J = 5$-th draw, resulting in posterior samples of size $20{,}000$ for inference.

From the RMSE-based calibration procedure, the selected tuning parameters were $a_{IS}^* = -41.3$ and $b_{IS}^* = -25.8$ for the IS algorithm, $a_{IMH}^* = -21.9$ and $b_{IMH}^* = -14.3$ for the IMH algorithm, and $a_{RWM}^* = -2{,}6$ and $b_{RWM}^* = -2.6$ for the RWM algorithm. Additional results, including illustrations of the log-likelihood and score functions highlighting the MLE, details of the calibration procedure used to select the tuning parameters a and b, and supplementary convergence diagnostics for the IMH and RWM algorithms, are provided in Appendix C of the Supplementary Material.

Table 7 presents the point estimates and the corresponding $RMSE_{KM}(\cdot)$ values obtained under each estimation procedure. Consistent with the simulation results, the point estimates are very similar across methods. Moreover, all estimators yield small $RMSE_{KM}(\cdot)$ values, indicating that the fitted XGamma survival functions closely match the Kaplan–Meier estimator. The Bayesian estimators obtained under the LINEX and generalized entropy loss functions produce the smallest $RMSE_{KM}(\cdot)$ values, although the differences relative to the remaining estimators are negligible. Furthermore, the Bayesian estimates obtained under the LINEX and generalized entropy loss functions are virtually identical across the IS, IMH, and RWM algorithms, coinciding up to four decimal places. This suggests that the resulting inference is largely insensitive to the posterior sampling scheme adopted.

Table 7
Point Estimates, Values and Interval Estimates

Results	**Method**									
	ML	IS			IMH			RWM		
	$\hat{\theta}_{ML}$	$\hat{\theta}_{SE}$	$\hat{\theta}_{LI}$	$\hat{\theta}_{GE}$	$\tilde{\theta}_{SE}$	$\tilde{\theta}_{LI}$	$\hat{\theta}^*_{GE}$	$\tilde{\theta}^*_{SE}$	$\tilde{\theta}^*_{LI}$	$\hat{\theta}^*_{GE}$
Estimate	**0.6003**	**0.5783**	**0.6041**	**0.6041**	**0.5804**	**0.6041**	**0.6041**	**0.5993**	**0.6042**	**0.6042**
$RMSE_{KM}(\cdot)$	**0.0214**	**0.0273**	**0.0213**	**0.0213**	**0.0264**	**0.0213**	**0.0213**	**0.0215**	**0.0213**	**0.0213**

Table 8 presents the 95% confidence and credible intervals for θ. All intervals exhibit substantial overlap, indicating good agreement among the competing inferential procedures. As expected, the HPD intervals are shorter than the corresponding equal-tail intervals. The IS-based intervals are the most precise, while the MLE and RWM approaches produce slightly wider intervals. Furthermore, as observed in the simulation studies, the IMH algorithm exhibited poorer mixing properties than the RWM algorithm (see Appendix B.2 of the Supplementary Material). Nevertheless, the resulting credible intervals are very similar across the Bayesian methods, suggesting that the reduced mixing efficiency of the IMH sampler has little practical impact on the final inference. Overall, the interval estimates support the conclusions obtained from the point estimates.

Table 8
Confidence and Credible Intervals

Method	**Interval**	**Lower**	**Upper**	**Length**
MLE	Wald	0.4818	0.7187	0.2369
IS	Equal-tail	0.4822	0.6355	0.1533
	HPD	0.4971	0.6383	0.1411
IMH	Equal-tail	0.4825	0.6772	0.1948
	HPD	0.4988	0.6772	0.1724
RWM	Equal-tail	0.4852	0.7231	0.2379
	HPD	0.4800	0.7157	0.2357

Figure 2 displays the Kaplan–Meier survival curve together with the fitted survival functions $S(t \mid \theta)$ corresponding to $\hat{\theta}_{ML}$, $\hat{\theta}_{SE}$, $\hat{\theta}_{LI}$, and $\hat{\theta}^{*}_{GE}$. To improve readability, the curves associated with the other estimators are omitted, since the corresponding estimates coincide with $\hat{\theta}_{LI}$ up to four decimal places. Consistent with the small $RMSE_{KM}(\cdot)$ values reported in Table 7, all fitted survival curves closely follow the Kaplan–Meier estimator throughout the observed time range, providing further evidence of the satisfactory fit of the XGamma model to the leukemia data.

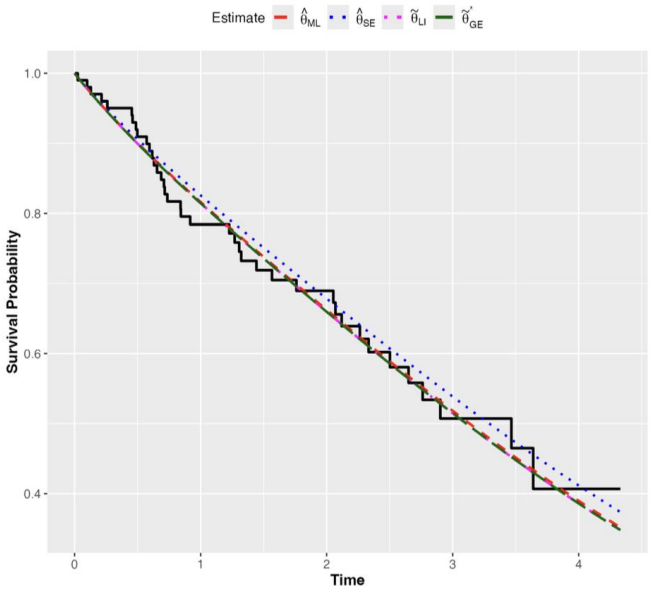


Figure 2: Kaplan-Meier Curve and Estimated Survival Curves

Table 9 summarizes the predictive performance of the estimators in terms of the root mean squared predictive error (RMSPE), coverage probability (CP), and average interval length (AL). As can be seen, the MLE, BE-IMH, and BE-RWM estimators exhibit very similar predictive performance, with a slight advantage for the MLE, which attains the smallest RMSPE value. Moreover, these methods present identical coverage probabilities ($CP \approx 0.9487$), which are very close to the nominal 95% level, indicating that the corresponding predictive intervals have similar calibration properties.

Table 9
Predictions Intervals - Equals Tails

Method	**Measures**		
2-4	RMSPR	CP	AL
MLE	2.7048	0.9487	11.0232
BE-IS	4.3742	0.8974	15.3154
BE-MH	2.7744	0.9487	11.3651
BE-RWM	2.7562	0.9487	11.3587

The BE-IS estimator exhibits a noticeably larger RMSPE and a slightly lower coverage probability. A possible explanation is the high proportion of censored observations in this dataset, which reduces the amount of information available for posterior learning and may adversely affect the accuracy of the importance-sampling approximation. Nevertheless, the predictive intervals produced by BE-IS remain informative, with an empirical coverage probability of approximately 90%, despite falling below the nominal 95% level.

Regarding the average interval length, the MLE method produced the shortest predictive interval, indicating greater efficiency. The slightly wider intervals obtained from the Bayesian approaches reflect a more conservative behavior in uncertainty quantification. Overall, the results indicate a slight advantage for the maximum likelihood estimator in terms of both predictive accuracy and interval efficiency. However, the Bayesian methods remain highly competitive, yielding similar performance and demonstrating stability in predictive inference.

6.2 Data set 2: Laryngeal Cancer Data Set

This dataset consists of survival times (in months) for $n = 90$ patients diagnosed with laryngeal cancer, of whom 40 are right-censored. To improve numerical stability, the survival times were converted to years by dividing the original observations by 12. This transformation preserves the distributional properties of the data while facilitating numerical computation.

Table 10 reports the point estimates and the corresponding $RMSE_{KM}(\cdot)$ values obtained by the maximum likelihood and Bayesian procedures. The selected values of the tuning parameters associated with the LINEX and generalized entropy loss functions are reported in Appendix C.2 of the Supplementary Material. As in the previous application, the point estimates are virtually identical across all methods, yielding an estimated value of $\theta \approx 2.50$. Moreover, the $RMSE_{KM}(\cdot)$ values coincide up to four decimal places, indicating that the estimated survival functions closely approximate the Kaplan–Meier curve, with no practically significant differences among the estimators.

Table 10
Point Estimates and Values

Results	**Method**									
	ML	IS			IMH			RWM		
	$\hat{\theta}$	$\hat{\theta}_{SE}$	$\hat{\theta}_{LI}$	$\hat{\theta}_{GE}$	$\tilde{\theta}_{SE}$	$\tilde{\theta}_{LI}$	$\hat{\theta}^*_{GE}$	$\tilde{\theta}^*_{SE}$	$\tilde{\theta}^*_{LI}$	$\hat{\theta}^*_{GE}$
Estimate	**2.4861**	**2.5110**	**2.5010**	**2.5004**	**2.5009**	**2.4973**	**2.5009**	**2.4857**	**2.5009**	**2.5008**
$RMSE_{KM}(\cdot)$	**0.0280**	**0.0280**	**0.0280**	**0.0280**	**0.0280**	**0.0280**	**0.0280**	**0.0280**	**0.0280**	**0.0280**

Table 11 reports the corresponding confidence and credible intervals. All intervals are very similar, providing additional evidence of the stability of the inferential procedures. The differences in interval length are small and do not affect the overall inferential conclusions.

These results are further illustrated in Figure 11, which displays the Kaplan–Meier survival curve together with the fitted survival functions corresponding to $\theta \in \{\hat{\theta}_{ML}, \hat{\theta}_{SE}, \tilde{\theta}_{LI}, \tilde{\theta}_{GE}\}$. The estimated curves are virtually indistinguishable and closely follow the Kaplan–Meier estimator, providing further evidence of the excellent fit of the XGamma model to these data.

Table 11
Confidence and Credible Intervals

Method	Interval	Lower	Upper	Length
MLE	Wald	1.9436	3.0287	1.0851
IS	Equal-tail	200421	0.0657	1.0235
	HPD	2.0421	2.9721	0.9300
IMH	Equal-tail	2.0032	3.0724	1.0693
	HPD	1.9842	2.0908	0.9967
RWM	Equal-tail	1.9796	3.0591	1.0795
	HPD	1.9540	3.0242	1.0702

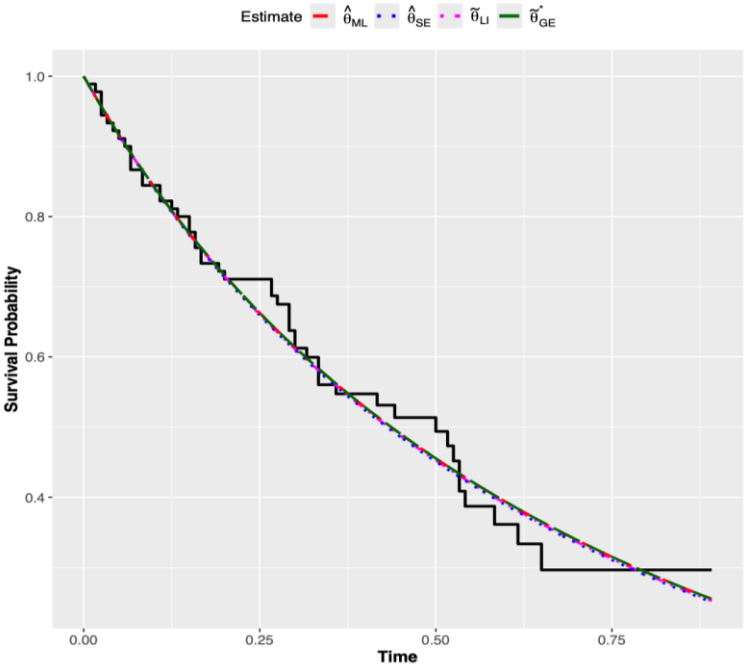


Figure 3: Kaplan-Meier Curve and Estimated Survival Curves

Table 12 summarizes the predictive performance of the estimators. All methods achieve the same coverage probability ($CP = 0.96$), indicating well-calibrated predictive intervals. The BE-IS estimator attains the smallest RMSPE and average interval length (AL), suggesting a modest advantage in predictive accuracy and efficiency. Overall, the differences among the methods are small.

Table 12
Predictions Measures

Method	Measures		
	RMSPR	**CP**	**AL**
MLE	0.4372	0.96	2.2677
BE-IS	0.3298	0.96	1.7788
BE-IMH	0.4445	0.96	2.3342
BE-RWM	0.4458	0.96	2.3522

Final Remarks

In this article, we investigated maximum likelihood and Bayesian inference for the XGamma distribution in the presence of right-censored survival data. We established the existence and uniqueness of the maximum likelihood estimator and considered Bayesian estimation under squared-error, LINEX, and generalized entropy loss functions. Since closed-form expressions for the estimators are not available, numerical procedures based on importance sampling (IS), independent Metropolis–Hastings (IMH), and random-walk Metropolis (RWM) algorithms were developed and implemented.

The simulation studies showed that all inferential procedures produced accurate estimates of the model parameter and achieved satisfactory interval estimation performance. As expected, estimation accuracy improved with increasing sample size and deteriorated with higher censoring proportions. The maximum likelihood estimator and the Bayesian estimator under squared-error loss generally yielded the smallest absolute estimation errors, whereas the Bayesian estimators under LINEX and generalized entropy loss functions often provided smaller RMSE values with respect to the Kaplan–Meier estimator. Overall, the differences among the IS, IMH, and RWM implementations were negligible from an inferential perspective, indicating that all three algorithms provide reliable approximations to the posterior distribution.

The predictive analysis also revealed similar performance among the competing methods. In particular, the Bayesian predictive procedures achieved coverage probabilities comparable to those of the maximum likelihood approach, while in some scenarios the IS-based procedure produced slightly shorter predictive intervals and lower predictive errors. These results suggest that the proposed methods provide reliable predictive inference for future survival observations.

The two real-data applications further confirmed the practical usefulness of the XGamma distribution. In both datasets, the maximum likelihood and Bayesian procedures produced very similar point and interval estimates, leading to fitted survival curves that closely matched the Kaplan–Meier estimator. Moreover, the Bayesian procedures based on

IS, IMH, and RWM yielded nearly identical inferential conclusions, despite differences in computational efficiency.

Overall, the results indicate that the XGamma distribution provides a flexible and effective model for right-censored survival data and that both maximum likelihood and Bayesian approaches constitute viable inferential alternatives. The choice among the Bayesian estimators should be guided by the analyst's objective, since different loss functions emphasize different aspects of estimation accuracy and predictive performance.

Future research may extend the proposed framework to other censoring mechanisms, alternative prior specifications, and hierarchical Bayesian formulations in which the asymmetry parameters of the LINEX and generalized entropy loss functions are estimated jointly with the model parameter. Another promising direction is the assessment of the robustness of the proposed procedures under model misspecification and competing lifetime distributions.

All simulation studies and real-data applications were carried out in the R software environment. To ensure full reproducibility, the complete source code used to generate the numerical results, tables, and figures has been made freely available through a public GitHub repository (https://github.com/erlandsonsaraiva09/XGamma-RightCensures).

ACKNOWLEDGMENTS

The authors are grateful for the financial support received from the Brazilian institution CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Finance Code 001, process number 402154/2023-1). The first author acknowledges FUNDECT (Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, TO number 120/2024, SIAFIC 818, process number 83/026.835/2024). The third author acknowledges FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, process number 2025/10988-0). The first two authors also extend their appreciation to the Federal University of Mato Grosso do Sul and its Institute of Mathematics for their support. The last two authors acknowledge the Institute of Mathematical and Computer Sciences (ICMC), University of São Paulo (USP), and the Department of Statistics (DEs) of the Federal University of São Carlos (UFSCar) for their support.

The authors would like to express their sincere gratitude to the Editor and the two anonymous reviewers for their careful reading of the manuscript and for their valuable comments and suggestions, which substantially improved the quality and clarity of this work.

REFERENCES

1. Assane, C.C., Pereira, B.D.B. and Pereira, C.A.D.B. (2018). Bayesian significance test for discriminating between survival distributions. Communications in Statistics-Theory and Methods, 47(24), 6095-6107.
2. Balakrishnan, N. and Aggarwala, R. (2000). Progressive Censoring. Theory, Methods and Applications. Birkhauser, Boston.

3. Banerjee, S. and Carlin, B.P. (2014). Hierarchical Modeling and Analysis for Spatial Data. 2nd ed., Chapman & Hall/CRC, Boca Raton.
4. Basu, A.P. and Ebrahimi, N. (1991). Bayesian Approach to Life Testing and Reliability Estimation Using Asymmetric Loss Functions. Journal of Statistical Planning and Inference, 29(1), 21-31.
5. Bernardo, J.M. and Smith, A.F.M. (1994). Bayesian Theory. John Wiley & Sons, Chichester.
6. Calabria, R. and Pulcini, G. (1996). Point Estimation Under Asymmetric Loss Functions for Left-Truncated Exponential Samples. Communications in Statistics – Theory and Methods, 25(3), 585-600.
7. Casella, G. and Berger, R.L. (2002). Statistical Inference. 2nd ed. Duxbury, Pacific Grove, CA.
8. Chen, M.H. and Shao, Q.M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. Journal of Computational and Graphical Statistics, 8, 69-92.
9. Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. The American Statistician, 49, 327-335.
10. Colosimo, E.A. and Giolo, S.R. (2024). Análise de sobrevivência aplicada. Blücher, São Paulo.
11. Crowther, M.J. and Lambert, P.C. (2019). Survival analysis and interpretation of time-to-event data: The tortoise and the hare. Journal of the American Medical Association (JAMA), 321(14), 1401-1402.
12. Dey, D.K. and Liu P.S.L. (1992). On comparison of estimators in a generalized life model Microelectron. Microelectronics Reliability, 32(1-2), 207-221.
13. Dey, D.K., Gosh, M. and Srinivasan, C. (1987). Simultaneous estimation of parameters under entropy loss. Journal of Statistical Planning and Inference, 15, 347-363.
14. Fletcher, R. and Reeves, C.M. Function minimization by conjugate gradients. Computer Journal, 7, 148-154.
15. Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. Statistical Science, 7, 457-511.
16. Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (1995). Bayesian data analysis. Vol. 2, FL: CRC press Boca Raton.
17. Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008). Lindley Distribution and Its Application. Mathematics and Computers in Simulation, 78(4), 493-506.
18. Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996). Markov Chain Monte Carlo in Practice. Chapman and Hall: London, UK.
19. Hastings, W.K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. Biometrika, 57, 97-109.
20. Jaheen, Z.F. (2005). Bayesian Estimation of the Parameters of the Generalized Exponential Distribution Under Asymmetric Loss Functions. Communications in Statistics – Theory and Methods, 34(4), 1015-1028.
21. Kloek, T. and van Dijk, H.K. (1978). Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo. Econometrica, 46(1), 1-19.
22. Lawless, J.F. (2003). Statistical Models and Methods for Life time Data. Wiley, New York.
23. Li, Y., Lee, J. and Kottas, A. (2022). Bayesian nonparametric Erlang mixture modeling for survival analysis. Computational Statistics & Data Analysis, 188, 107874.

24. Lindley, D.V. (1958). Fiducial Distributions and Bayes' Theorem. Journal of the Royal Statistical Society, Series B, 20(1), 102-107.
25. Mohammed, Y.A., Yatim, B. and Ismail, S. (2014). Mixture Model of the Exponential, Gamma and Weibull Distributions to Analyse Heterogeneous Survival Data. Journal of Scientific Research and Reports, 4(5), 439-453.
26. Nelder, J.A. and Mead, R.A simplex algorithm for function minimization. Computer Journal, 7, 308-313.
27. Nelson, W. (2009). Accelerated Testing: Statistical Models, Test Plans, and Data Analysis. Wiley Series in Probability and Statistics. John Wiley & Sons.
28. R Core Team (2026). R: A Language and Environment for Statistical Computing. Available online: https://www.R-project.org
29. Saraiva, E.F. and Suzuki, A.K. (2017). Bayesian computational methods for estimation of two-parameters Weibull distribution in presence of right-censored data. Chilean Journal of Statistics, 8(2), 25-43.
30. Sen, S., Maiti, S.S. and Chandra, N. (2016). The xgamma Distribution: Statistical Properties and Application. Journal of Modern Applied Statistical Methods, 15(1), DOI: 10.22237/jmasm/1462077420
31. Singh, S.K. and Singh, U. (2013). Bayesian Estimation of Weibull Parameters Under General Entropy Loss Function. Journal of Probability and Statistics, 2013, Article ID 589640.
32. Srinivasan, R. Importance sampling: Applications in communications and detection. Berlin: Springer-Verlag, 2002.
33. Tierney, L. (1994). Markov chains for exploring posterior distributions. The Annals of Statistics, 22(4), 1701-1762. doi:10.1214/aos/1176325750.
34. Varian, H.R. (1975). A Bayesian Approach to Real Estate Assessment. In: L. J. Savage, S. E. Feinberg and A. Zellner, Eds., Studies in Bayesian Econometrics and Statistics: In Honor of L. J. Savage. North-Holland Pub. Co., Amsterdam, pp. 195-208.
35. Zellner, A. (1986). A Bayesian Estimation and Prediction Using Asymmetric Loss Function. Journal of the American Statistical Association, 8, 446-451.