Statistics and Its Interface Volume 13 (2020) 139–149

Reference Bayesian analysis for the generalized lognormal distribution with application to survival data

Vera L. D. Tomazella*, Sandra R. de Jesus, Francisco Louzada, Saralees Nadarajah†, and Pedro L. Ramos‡

This paper proposes a reference Bayesian approach for the estimation of the parameters of the generalized lognormal distribution in the presence of survival data. It is shown that the reference prior leads to a proper posterior distribution while the Jeffreys prior leads to an improper posterior. Simulation studies were performed to analyze the frequentist properties of credible intervals from the reference posterior distribution, considering complete and censored data. The proposed methodology is illustrated using two real datasets.

AMS 2000 subject classifications: Primary 62F15, 62N01; secondary 62N05.
Keywords and phrases: Bayesian inference, Generalized lognormal distribution, Lognormal distribution, Reference prior, Jeffreys prior, Censored data.

1. INTRODUCTION

The lognormal distribution has been used in many different aspects of life sciences to model positive skewness data. Most aspects include, for example, reliability and survival analysis [14, 11, 8, 9, 2]. However, when the data come from a distribution with tails having a significant degree of positive skewness, the lognormal distribution does not sufficiently model the data. A good alternative is to consider models with more flexible tails relative to the lognormal distribution, such as the generalized lognormal distribution (logGN), which includes the lognormal distribution as a special case [19].

The logGN distribution is considered a tool for obtaining robust estimates and it represents a viable alternative to analyze data that adhere to a lognormal distribution, due to its flexibility [18]. This distribution has been used in different contexts, with different parameterizations, but the inferential procedures for the parameters of the distribution present problems. In the classical context, the parameterization that considers the threshold or location parameter introduces an unusual feature in the likelihood function. Hill [14] showed that the threshold parameter tends to the smallest observed value of the response variable and that the maximum likelihood estimates of the remaining parameters are inconsistent. Other references to the estimation problems in the classical context include [13, 10, 23, 17].

In the Bayesian context, the methods presented in the literature for the estimation of the parameters of the logGN distribution are restricted and applied to particular cases. Hill [14] considered subjective priors. Upadhyay and Peshwani [25] used the Jeffreys first rule as a noninformative prior. However, Jeffreys himself criticized his first rule (see Kass and Wasserman [16]) since the proposed prior is invariant only to power transformations and is not obtained by formal rules. Martín and Pérez [18] considered an invariant Jeffreys prior to perform inference on the parameters of interest. Despite the fact that the Jeffreys prior is not adequate in many situations and can lead to marginalization paradoxes and strong inconsistencies (see Bernardo [5, pg. 41] and the references therein), the authors did not prove that the posterior obtained is proper (see Ramos et al. [21] for a detailed discussion). Here, we show that the posterior obtained using the Jeffreys prior is improper and should not be used in a Bayesian analysis.

To overcome this problem, we consider estimation of parameters using reference priors introduced by Bernardo [4] and [3, 5]. This method describes a way to find an objective prior that maximizes the lack of information. In this case, the estimation process depends only on the distribution used and the observed data. The reference prior produces a posterior reference distribution with important properties such as generality, invariance, consistent marginalization and consistent sampling properties [5]. We prove that the reference posterior obtained is proper and can be used to obtain posterior estimates. Finally, while the above references consider estimation for only complete data, we extend the proposed inference for censored data.

This paper is organized as follows. Section 2 presents the logGN distribution, its survival and hazard functions, and likelihood function for complete data. Sections 3 and 4 present prior reference distributions and Jeffreys priors, respectively, for the logGN distribution. Section 5 considers

*First supporter of the project.
†Corresponding author.
‡Pedro L. Ramos acknowledges support from to the São Paulo State Research Foundation (FAPESP Proc. 2017/25971-0).

the presence of censored data. In Section 6, some model selection criteria are presented. Sections 7 and 8 present a simulation study and some applications to real datasets, respectively.

2. GENERALIZED LOGNORMAL DISTRIBUTION

The logGN distribution was introduced by Vianelli [27, 28] as a family of distributions having multiplicative errors of order r. Other parameterizations can be found in the literature; see, for example, Hill [14] and Chen [8]. Here, the logGN distribution is obtained through the exponential transformation of a random variable that follows the generalized normal distribution studied by Nadarajah [19].

It is understood that the random variable X has a logGN distribution if its probability density function (pdf) is given by

$$f(x|\mu,\tau,s) = \frac{s}{2\,x\,\tau\Gamma\left(\frac{1}{s}\right)} \exp\left\{-\left|\frac{\log x-\mu}{\tau}\right|^{s}\right\} \tag{1}$$

with $x > 0,\ \mu \in \Re,\ \tau,\ s > 0$. The logGN distribution is the distribution of $X = \exp(Y)$ when Y is a random variable following the generalized normal distribution.

The logGN distribution contains the lognormal distribution as a particular case when $s = 2$ and τ is replaced by $\sqrt{2}\,\tau$. The log-Laplace distribution is the particular case for $s = 1$.

Martín and Pérez [18] argued that the logGN distribution presents better fits to data relative to the lognormal distribution when $s \in (1, 2) \cup (2, 3)$. The logGN distribution allows for more flexible kurtosis than the lognormal distribution. Furthermore, the capacity of the logGN distribution to provide a precise fit to the data depends on its shape.

Zhu and Zinde-Walsh [29] proposed a reparameterization of the asymmetric exponential power distribution that allows one to observe the effect of the shape parameter. Using a similar reparameterization, $\sigma = \tau\,\Gamma\left(1+\frac{1}{s}\right)$, the logGN distribution in (1) can be written as

$$f(x|\boldsymbol{\theta}) = 2^{-1}\,x^{-1}\,\sigma^{-1}\exp\left\{-\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right)^{s}\right\}, \tag{2}$$

where $x > 0,\ \mu \in \Re,\ \sigma,\ s > 0$. The corresponding survival and hazard functions are

$$S(x|\boldsymbol{\theta}) = \begin{cases} 1 - \dfrac{\Gamma\left(\frac{1}{s}, \left(\frac{\Gamma\left(1+\frac{1}{s}\right)(\mu-\log(x))}{\sigma}\right)^{s}\right)}{2\,s\,\Gamma\left(1+\frac{1}{s}\right)}, & \text{if} \quad 0 < x \le e^{\mu}, \\[2ex] \dfrac{\Gamma\left(\frac{1}{s}, \left(\frac{\Gamma\left(1+\frac{1}{s}\right)(\log(x)-\mu)}{\sigma}\right)^{s}\right)}{2\,s\,\Gamma\left(1+\frac{1}{s}\right)}, & \text{if} \quad x > e^{\mu}, \end{cases}$$

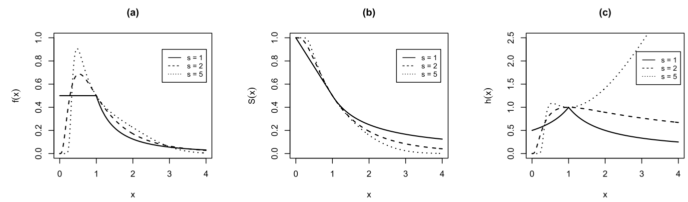

Figure 1. (a) Density function; (b) Survival function; (c) Hazard function.

and

$$h(x|\boldsymbol{\theta}) = \begin{cases} \dfrac{s\,\Gamma\left(1+\frac{1}{s}\right)\exp\left\{-\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right)^{s}\right\}}{x\sigma\left\{2s\Gamma\left(1+\frac{1}{s}\right)-\Gamma\left(\frac{1}{s},\left(\frac{\Gamma\left(1+\frac{1}{s}\right)(\mu-\log(x))}{\sigma}\right)^{s}\right)\right\}}, & \text{if } 0<x\le e^{\mu}, \\[2ex] \dfrac{s\,\Gamma\left(1+\frac{1}{s}\right)\exp\left\{-\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right)^{s}\right\}}{x\;\sigma\;\Gamma\left(\frac{1}{s},\left(\frac{\Gamma\left(1+\frac{1}{s}\right)(\log(x)-\mu)}{\sigma}\right)^{s}\right)}, & \text{if } x > e^{\mu}, \end{cases}$$

respectively, where $\Gamma(\cdot,\cdot)$ denotes the upper incomplete gamma function.

Figure 1 plots the density function in (2), the survival function and the hazard function for different values of s, assuming $\mu = 0$ and $\sigma = 1$.

2.1 Likelihood for complete data

Let $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ be a random sample from the distribution in (2). The likelihood function is given by

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = 2^{-n}\exp\left\{-\sum_{i=1}^{n}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^{s}\right\} \times \sigma^{-n}\prod_{i=1}^{n}x_i^{-1}. \tag{3}$$

The log-likelihood function is given by

$$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = -n\log 2 - n\log\sigma - \sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^{s}. \tag{4}$$

The first-order derivatives of the log-likelihood function for a single observation are

$$\frac{\partial\log L}{\partial\mu} = \frac{s\,\Gamma\left(1+\frac{1}{s}\right)}{\sigma}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right)^{s-1} \times \operatorname{sign}(\log(x)-\mu), \tag{5}$$

$$\frac{\partial\log L}{\partial\sigma} = -\frac{1}{\sigma} + \frac{s}{\sigma}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right)^{s} \tag{6}$$

and

(7)

$$\frac{\partial \log L}{\partial s} = -\left[\log\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right) - \frac{\Psi\left(1+\frac{1}{s}\right)}{s}\right] \times \left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x)-\mu|}{\sigma}\right)^{s},$$

where $\Psi(s) = \frac{\Gamma^{'}(s)}{\Gamma(s)}$ denotes the digamma function.

The Fisher information matrix is obtained by the score function. This matrix is useful to get the reference *priors* for the model parameters. The elements of the Fisher information matrix for the distribution in (4) are obtained by the following proposition.

Proposition 2.1. *Let $I(\theta)$ denote the Fisher information matrix, with $\boldsymbol{\theta} = (\boldsymbol{\theta})$. The elements of the Fisher information matrix,*

$$I_{ij}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 \log f(x|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right) = E\left(\frac{\partial \log f(x|\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial \log f(x|\boldsymbol{\theta})}{\partial\theta_j}\right),$$

where $i, j = 1, 2, 3$, with $I_{ij}(\boldsymbol{\theta}) = I_{ji}(\boldsymbol{\theta})$ and θ_j the jth element of $\boldsymbol{\theta}$ are given by

$$I_{11}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L}{\partial \mu}\right)^2\right] = \frac{n\Gamma\left(\frac{1}{s}\right)\Gamma\left(2-\frac{1}{s}\right)}{\sigma^2},$$

$$I_{12}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L}{\partial \mu}\right)\left(\frac{\partial \log L}{\partial \sigma}\right)\right] = 0,$$

$$I_{13}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L}{\partial \mu}\right)\left(\frac{\partial \log L}{\partial s}\right)\right] = 0,$$

$$I_{22}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L}{\partial \sigma}\right)^2\right] = \frac{ns}{\sigma^2},$$

$$I_{23}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L}{\partial \sigma}\right)\left(\frac{\partial \log L}{\partial s}\right)\right] = -\frac{n}{\sigma s},$$

$$I_{33}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L}{\partial s}\right)^2\right] = \frac{n}{s^3}\left\{\left(1+\frac{1}{s}\right)\Psi^{'}\left(1+\frac{1}{s}\right)\right\},$$

where $s > 1$ and $\Psi'(s) = \frac{\partial}{\partial s}\Psi(s)$ denotes the trigamma function. The restriction $s > 1$ ensures that the elements $I_{ij}(\boldsymbol{\theta})$ for $i, j = 1, 2, 3$ are finite and the information matrix $I(\boldsymbol{\theta})$ is positive definite.

The proof is a consequence of Proposition 5 in Zhu and Zinde-Walsh [29]. Then, the Fisher's information matrix is given by

(8)

$$I(\boldsymbol{\theta}) = n\begin{bmatrix} \frac{\Gamma\left(\frac{1}{s}\right)\Gamma\left(2-\frac{1}{s}\right)}{\sigma^2} & 0 & 0 \\ 0 & \frac{s}{\sigma^2} & -\frac{1}{\sigma\, s} \\ 0 & -\frac{1}{\sigma\, s} & \frac{1}{s^3}\left\{\left(1+\frac{1}{s}\right)\Psi^{'}\left(1+\frac{1}{s}\right)\right\} \end{bmatrix}.$$

The corresponding inverse Fisher's information matrix is given by

(9)

$$B(\boldsymbol{\theta}) = \frac{1}{n}\begin{bmatrix} \frac{\sigma^2}{\Gamma(\frac{1}{s})\Gamma(2-\frac{1}{s})} & 0 & 0 \\ 0 & \frac{\sigma^2}{s\left[1-\frac{s}{(1+s)\Psi^{'}\left(1+\frac{1}{s}\right)}\right]} & \frac{\sigma s^2}{-s+(1+s)\Psi^{'}\left(1+\frac{1}{s}\right)} \\ 0 & \frac{\sigma s^2}{-s+(1+s)\Psi^{'}\left(1+\frac{1}{s}\right)} & \frac{s^4}{-s+(1+s)\Psi^{'}\left(1+\frac{1}{s}\right)} \end{bmatrix}.$$

3. REFERENCE ANALYSIS IN THE MULTIPARAMETER CASE

Reference analysis was introduced by Bernardo [4] and [3, 5]. The reference prior maximizes the lack of information on the quantity of interest. An important characteristic of the Berger-Bernardo method is the different treatment given to parameters of interest and nuisance parameters. The construction of the reference prior, in the presence of nuisance parameters, must be made using an ordered parameterization. The parameter of interest is selected and the procedure below is followed.

Corollary 3.1. *Let $f(x|\phi, \boldsymbol{\lambda}), x \in X, \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m), \phi \in \Phi, \boldsymbol{\lambda} \in \Lambda = \prod_{j=1}^{m}\Lambda_j$ be a probability model with $m + 1$ real-valued parameters and let ϕ be the quantity of interest. Suppose that the posterior distribution of $(\phi, \lambda_1, \ldots, \lambda_m)$ is asymptotically normal with covariance matrix $B(\widehat{\phi}, \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m)$, where $H = B^{-1}$, B_j is the upper left $j \times j$ submatrix of B, $H_j = B_j^{-1}$, and $h_{jj}(\phi, \lambda_1, \ldots, \lambda_m)$ is the lower right element of H_j. If the nuisance parameter spaces Λ_i do not depend on $\{\phi, \lambda_1, \ldots, \lambda_{i-1}\}$, and the functions $h_{11}, h_{22}, \ldots, h_{mm}, h_{m+1,m+1}$ factorize in the form*

$$h_{11}^{\frac{1}{2}}(\phi, \boldsymbol{\lambda}) = f_0(\phi)g_0(\lambda_1, \ldots, \lambda_m)$$

and

$$h_{i+1,i+1}^{\frac{1}{2}}(\phi, \boldsymbol{\lambda}) = f_i(\lambda_i)g_i(\phi, \lambda_i, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_m)$$

$i = 1, \ldots, m$, then $\pi^{\mathsf{R}}(\phi) \propto f_0(\phi)$ and

$$\pi^{\mathsf{R}}(\lambda_i|\phi, \lambda_1, \ldots, \lambda_{i-1}) \propto f_i(\lambda_i),\ \ i = 1, \ldots, m.$$

There is no need for compact approximations, even if $\pi^{\mathsf{R}}(\lambda_i|\phi, \lambda_1, \ldots, \lambda_{i-1})$ are not proper.

Under appropriate regularity conditions [5] the posterior distribution of $(\phi, \lambda_1, \ldots, \lambda_m)$ is asymptotically normal with mean $(\widehat{\phi}, \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m)$ and covariance matrix $B(\widehat{\phi}, \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m)$, where $B(\widehat{\phi}, \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m) = I^{-1}(\widehat{\phi}, \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m)$ and $B_{11}^{-\frac{1}{2}}(\phi, \lambda_1, \ldots, \lambda_m) = h_{11}^{\frac{1}{2}}(\phi, \lambda_1, \ldots, \lambda_m)$ and $I(\phi, \lambda_1, \ldots, \lambda_m)$ is the corresponding $(m+1)\times(m+1)$ Fisher information matrix. Furthermore, $H(\phi, \lambda_1, \ldots, \lambda_m) = I(\phi, \lambda_1, \ldots, \lambda_m)$ and the reference prior may be computed from the elements of $I(\phi, \lambda_1, \ldots, \lambda_m)$.

3.1 Reference analysis for the distribution parameters

The parameter vector ($\boldsymbol{\theta}$) is ordered and divided into 3 distinct groups, according to their inferential importance. We consider here the case in which μ is the parameter of interest and σ and s are the nuisance parameters. To obtain a joint reference prior for the parameters μ, σ and s, the following ordered parameterization was adopted:

$$\pi^{\mathsf{R}}(\boldsymbol{\theta}) = \pi^{\mathsf{R}}(s|\mu,\sigma)\pi^{\mathsf{R}}(\sigma|\mu)\pi^{\mathsf{R}}(\mu).$$

Consider the Fisher matrix in (8), the inverse Fisher matrix in (9) and Corollary 1. Let $H(\boldsymbol{\theta}) = I(\boldsymbol{\theta})$, it follows that

$$h_{33}(\boldsymbol{\theta}) = \frac{1}{s^3}\left\{\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right\}$$

and

$$h_{33}^{\frac{1}{2}}(\boldsymbol{\theta}) = \sqrt{\frac{1}{s^3}\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)} = f_2(s)g_2(\mu,\sigma).$$

Then,

$$\pi^{\mathsf{R}}(s|\mu,\sigma) \propto s^{-\frac{3}{2}}\left[\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right]^{\frac{1}{2}}.$$

Let $H_2(\boldsymbol{\theta}) = B_2^{-1}(\boldsymbol{\theta})$, where $B_2(\boldsymbol{\theta})$ is the upper left 2×2 submatrix of $B(\boldsymbol{\theta})$, it follows that

$$h_{22}(\boldsymbol{\theta}) = \frac{1}{\sigma^2}\; s\left[1-\frac{s}{(1+s)\Psi'\left(1+\frac{1}{s}\right)}\right]$$

and

$$h_{22}^{\frac{1}{2}}(\boldsymbol{\theta}) = \frac{1}{\sigma}\sqrt{s\left[1-\frac{s}{(1+s)\Psi'\left(1+\frac{1}{s}\right)}\right]} = f_1(\sigma)g_1(\mu,s).$$

Then,

$$\pi^{\mathsf{R}}(\sigma|\mu) \propto \frac{1}{\sigma}.$$

Finally, let $h_{11}(\boldsymbol{\theta}) = B_{11}^{-1}(\boldsymbol{\theta})$, it follows that

$$h_{11}(\boldsymbol{\theta}) = \frac{\Gamma\left(\frac{1}{s}\right)\Gamma\left(2-\frac{1}{s}\right)}{\sigma^2}$$

and

$$h_{11}^{\frac{1}{2}}(\boldsymbol{\theta}) = \sqrt{\frac{\Gamma\left(\frac{1}{s}\right)\Gamma\left(2-\frac{1}{s}\right)}{\sigma^2}} = f_0(\mu)g_0(\sigma,s).$$

Then,

$$\pi^{\mathsf{R}}(\mu) \propto 1.$$

Therefore, a joint reference prior for the ordered parameter is given by

$$\text{(10)}\qquad \pi^{\mathsf{R}}(\boldsymbol{\theta}) \propto \frac{1}{\sigma}s^{-\frac{3}{2}}\left[\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right]^{\frac{1}{2}},$$

where $\mu \in \Re, \sigma \in \Re^+$ and $s > 1$.

Using the likelihood function (4) and the reference prior (10), we obtain the joint posterior distribution for μ, σ and s, (11)

$$\pi^{\mathsf{R}}(\mu,\sigma,s|\mathbf{x}) \propto \sigma^{-(n+1)}s^{-\frac{3}{2}}\left[\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right]^{\frac{1}{2}} \times \exp\left\{-\sum_{i=1}^{n}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^{s}\right\}.$$

The posterior conditional probability densities are (12)

$$\pi^{\mathsf{R}}(\mu|\sigma,s,\mathbf{x}) \propto \exp\left\{-\sum_{i=1}^{n}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^{s}\right\},$$

(13)

$$\pi^{\mathsf{R}}(\sigma|\mu,s,\mathbf{x}) \propto \sigma^{-(n+1)}\exp\left\{-\sum_{i=1}^{n}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^{s}\right\}$$

and (14)

$$\pi^{\mathsf{R}}(s|\mu,\sigma,\mathbf{x}) \propto s^{-\frac{3}{2}}\left[\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right]^{\frac{1}{2}} \times \exp\left\{-\sum_{i=1}^{n}\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^{s}\right\}.$$

The densities in (12), (13) and (14) do not belong to any known parametric family and the estimates of the parameters of interest can only be obtained numerically. The posterior densities will be evaluated by Monte Carlo simulation in a Markov Chain (MCMC).

4. THE PROBLEM WITH THE JEFFREYS PRIOR

The Jeffreys prior is a method used to obtain an objective prior which is invariant under injective transformations. This method is proportional to the square root of the determinant of the Fisher information matrix. Despite the property of invariance, the Jeffreys prior may be improper and can lead to an improper posterior.

In the multiparameter case, where $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_m)$, the elements of the Fisher information matrix $I(\theta)$ are defined by

$$I_{ij}(\boldsymbol{\theta}) = E\left(\frac{\partial \log f(x|\boldsymbol{\theta})}{\partial \theta_i}\frac{\partial \log f(x|\boldsymbol{\theta})}{\partial \theta_j}\right) \quad i,j=1,\ldots,m.$$

The use of the Jeffreys rule in the multiparameter case is often inadequate. The assumption of a priori independence between parameters of different nature, and the separate use of Jeffreys rule for the specification of marginal distributions may give different results than what will be obtained by the Jeffreys principle. The Jeffreys prior obtained from the square root of the determinant of the Fisher information matrix of (8) is

(15)
$$\pi^{\mathsf{J}}(\boldsymbol{\theta}) \propto \left[\Gamma\left(\frac{1}{s}\right)\Gamma\left(2-\frac{1}{s}\right)\left\{\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)-1\right\}\right]^{\frac{1}{2}} \times \sigma^{-2}s^{-1}\ .$$

The prior in (15) leads to an improper posterior distribution. Below, we present definitions and lemmas, similar to those in [22] in order to prove that the prior distributions of (10) and (15) lead to a proper and improper posterior distributions.

The prior density for $(\mu,\sigma,s)\in\Omega=\Re\times(0,\infty)\times(1,\infty)$ is given by

$$\pi(\mu,\sigma,s)\propto\frac{\pi(s)}{\sigma^a},\quad a\in\Re, \tag{16}$$

where a is a hyperparameter and $\pi(s)$ is the marginal prior of the shape parameter, for several choices of $\pi(s)$ and a.

The Jeffreys prior and reference prior are of the form (16) with

$$a=1 \text{ and } \pi^{\mathsf{R}}(s)\propto s^{-\frac{3}{2}}\left[\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right]^{\frac{1}{2}} \tag{17}$$

and

(18)
$$\pi^{\mathsf{J}}(s) \propto s^{-1}\left[\Gamma\left(\frac{1}{s}\right)\Gamma\left(2-\frac{1}{s}\right)\left\{\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)-1\right\}\right]^{\frac{1}{2}}$$

with $a=2$.

The posterior distribution associated with the prior in (16) is proper if and only if

$$\int_1^\infty L(s|\boldsymbol{x})\pi(s)ds<\infty, \tag{19}$$

where $L(s|\boldsymbol{x})$ is the integrated likelihood for s given by

$$\int_\Re\int_0^\infty L(\mu,\sigma,s|\boldsymbol{x})\,\sigma^{-a}\,d\sigma\,d\mu. \tag{20}$$

Lemma 4.1. *The marginal prior for s given in equations (17) and (18) is a continuous function in $[1,\infty]$. As $s\to\infty$, we have $\pi^{\mathsf{R}}(s)=O(s^{-3/2})$ and $\pi^{\mathsf{J}}(s)=O(s^{-1})$.*

Proof. A direct inspection of (17) and (18) shows continuity in $[1,\infty]$. As $s\to\infty$, we have that $\Gamma\left(\frac{1}{s}\right)=O(s)$ and $\Psi'\left(1+\frac{1}{s}\right)\to 1.6449$. □

Lemma 4.2. *For $n>2-a$, the likelihood for s under the class of priors (16) is a continuous function in $[1,\infty]$ and $L(s|\boldsymbol{x})=O(1)$ as $s\to\infty$.*

Proof. See Appendix A. □

Proposition 4.3. *The reference prior given in (17) yields a proper posterior distribution and the Jeffreys prior in (18) leads to an improper posterior distribution.*

Proof. See Appendix B. □

5. PRESENCE OF CENSORED DATA

Let $\boldsymbol{x}$ denote a random sample with complete and censored survival times. The sample is divided into two sets, $\boldsymbol{x}_o=(x_1,\ldots,x_r)'$ containing r uncensored observations, and $\boldsymbol{x}_c=(x_{r+1},\ldots,x_n)'$ containing $n-r$ censored observations, therefore $\boldsymbol{x}=\boldsymbol{x}_o\cup\boldsymbol{x}_c$. The likelihood function for μ,σ and s is given by

$$L(\mu,\sigma,s|\mathbf{x})=\prod_{i=1}^r f(x_i|\mu,\sigma,s)\prod_{j=r+1}^n S(x_j|\mu,\sigma,s),$$

where $f(x_i|\mu,\sigma,s)$ and $S(x_j|\mu,\sigma,s)$ are the density and survival functions, respectively. Thus, the likelihood function may be written as

$$L(\mu,\sigma,s|\boldsymbol{x})=\prod_{i=1}^r f(x_i|\mu,\sigma,s)\prod_{j=r+1}^n\int_{x_j^*}^\infty f(x_j|\mu,\sigma,s)dx_j,$$

where $f(x|\mu,\sigma,s)$ is given in (4).

Considering the joint reference prior distribution given in (10), we obtain the joint posterior distribution for μ, σ and s

$$\pi^{\mathsf{R}}(\mu,\sigma,s|\boldsymbol{x}_o,\boldsymbol{x}_c)\propto P_oP_c, \tag{21}$$

where

$$P_o=\sigma^{-(r+1)}s^{-\frac{3}{2}}\left[\left(1+\frac{1}{s}\right)\Psi'\left(1+\frac{1}{s}\right)\right]^{\frac{1}{2}} \times\exp\left\{-\sum_{i=1}^r\left(\frac{\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|}{\sigma}\right)^s\right\}$$

and

$$P_c=\prod_{j=r+1}^n\int_{x_j^*}^\infty f(x_j|\mu,\sigma,s)dx_j.$$

There is no closed form representation for the posterior density of (21). Furthermore, the determination of the posterior distribution is complicated due to the integral in P_c. To solve this problem, $\boldsymbol{x}_c$ is considered as a set of unknown observations [26, 25].

The posterior distributions of the parameters and the censored observations are given by

$$\pi^{\mathsf{R}}(\mu, \sigma, s|\boldsymbol{x}_o, \boldsymbol{x}_c) = \pi^{\mathsf{R}}(\mu, \sigma, s|\boldsymbol{x}) \tag{22}$$

and

$$\pi^{\mathsf{R}}(\boldsymbol{x}_c|\mu, \sigma, s, \boldsymbol{x}_o) = \pi^{\mathsf{R}}(\boldsymbol{x}_c|\mu, \sigma, s), \tag{23}$$

respectively. The expression (22) corresponds to the joint posterior distribution, given by (11), and its generation is performed by MCMC. The expression (23) corresponds to the joint distribution of independent censored observations, and its generation uses the logGN distribution lower truncated at the censored value; in other words, the generated value must be greater than the observed value of the variable censored at the time of analysis.

6. SIMULATION STUDY

This section presents some frequentist properties of Bayesian estimators, utilizing the reference prior approach for complete data and censored data. To consider the approach for complete data lifetimes were generated from the logGN distribution with sample sizes $n = (100, 300, 500)$ and $1,000$ in accordance with (4) with parameters $\mu = 1.5, \sigma = 0.5$ and $s = 2.5$. The posterior samples were generated in accordance with (11), by the Metropolis-Hastings technique through the MCMC implemented in the R software. A single chain of dimension 15,000 was considered for each parameter. A burn-in of 5,000 was adopted in order to eliminate the effect of initial values, resulting in a sample of size 10,000. The convergence of the chain was checked by the criterion proposed by Geweke [12]. For each set of simulated data, an average of the estimates of the parameters, the mean square error and the coverage probability of the 95% HPD credibility intervals were computed.

The approach for censored data was performed in two stages: The first step generated lifetimes and censored times. The lifetimes denoted by $x_i, i = 1, \ldots, n$ were generated with samples sizes $n = (100, 300, 500)$ and (1,000) in accordance with (4) with parameters $\mu = 1.5, \sigma = 0.5$ and $s = 2.5$. The censoring time c_i for i-th individual was generated in accordance with (2) with parameters $\mu = \alpha, \sigma = 0.5$ and $s = 2.5$ with α controlling the percentage of censored observations. The pair is (t_i, δ_i), where $t_i = \min(x_i, c_i)$ and δ_i is equal to 1 if $x_i \leq c_i$ and equal to 0 if $x_i > c_i$.

The second step generated posterior samples by the Metropolis-Hasting technique. New values for the censored observations were generated using the expression (23), which corresponds to the logGN distribution lower truncated at the censored value. The posterior samples were generated by expression (22) through the MCMC.

Similar to the complete data, each sample size was simulated 1,000 times and an average of the parameter estimates, the mean square error and the coverage probability of the 95% HPD credible intervals were computed.

Table 1. The coverage probability of the 95% HPD credible intervals for each sample size and each parameter of the distribution considering complete and censored data

Data	Parameters	n 100	300	500	1000
Complete	$\mu = 1.5$	92.5	93.3	94.3	94.6
	$\sigma = 0.5$	82.8	94.8	94.6	94.4
	s = 2.5	81.6	93.0	94.0	94.0
10% of censure	$\mu = 1.5$	92.6	94.7	94.5	93.7
	$\sigma = 0.5$	85.5	94.4	94.7	94.8
	s = 2.5	84.3	94.8	95.3	95.5
20% of censure	$\mu = 1.5$	91.7	94.5	94.4	96.0
	$\sigma = 0.5$	82.6	95.6	94.7	95.1
	s = 2.5	80.8	94.8	94.6	94.2
40% of censure	$\mu = 1.5$	93.4	94.4	93.4	93.2
	$\sigma = 0.5$	82.6	96.1	95.0	96.0
	s = 2.5	82.8	95.5	95.8	95.5

Table 1 shows the coverage probability of the 95% HPD credibility intervals for each sample size and parameter. The empirical convergences are similar and close to the nominal level when $n = 300$. The presence of censoring impacts negatively. Figure 2 shows the variance and mean squared error of the estimators of σ and s for different samples. As expected, as the sample size grows, the variance and the mean squared error of estimators decrease. However, this reduction is more marked for the estimator of σ in all considered situations.

7. APPLICATION

In this section the proposed methodology is applied to two real datasets, considering the distribution in the presence of complete and censored data.

Complete data. In order to illustrate the methodology presented for complete data, we consider a dataset of survival times, in months, of 184 patients with lung cancer (Overduin [20]). The goal is to compare the fits of the logGN and lognormal distributions.

In the Bayesian context there are a variety of criteria that can be adopted to select the best fit between a collection of models. This paper considers the following criteria: Deviation Information Criterion (DIC) proposed by Spiegelhalter et al. [24], the Expected Akaike Information Criterion (EAIC) proposed by Brooks [6] and the Expected Bayesian

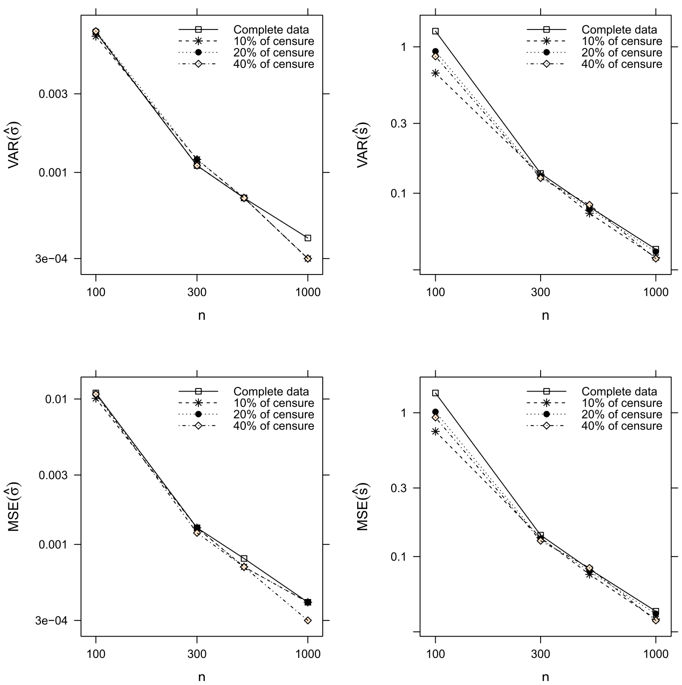

Figure 2. Variance and Mean Squared Error (MSE) of the Bayesian estimators for different samples.

Information Criterion (EBIC) proposed by Carlin and Louis [7]. The logGN and lognormal distributions were fitted to data via the Bayesian reference process. The posterior samples were generated by the Metropolis-Hastings technique, similar to the simulation study. A single chain of dimension 150,000 was considered for each parameter, discarding the first 75,000 iterations to eliminate the effect of initial values, and to avoid correlation problems, a space with a size of 15 was used, resulting in a final sample size 5,000. The convergence of the chain was verified by the Geweke criterion. Table 2 shows the posterior summaries for the parameters for both distributions and the model selection criteria. The logGN distribution is the most suitable to represent the data, as it displays better performance than the lognormal distribution for all criteria used.

Table 2. Posterior mean and 95% HDP intervals for the parameters of the model and Bayesian comparison criteria for the data on patients with lung cancer

Model	$\boldsymbol{\theta}$	Mean	HDP (95%)	DIC	EAIC	EBIC
	μ	2.87	(2.79; 2.94)			
logNG	σ	0.55	(0.43; 0.67)	1334.0	1330.2	1334.5
	s	1.62	(1.08; 2.14)			
logN	μ	2.86	(2.79; 2.93)	1378.5	1377.8	1384.3
	σ	0.44	(0.39; 0.49)			

Figure 3 shows the predicted posterior distributions for both distributions in the left panel. The estimated survival

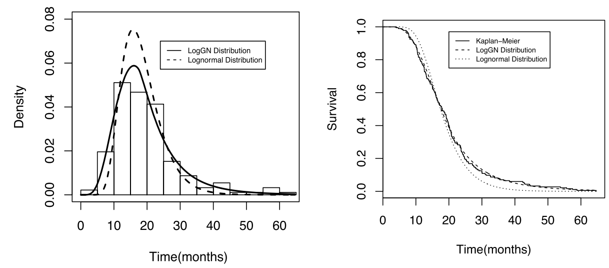

Figure 3. (a) Predicted posterior distributions of the logGN and lognormal distributions; (b) Kaplan-Meier curve with estimated survival functions of the logGN and lognormal distributions.

curves by Kaplan-Meier and by both models are shown in the right panel. We observe that the logGN distribution is the most suitable to describe the data. This is expected due to the flexibility gained by the extra parameter.

Censored data. In order to illustrate the methodology presented in Section 5 in the presence of censoring we consider a dataset from Kalbfleisch and Prentice [15]. The data refers to the study of two treatment regimes applied to 137 patients with advanced inoperable lung cancer. The survival time of the patients was measured in days. The data exhibits 9 censored observations. The goal is to fit the logGN and lognormal distributions to the dataset and to compare their efficiencies.

The logGN and lognormal distributions were fitted to the data via the Bayesian reference process. The posterior samples were generated by the Metropolis-Hastings technique, similar to the simulation study. A single chain of dimension 180,000 was considered for each parameter, discarding the first 80,000 iterations to eliminate the effect of initial values, and to avoid correlation problems, a space with a size of 20 was used, resulting in a final sample size 5,000. The convergence of the chain was confirmed using the Geweke criterion.

Table 3. Posterior mean and 95% HDP intervals for the parameters of the model and Bayesian comparison criteria for the inoperable lung cancer data

Model	$\boldsymbol{\theta}$	Mean	(95%) HDP	DIC	EAIC	EBIC
	μ	4.17	(4.07; 4.27)			
logGN	σ	1.67	(1.54; 1.81)	1500.9	1506.0	1509.8
	s	1.89	(1.63; 2.15)			
logN	μ	4.15	(4.05; 4.25)	1553.1	1556.3	1562.1
	σ	1.16	(1.10; 1.23)			

Table 3 shows the posterior summaries for the parameters of both distributions and the model selection criteria. The logGN distribution is the most suitable to represent the data, as it displays better performance than the lognormal distribution considering all the criteria. Figure 4 shows

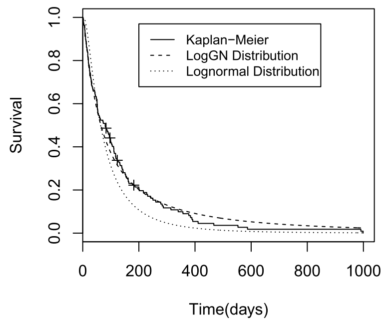

Figure 4. Kaplan-Meier curve with estimated survival functions of the logGN and lognormal distributions for the inoperable lung cancer data.

the estimated survival curves by Kaplan-Meier and by both models. The proposed distribution provides a better fit than the lognormal distribution.

8. DISCUSSION

In this paper, we have presented the logGN distribution from the standpoint of reference Bayesian analysis in the presence of survival data. The Jeffreys prior and reference prior were considered for the logGN distribution. However, Jeffreys prior leads to an improper posterior distribution and can not be used in a Bayesian analysis.

To overcome this problem we considered the reference analysis that provides a general method for finding an objective prior distribution that maximizes the lack of information. We proved that the posterior obtained is proper and can be used to deduce posterior summaries. Since the reference priors do not take into account the opinions of experts, but consider the assumed distribution and the observed data, the reference prior can conveniently be used as a reference for the posterior estimates.

The approach presented is a viable alternative to fit many types of survival data, given the flexibility of the logGN distribution and the various shapes its survival function can assume, depending on the values of the shape parameter. Simulation studies were performed to verify the adequacy of the proposed inference method, considering the presence of complete and censored data, for different sample sizes. The simulated results showed good frequentist properties even for moderate sample sizes. Real data applications showed that the logGN distribution outperformed the lognormal one, regardless of the model selection criterion.

There are a large number of possible extensions of this current work. An objective Bayesian analysis for a generalized lognormal regression model will be considered. Different Markov chain Monte Carlo techniques can be used to improve the convergence of the regression model. These techniques should be compared under different n and p. Another extension is a diagnostic analysis to assess goodness of fit from the assumption of the generalized lognormal regression model.

APPENDIX A: PROOF OF LEMMA 4.2

It is known that σ can be obtained analytically. Integrating σ, we obtain the integrated likelihood function for (μ, s),

(24)
$$L(\mu, s|\boldsymbol{x}) = \int_0^\infty L(\mu, \sigma, s|\boldsymbol{x})\pi(\sigma)d\sigma \propto s^{-1}\Gamma\left(\frac{n+a-1}{s}\right) \times \left[\sum_{i=1}^n \left(\Gamma\left(1+\frac{1}{s}\right)|\log(x_i)-\mu|\right)^s\right]^{\frac{-(n+a-1)}{s}}.$$

Considering the likelihood $L(\mu, s|\boldsymbol{x})$, we integrate μ, obtaining

$$L(s|\boldsymbol{x}) = \int_{\Re} L(\mu, s|\boldsymbol{x})\pi(\mu)d\mu \propto s^{-1}\Gamma\left(\frac{n+a-1}{s}\right)\Gamma\left(1+\frac{1}{s}\right)^{-(n+a-1)} \times \int_{\Re}\left[\sum_{i=1}^n |\log(x_i)-\mu|^s\right]^{\frac{-(n+a-1)}{s}} d\mu. \quad (25)$$

Assuming that $Y = \log(X)$, where Y is a generalized normal random variable, we have that $\int_{\Re}\left[\sum_{i=1}^n |y_i-\mu|^s\right]^{-(n+a-1)} d\mu$ is limited. Thus, $y_i = \ log(x_i)$. We define the following functions: let

$$A(\mu, s) = n\left(\max|y_i| + |\mu|\right)^s$$

and

$$B(\mu, s) = \begin{cases} n|\overline{y}-\mu^*|^s, \ if \ \mu \in L1, \\ n|\overline{y}-\mu|^s, \ if \ \mu \in L2, \end{cases}$$

where $\mu^* = \arg\min_\mu \sum_{i=1}^n |y_i-\mu|^s$, $L1 = \{\mu \in \Re : |\mu| < |\overline{y}|\}$ and $L2 = \{\mu \in \Re : |\mu| \geq |\overline{y}|\}$.

It is observed that

I) $\sum_{i=1}^n |y_i-\mu|^s \leq \sum_{i=1}^n (|y_i| + |\mu|)^s \leq \sum_{i=1}^n (\max|y_i| + |\mu|)^s = n(\max|y_i| + |\mu|)^s = A(\mu, s)$.

II) The function $|\cdot|^s$ for $s \geq 1$ is convex in $\Re$. By Jensen's inequality, we have

$$\sum_{i=1}^n |y_i-\mu|^s \geq \left|\sum_{i=1}^n (y_i-\mu)\right|^s = n\,|\overline{y}-\mu|^s.$$

If $\mu \in L1$ it follows that $\sum_{i=1}^n |y_i-\mu|^s \geq n|\overline{y}-\mu^*|^s$. If $\mu \in LS$ we have $\sum_{i=1}^n |y_i-\mu|^s \geq n|\overline{y}-\mu|^s$.

Thus, $\sum_{i=1}^{n}|y_i-\mu|^s \geq B(\mu,s)$. Considering I and II, it follows that $B(\mu,s) \leq \sum_{i=1}^{n}|y_i-\mu|^s \leq A(\mu,s)$. Therefore,

$$\int_{\Re}[A(\mu,s)]^{-\frac{(n+a-1)}{s}}d\mu \leq \int_{\Re}\left[\sum_{i=1}^{n}|y_i-\mu|^s\right]^{-\frac{(n+a-1)}{s}}d\mu \leq \int_{\Re}[B(\mu,s)]^{-\frac{(n+a-1)}{s}}d\mu.$$

We calculate the integral above on the left and right hand sides. For the integral on the left hand side, we have

$$\int_{\Re}[A(\mu,s)]^{-\frac{(n+a-1)}{s}}d\mu = \int_{\Re}[n(\max|y_i|+|\mu|)^s]^{-\frac{(n+a-1)}{s}}d\mu = 2\,n^{-\frac{(n+a-1)}{s}}\int_0^{\infty}(\max|y_i|+\mu)^{-(n+a-1)}d\mu.$$

Let $z=\max|y_i|+\mu$, then $dz=d\mu$. Thus, we have that

$$\int_{\Re}[n(\max|y_i|+|\mu|)^s]^{-\frac{(n+a-1)}{s}}d\mu = 2\left.\frac{n^{-\frac{(n+a-1)}{s}}z^{-(n+a-2)}}{-(n+a-2)}\right|_{\max|y_i|}^{\infty} = 2n^{-\frac{(n+a-1)}{s}}\left[(n+a-2)(\max|y_i|)^{n+a-2}\right]^{-1}. \tag{26}$$

Therefore,

$$\int_{\Re}[n(\max|y_i|+|\mu|)^s]^{-\frac{(n+a-1)}{s}}d\mu = n^{-\frac{(n+a-1)}{s}}f_1(y), \tag{27}$$

where $f_1(y)=\frac{2}{(n+a-2)(\max|y_i|)^{n+a-2}}$ does not depend on s.

For the integral on the right hand side, we have

$$\int_{\Re}[B(\mu,s)]^{-\frac{(n+a-1)}{s}}d\mu = \int_{L1}[n|\overline{y}-\mu^*|^s]^{-\frac{(n+a-1)}{s}}d\mu + \int_{L2}[n|\overline{y}-\mu|^s]^{-\frac{(n+a-1)}{s}}d\mu = n^{-\frac{(n+a-1)}{s}}\left[\int_{L1}|\overline{y}-\mu^*|^{-(n+a-1)}d\mu + \int_{L2}|\overline{y}-\mu|^{-(n+a-1)}d\mu\right]. \tag{28}$$

Therefore,

$$\int_{\Re}[B(\mu,s)]^{-\frac{(n+a-1)}{s}}d\mu = n^{-\frac{(n+a-1)}{s}}f_2(y), \tag{29}$$

where $f_2(y)=\int_{L1}|\overline{y}-\mu^*|^{-(n+a-1)}d\mu+\int_{L2}|\overline{y}-\mu|^{-(n+a-1)}d\mu$ does not depend on s.

Considering (26) and (28), we have

$$f_1(y) \leq n^{-\frac{(n+a-1)}{s}}\int_{\Re}\left[\sum_{i=1}^{n}|y_i-\mu|^s\right]^{-\frac{(n+a-1)}{s}}d\mu \leq f_2(y). \tag{30}$$

Therefore,

$$\int_{\Re}\left[\sum_{i=1}^{n}|y_i-\mu|^s\right]^{-\frac{(n+a-1)}{s}}d\mu = O\left(n^{-\frac{(n+a-1)}{s}}\right). \tag{31}$$

It is known that $y_i=\log(x_i)$. Therefore,

$$\int_{\Re}\left[\sum_{i=1}^{n}|\log(x_i)-\mu|^s\right]^{-\frac{(n+a-1)}{s}}d\mu = O\left(n^{-\frac{(n+a-1)}{s}}\right). \tag{32}$$

The above result will allow the study of the behavior of the integrated likelihood for s. Inserting the resulting value from (30) into (24), we have

$$L(s|\boldsymbol{x}) \propto s^{-1}\Gamma\left(\frac{n+a-1}{s}\right)\Gamma\left(1+\frac{1}{s}\right)^{-(n+a-1)} \times O\left(n^{-\frac{(n+a-1)}{s}}\right). \tag{33}$$

To study the behavior of the integrated likelihood for s, we consider the following result: $\frac{1}{\Gamma(z)} \approx z$ as z approaches 0 [1]. Therefore, as $s\to\infty$, we have $\Gamma\left(\frac{1}{s}\right)\approx s$. Moreover, considering the expansion of the first order Taylor series of $\log\Gamma(1+z)$ for values near $z=0$, we have $\log\Gamma(1+z)\approx z\Psi(1)$, where $\Psi(1)\approx -0.57721$. Thus, $\Gamma\left(1+\frac{1}{s}\right)\approx e^{\frac{\Psi(1)}{s}}$ for large values of s. So, as $s\to\infty$, we have $\Gamma\left(\frac{n+a-1}{s}\right)\approx\frac{s}{n+a-1}$. Therefore,

$$L(s|\boldsymbol{x}) \approx s^{-1}\frac{s}{n+a-1}\left(e^{\frac{\Psi(1)}{s}}\right)^{-(n+a-1)}O\left(n^{-\frac{(n+a-1)}{s}}\right) \approx e^{\frac{-\Psi(1)(n+a-1)}{s}}O\left(n^{-\frac{(n+a-1)}{s}}\right) = O\left(e^{\frac{-(n+a-1)}{s}\{\Psi(1)+\log n\}}\right)=O(1). \tag{34}$$

The proof is complete. □

APPENDIX B: PROOF OF PROPOSITION 4.3

The proof will consider the condition given in (18) and Lemmas 4.1 and 4.2.

Considering the reference prior given in Lemma 4.1, Lemma 4.2 and condition (18), the posterior reference distribution is proper if and only if

$$\int_1^{\infty}O(1)O\left(s^{-\frac{3}{2}}\right)ds<\infty.$$

Thus,

$$\int_1^{\infty}O(1)O\left(s^{-\frac{3}{2}}\right)ds = \int_1^{\infty}s^{-\frac{3}{2}}ds = -2 < \infty. \tag{35}$$

Therefore, the reference prior leads to a proper posterior distribution.

Considering the Jeffreys prior given in Lemma 4.1, Lemma 4.2 and condition (18), the Jeffreys posterior distribution is proper if and only if

$$\int_1^{\infty} O(1)O\left(s^{-1}\right) ds < \infty.$$

Thus,

$$\int_1^{\infty} O(1)O\left(s^{-1}\right) ds = \int_1^{\infty} s^{-1} ds = \infty. \quad (36)$$

Therefore, the Jeffreys prior leads to an improper posterior distribution, completing the proof. □

ACKNOWLEDGEMENTS

The authors would like to thank the Editor and the two referees for careful reading and comments which greatly improved the paper.

Received 20 September 2018

REFERENCES

[1] Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions*. New York: Dover, p. 256. MR0208797
[2] Basak, P., Basak, I., and Balakrishnan, N. (2009). Estimation for the three-parameter lognormal distribution based on progressively censored data. *Computational Statistics & Data Analysis*, 53(10):3580–3592. MR2751948
[3] Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors. In *Bayesian Statistics 4 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.)*. Oxford: University Press, 35–60 (with discussion). MR1380269
[4] Bernardo, J. M. (1979). Reference posterior distributions for bayesian inference. *Journal of the Royal Statistical Society. Series B*, 41:113–147 (with discussion). Reprinted in Bayesian Inference (N. G. Polson and G. C. Tiao, eds.), Brookfield, VT: Edward Elgar (1995), 229–263. MR0547240
[5] Bernardo, J. M. (2005). *Reference Analysis*, volume 25. Elsevier. MR2490522
[6] Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde (2002). *Journal of the Royal Statistical Society. Series B - Statistical Methodology*, 64:616–618. MR1979380
[7] Carlin, B. P. and Louis, T. A. (2001). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall/CRC, Boca Raton, second edition. MR1427749
[8] Chen, G. (1995). Generalized log-normal distributions with reliability application. *Computational Statistics and Data Analysis*, 19:309–319. MR1323273
[9] Chen, K. and Papadopoulos, A. (1997). Shortest Bayes credibility intervals for the lognormal failure model. *Microelectron Reliability*, 37:1859–1863.
[10] Cohen, A. C. and Whitten, B. J. (1980). Estimation in the three-parameter lognormal distribution. *Journal of the American Statistical Association*, 75:399–404. MR0577372
[11] Cohen, A. C. and Whitten, B. J. (1988). *Parameter Estimation in Reliability and Life Span Models*. Marcel Dekker, New York. MR1022663
[12] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics 4 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.)*. Oxford: University Press, 625–631. MR1380276
[13] Harter, H. L. and Moore, A. H. (1966). Local-maximum-likelihood estimation of the parameters of the three-parameter lognormal populations from complete and censored samples. *Journal of the American Statistical Association*, 61:842–851. MR0203853
[14] Hill, B. M. (1963). The three-parameter lognormal distribution and bayesian analysis of a point-source epidemic. *Journal of the American Statistical Association*, 58:72–84. MR0144407
[15] Kalbfleisch, J. D. and Prentice, L. R. (2002). *The Statistical Analysis of Failure Time Data*. Wiley Series in Probability and Statistics. Wiley, 2 edition. MR1924807
[16] Kass, R. E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435):1343–1370.
[17] Komori, Y. and Hirose, H. (2004). Easy estimation by a new parametrization for the three-parameter lognormal distribution. *Journal of Statistical Computation and Simulation*, 74:63–74. MR2044181
[18] Martín, J. and Pérez, C. (2009). Bayesian analysis of a generalized lognormal distribution. *Computational Statistics & Data Analysis*, 53(4):1377–1387. MR2657098
[19] Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32(7):685–694. MR2119411
[20] Overduin, S. (2004). Use of the lognormal distribution for survival data: Inference and robustness. Master's thesis, Simon Fraser University.
[21] Ramos, P. L., Louzada, F., and Ramos, E. (2018). Posterior properties of the Nakagami-m distribution using noninformative priors and applications in reliability. *IEEE Transactions on Reliability*, 67(1):105–117.
[22] Salazar, E., Migon, H. S., and Ferreira, M. A. R. (2012). Bayesian analysis for exponential power regression model. *Sankhya - Series B*, 74:107–125. MR3021041
[23] Singh, V. P.; Cruise, J. F. and Ma, M. (1990). A comparative evaluation of the estimators of the three-parameter lognormal distribution by Monte Carlo simulation. *Computational Statistics and Data Analysis*, 10:71–85.
[24] Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statististical Society, Series B*, 64 (4):583–639. MR1979380
[25] Upadhyay, S. K. and Peshwani, M. (2001). Full posterior analysis of three parameter lognormal distribution using Gibbs sampler. *Journal of Statistical Computation and Simulation*, 71 (3):215–230. MR1886362
[26] Upadhyay, S. K., Vasishta, N., and Smith, A. F. M. (2001). Bayes inference in life testing and reliability via Markov chain Monte Carlo simulation. *The Indian Journal of Statistics. Series A*, 63 (1):15–40. MR1792622
[27] Vianelli, S. (1982). Sulle curve lognormali di ordine r quali famiglie di distribuzioni di errori di proporzione. *Statistica*, 42:155–176.
[28] Vianelli, S. (1983). The family of normal and lognormal distributions of order r. *Metron*, 41:3–10. MR0740130
[29] Zhu, D. and Zinde-Walsh, V. (2009). Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics*, 148:86–99. MR2494820

Vera L. D. Tomazella
Department of Statistics
Universidade Federal de São Carlos
São Carlos, São Paulo
Brazil
E-mail address: vera@ufscar.br

Sandra R. de Jesus
Multidisciplinary Health Institute
Federal University of Bahia
Vitória da Conquista, Bahia
Brazil
E-mail address: sandrarj@ufba.br

Francisco Louzada
Institute of Mathematical Science and Computing
University of São Paulo
São Carlos, São Paulo
Brazil
E-mail address: louzada@icmc.usp.br

Saralees Nadarajah
School of Mathematics
University of Manchester
Manchester
UK
E-mail address: Saralees.Nadarajah@manchester.ac.uk

Pedro L. Ramos
Institute of Mathematical Science and Computing
University of São Paulo
São Carlos, São Paulo
Brazil
E-mail address: pedrolramos@usp.br