2016

OBJECTIVE BAYESIAN ANALYSIS OF A GENERALIZED LOGNORMAL DISTRIBUTION

Shengnan Li
sli3@mtu.edu

OBJECTIVE BAYESIAN ANALYSIS OF A GENERALIZED LOGNORMAL DISTRIBUTION

By

Shengnan Li

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Mathematical Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY

2016

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Mathematical Sciences.

Department of Mathematical Sciences

Thesis Advisor:	*Dr. Min Wang*
Committee Member:	*Dr. Jianping Dong*
Committee Member:	*Dr. Qingli Dai*
Department Chair:	*Dr. Mark S. Gockenbach*

Dedication

To my parents, advisor and friends

who didn't hesitate to criticize my work at every stage - without which I would neither be who I am nor would this work be what it is today.

Contents

List of Figures

List of Tables

Abstract

The generalized lognormal distribution plays an important role in various aspects of life testing experiments. We examine Bayesian analysis of this distribution using objective priors (in the general sense of priors constructed using some formal rules) for the model parameters in this paper. Specifically, the derivation of explicit expressions for multiple types of the Jeffreys priors, the reference priors with different group ordering of the parameters, and the first-order matching priors. We investigate the important issue of proper posterior distributions. It is shown that only two of them lead to proper posterior distributions. Monte Carlo simulations are conducted to compare the performances of the Bayesian approaches under the various priors. Last, a real-world data case will be shown to illustrate the theoretical analysis.

Chapter 1

Introduction

Lognormal distribution is utilized in many different aspects of life sciences, including biology, ecology, and reliability/survival analysis as well as in economics, finance, and risk analysis. This is mainly because of its various attractive properties and its suitable fit for many experimental data, especially when the assumption of symmetry is not appropriate. For example, the lengths of incubation periods (time from exposure to the point at which first symptoms appear) of infectious diseases usually fit closely to the lognormal distribution. We here refer the interested readers to [1] for details on this topic.

If a random variable Y has the lognormal distribution, the random variable $X = \exp(Y)$ is normally distributed. Recently, [18] studied a generalized form of the lognormal distribution. The authors have generalized the two-parameter log-normal distribution to the three-parameter generalized lognormal (for short, logGN) distribution with the additional parameter, which provides a more suitable transformation for analyzing asymmetric data sets. Thus, the logGN distribution can adequately model the data whereby the lognormal distribution may not be absolutely suitable. As an illustration, [18] showed its superior performance on the analysis of life cycle data belonging to the field of engineering. Consequently, besides the classical lognormal distribution, the logGN distribution can be viewed as another one of the important skewed distribution for analyzing the data from different fields, atmospheric sciences, environmental sciences, microbiology, reliability/survival analysis; see, for example, [7], [24].

We say that the random variable $X = \log(Y)$ is generalized normal (GN) distribution if a random variable Y follows the logGN distribution. The probability density function (pdf) of the logGN distribution with parameters μ, σ, and s is

$$f(y \mid \mu, \sigma, s) = \frac{s}{2y\sigma\Gamma(1/s)} \exp\left(-\left|\frac{\log y - \mu}{\sigma}\right|^s\right), \tag{1.1}$$

where $y > 0, -\infty < \mu < \infty, \sigma > 0$, and $s \geq 1$. When $s = 2$, the logGN distribution reduces to the lognormal one, showing its more flexibility to experimental data than the lognormal one. This distribution also include the logLaplace distribution as a particular case by taking $s = 1$.

Figure 1.1 shows the visual representation of logGN's density and hazard rate function with different choices of the parameter s with $\mu = 1, \sigma = 0.5$. It can be seen from the two figures that the logGN distribution has very flexible shapes of density and hazard rate functions based on different combinations of the unknown parameters. In addition, we observe that as y tends to infinity, the density of the logGN distribution approaches 0, indicating that the logGN distribution is suitable for modeling the data under the situation in which the large values of y are not of interest.

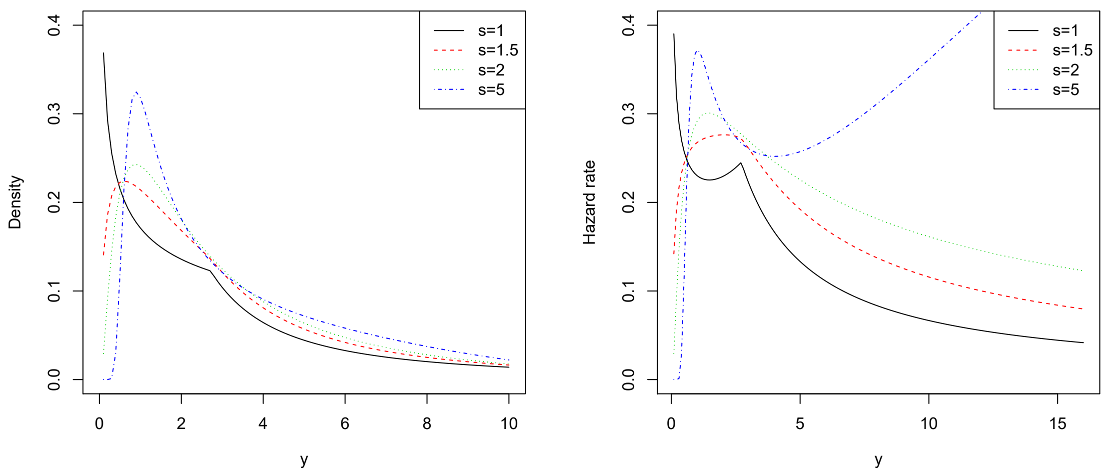

Figure 1.1: logGN's Probability density function and hazard rate plots for $\mu = 1$, and $\sigma = 0.5$ and different values of $\{s = 1, 1.5, 2, 5\}$.

We shall thus be interested in estimating the three unknown parameters of the logGN distribution from both the frequentist and Bayesian frameworks. Here, we contemplate objective Bayesian analysis of the logGN distribution using objective priors for the unknown parameters, which are constructed using some formal rules. To the best of our knowledge, there are just few Bayesian steps for analysing the logGN distribution in the literature. [18] derived the independence Jeffreys prior by treating the three parameters independently and provided a simple approximated form to this prior, whereas they did not consider other commonly used objective priors based on other formal rules. Additionally, they did not investigate the important issue of whether the considered priors result in proper posterior distribution. Later on, we will show that the prior used by [18] results in an improper posterior distribution. Thus, special attention should be paid when we use improper priors for the unknown parameters. In this paper, we have different types of the Jeffreys priors, the reference priors with different group ordering of the parameters, and the first-order matching priors and study their posterior proprieties under these improper priors. This study is quite important from both theoretical and practical viewpoints, because the results not only prevent researchers from making invalid statistical inference from improper posterior distributions, but also provide a guideline to perform Bayesian analysis for the logGN distribution using objective priors of the unknown parameters.

The rest of the paper has the following sections. In Section 2, we consider various objective priors of the model parameters constructed using some formal rules and provide a general form of the various priors under consideration. In Section 3, we investigate the issue of whether these improper priors result in proper posterior distributions. In Section 4, we develop an efficient Gibbs sampler algorithm for posterior computation. In Section 5, computational simulations are conducted to compare the various priors and the maximum-likelihood estimation (MLE). A real data application is presented in Section 6. Concluding remarks and future work are illustrated in the fianl section. We will also provide more detailed proofs in the Appendix.

Chapter 2

Objective Bayesian analysis

Bayesian analysis begins with the prior specifications for the unknown parameters. In this section, we derive the three types of the Jeffreys priors (Section 2.0.1), the two types of reference priors for all possible model parameters (Section 2.0.2), and the general form of the first-order probability matching priors (Section 2.0.3).

2.0.1 Inside of the Jeffreys-type priors

When prior knowledge is missing, the noninformative priors of the unknown parameters is often preferred and are usually obtained from the expected Fisher information of the model. It can be seen from [18] that the Fisher information matrix of the logGN distribution is as following:

$$H(\varphi) = \begin{pmatrix} \frac{(s-1)s\Gamma(1-s^{-1})}{\sigma^2\Gamma(s^{-1})} & 0 & 0 \\ \\ 0 & \frac{s}{\sigma^2} & -\frac{A}{\sigma s} \\ \\ 0 & -\frac{A}{\sigma s} & \frac{A^2+B}{s^3} \end{pmatrix}, \tag{2.1}$$

where $\varphi = (\mu, \sigma, s)$, $\psi(\cdot)$ is the diagamma function, $A = 1 + \psi(1 + s^{-1})$ and $B = (1 + s^{-1})\psi'(1 + s^{-1}) - 1$.

Within the Bayesian framework, one of the commonly used noninformative priors is the Jeffreys (Jeffreys, 1998), which contains Jeffreys-rule prior and independence Jeffreys priors. For the logGN distribution, we consider the following two groups of the parameters: $\{(\mu), (\sigma, s)\}$ and $\{(\mu), (\sigma), (s))\}$. It will be shown that these different types of the Jeffreys priors can be unified as

$$\pi(\varphi) \propto \frac{\pi(s)}{\sigma^a}, \tag{2.2}$$

where $a \in \mathbb{R}$ is a hyper-parameter and $\pi(s)$ can be defined as the 'marginal' prior of the parameter s. We summarize these priors in the following theorem with proofs provided in the Appendix.

Theorem 1 *Consider the logGN distribution with the pdf in (1.1). The independence Jeffreys priors with the groupings* $\{(\mu),(\sigma),(s))\}$ *and* $\{(\mu),(\sigma,s)\}$*, and the Jeffreys-rule priors for* $\{\mu,\sigma,s\}$ *are marked as* $\pi^{\mathrm{J1}}(\varphi),\pi^{\mathrm{J2}}(\varphi)$*, and* $\pi^{J}(\varphi)$*, respectively. They are of the form (2.2) with*

$$a=1,\quad \pi^{\mathrm{J1}}(s)\propto s^{-1}\ \big[B\big]^{1/2}, \tag{2.3}$$

$$a=1,\quad \pi^{\mathrm{J2}}(s)\propto s^{-3/2}\ \big[A^2+B\big]^{1/2}, \tag{2.4}$$

$$a=2,\quad \pi^{\mathrm{J}}(s)\propto \left[\frac{s(s-1)\Gamma(1-s^{-1})}{\Gamma(s^{-1})}\right]^{1/2}\ \pi^{\mathrm{J1}}(s). \tag{2.5}$$

where $B=(1+s^{-1})\psi'(1+s^{-1})-1.$

As commented by [3], the Jeffreys-type priors may be unsatisfactory for multiparameter problems if we are only interested in a subset of the parameters with the rest treated as nuisance parameters, because it may result in some unsatisfied results. For instance, the frequentist coverage of Bayesian credible interval from the Jeffreys-rule priors may not reach the desired theoretical level. This motivates the study of alternative objective priors constructed based on some formal rules.

2.0.2 Inside of the reference priors

According to the influential paper of [5], the reference priors have been indicated as alternative tools for developing noninformative priors of the parameters. Note that this prior in problems involving multiple parameters depends on the different orderings of the unknown parameters. Since the Fisher information (2.1) does not depend on the location parameter μ, we can put it in anywhere. Thus, different orderings lead to the two types of the reference priors summarized in following theorem with proofs given in the Appendix.

Theorem 2 *Consider the logGN distribution with the pdf in (1.1). when $a = 1$, for the group orderings $(\mu, \sigma, s), (\sigma, s, \mu)$, and (σ, μ, s), the 'marginal' prior of the parameter s is*

$$\pi^{\mathrm{R1}}(s) \;\propto\; s^{-3/2}\;[A^2 + B]^{1/2}, \tag{2.6}$$

whereas for the group orderings $(\mu, s, \sigma), (s, \sigma, \mu)$, and (s, μ, σ) the 'marginal' prior of the parameter s is

$$\pi^{\mathrm{R2}}(s) \;\propto\; s^{-3/2}\;[B]^{1/2}. \tag{2.7}$$

It should be noted that π^{R1} is exactly the same as π^{J2} and that the expressions of other priors in Theorems 1 and 2 are quite similar. However, from the follwing results, we can get different answers, especially when we have smaller sample size. In particular, we will show that they behave differently in terms of the probability matching criteria defined by [11] from a theoretical point of view.

2.0.3 Inside of the First-order matching priors

A prior distribution under which the posterior probabilities of specific districts exactly or approximately coincide with their coverage probabilities is called a probability matching prior. More examples can be seen in the paper of [10], [11]. Since the shape parameter s is very important, we could develop first-order matching prior's general form, when s is the parameter of interest. The result helps us to choose a better prior of the Bayesian estimation.

Since we are interested in estimating the parameter s, we arrange the Fisher information $H(\varphi)$ in terms of the group ordering (s, σ, μ). The Fisher information matrix can then be rewritten as

$$H(\varphi) = \begin{pmatrix} \frac{A^2+B}{s^3} & -\frac{A}{\sigma s} & 0 \\ -\frac{A}{\sigma s} & \frac{s}{\sigma^2} & 0 \\ 0 & 0 & \frac{(s-1)s\Gamma(1-s^{-1})}{\sigma^2\Gamma(s^{-1})} \end{pmatrix}.$$

An orthogonal reparameterization of (s, σ, μ). Let $\theta_1 = s$, $\theta_2 = g(s, \sigma)$, and $\theta = (\theta_1, \theta_2, \theta_3)$ can simplify the procedure. We need to obtain a solution for

$$H(s, \sigma, \mu) = \begin{pmatrix} 1 & \frac{\partial g}{\partial s} & 0 \\ 0 & \frac{\partial g}{\partial \sigma} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_{\theta_1,\theta_1} & 0 & 0 \\ 0 & I_{\theta_2,\theta_2} & 0 \\ 0 & 0 & I_{\theta_3,\theta_3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{\partial g}{\partial s} & \frac{\partial g}{\partial \sigma} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then several differential equations can be derived as followings:

$$I_{\theta_1,\theta_1} + \left(\frac{\partial g}{\partial s}\right)^2 I_{\theta_2,\theta_2} = \frac{B + A^2}{s^3},$$

$$\frac{\partial g}{\partial s}\frac{\partial g}{\partial \sigma} I_{\theta_2,\theta_2} = -\frac{A}{\sigma s},$$

$$\left(\frac{\partial g}{\partial \sigma}\right)^2 I_{\theta_2,\theta_2} = \frac{s}{\sigma^2},$$

which gives

$$\frac{\partial g}{\partial s} + \frac{A\sigma}{s^2}\frac{\partial g}{\partial \sigma} = 0.$$

It can be verified that $g(s, \sigma) = \sigma^{-1} \exp(-s^{-1})[\Gamma(1 + s^{-1})]^{-1}$ is a solution, leading to

$$\theta_1 = s, \quad \theta_2 = \sigma^{-1} \exp\left(-s^{-1}\right) \left[\Gamma(1 + s^{-1})\right]^{-1}, \quad \theta_3 = \mu.$$

The likelihood function of θ is defined as

$$L(\theta) = \frac{\theta_2^n}{2^n} \exp\left\{ \frac{n}{\theta_1} - [\theta_2 \exp\left(\frac{1}{\theta_1}\right) \Gamma\left(1+\frac{1}{\theta_1}\right)]^{\theta_1} \sum_{i=1}^{n} \left| \frac{y_i - \theta_3}{\sigma} \right|^{\theta_1} \right\}.$$

The corresponding Fisher information under the above orthogonal reparameterization is given by

$$R(\theta) = \mathrm{Diag}\left\{ b_1(\theta),\ b_2(\theta),\ b_3(\theta) \right\}, \tag{2.8}$$

where Diag$\{\cdot\}$ is a diagonal matrix with the diagonal elements

$$b_1(\theta) = [(1+\theta_1^{-1})\psi'(1+\theta_1^{-1}) + A^2 - 1]\ \theta_1^{-3},$$

$$b_2(\theta) = \theta_1 \theta_2^2 \exp(2\theta_1^{-1})\ [\Gamma(1+\theta_1^{-1})]^2,$$

$$b_3(\theta) = \frac{(\theta_1 - 1)\theta_1\ \Gamma(1-\theta_1^{-1})}{\Gamma(\theta_1^{-1})}\ \theta_2^2 \exp(2\theta_1^{-1})\ [\Gamma(1+\theta_1^{-1})]^2.$$

It follows that the form of first-order matching prior is

$$\begin{aligned} \pi(\theta) &\propto a_1(\theta)^{1/2}\ g(\theta_2, \theta_3) \\ &\propto \theta_1^{-3/2}\ [(1+\theta_1^{-1})\ \psi'(1+\theta_1^{-1}) + A^2 - 1]^{1/2}\ k(\theta_2, \theta_3), \end{aligned} \tag{2.9}$$

where $k(\cdot)$ is an arbitrary positive and differentiable function of θ_2 and θ_3. Moreover, by letting $k(\theta_2, \theta_3) = \sigma^{-1}$, a first-order matching prior reduces to

$$\pi(\sigma, \mu, s) \propto \sigma^{-1}\ s^{-3/2}\ [A^2 + B]^{1/2},$$

which shows that π^{R1} in (2.6) is the same as the first-order probability matching prior. Other objective priors (2.3), (2.5), and (2.7) given in Theorems 1 and 2 are not. This shows that even the priors looks similar, the behavior of these priors is quite different from a theoretical viewpoint.

Since we can choose different function for $k(\cdot)$, there are uncountable first-order matching priors for the model parameters. Thus, it is of interest to narrow down the subclass of the priors according to the second-order matching criterion (Mukerjee and Dey, 1993). After tedious algebraic implications of derivatives and expectations from the Fisher information in (2.8), it can be proved that on second-order matching priors in our case (2.9) when s is the parameter of interest.

It should be noted that the priors in Theorems 1 and 2 are all improper. We are thus interested in investigating the important issue of whether these improper priors result in proper posterior distributions, since statistical inference based on an improper posterior distribution is invalid (Mukerjee and Ghosh, 1997.

Chapter 3

Propriety of the posterior distributions

Suppose that $y_1, \cdots, y_n$ are independent and identically distribution (iid) random variables generated from the logGN distribution with the pdf in (1.1). We have the likelihood function as

$$L(\mu, s, \sigma) \propto \sigma^{-n} s^n \big[\Gamma(1/s)\big]^{-n} \exp\bigg\{-\sum_{i=1}^{n}\Big|\frac{y_i-\mu}{\sigma}\Big|^s\bigg\}. \tag{3.1}$$

The joint posterior distribution of the unknown parameters under the prior of the form (2.2) is given by

$$\pi^{\tau}(\mu, s, \sigma \mid \mathbf{y}) \propto L(\mu, s, \sigma)\sigma^{-a}\pi^{\tau}(s), \tag{3.2}$$

where $\mathbf{y} = (y_1, \cdots, y_n)'$ and $\tau = \{\mathrm{J}, \mathrm{J1}, \mathrm{J2}, \mathrm{R1}, \mathrm{R2}\}$ represents the objective prior under consideration. The joint posterior distribution of (μ, s, σ) is proper if and only if

$$\int_1^{\infty} \pi^{\tau}(s \mid \mathbf{y})\, ds < \infty, \tag{3.3}$$

where $\pi^{\tau}(s \mid \mathbf{y})$ is the marginal posterior distribution of the parameter s obtained by integrating with μ and σ

$$\pi^{\tau}(s \mid \mathbf{y}) \propto \pi^{\tau}(s) \int_{\mathbb{R}} \int_0^{\infty} L(\mu, s, \sigma)\sigma^{-a}\ d\sigma\ d\mu = \pi^{\tau}(s) L^{I}(s; \mathbf{y}).$$

The finite nature of the integral in (3.3) is related to the tail behavior of $\pi^{\tau}(s \mid \mathbf{y})$, which is determined by both the integrated likelihood $L^{I}(s; \mathbf{y})$ and the 'marginal' priors of the parameter s in Theorems 1 and 2. The following two lemmas play a key role in determining whether the posterior distributions of the unknown parameters are proper under these priors.

Lemma 1 *The marginal priors of the parameter s in Theorem 1 are continuous functions in $[1, \infty)$ and are such that $\pi^{\mathrm{J1}}(s) = O(s^{-1})$, $\pi^{\mathrm{J2}}(s) = O(s^{-3/2})$, and $\pi^{\mathrm{J}}(s) =$*

$O(s^{-1/2})$ as $s \to \infty$. Similarly $\pi^{\mathrm{R1}}(s) = \pi^{\mathrm{R2}}(s) = O(s^{-3/2})$ as $s \to \infty$.

Lemma 2 *When $n > 3-a$, the integrated likelihood of the parameter s is a continuous function in $[1, \infty)$ and is such that $L^I(s; \mathbf{y}) = O(1)$ as $s \to \infty$.*

The propriety of the posterior distribution of various priors in Theorems 1 and 2 are given by as followings.

Proposition 1 *Consider the logGN distribution with the pdf in (1.1). The Jeffreys-type priors π^J and π^{J1} in Theorem 1 lead to improper posterior distributions. Provided that $n > 2$, the prior π^{J2} and the two reference priors in Theorem 2 result in proper posterior distributions.*

It deserves mentioning that [18] derived the independence Jeffreys prior π^{J2} of the unknown model parameters. Because the expression of π^{J2} is quite complex, they provided a simple approximated form given by $\pi^{\mathrm{MP}}(s) \propto s^{-1/2}$. We observe from Lemmas 1 and 2 that this approximation may lead to an improper posterior distribution. Instead, we propose a valid appropriated form for the 'marginal' prior of s according to the tail behavior of the priors with respect to s. Specifically, we consider an approximated form given by

$$a = 1, \quad \pi^{\mathrm{RA}}(s) \propto s^{-3/2}.$$

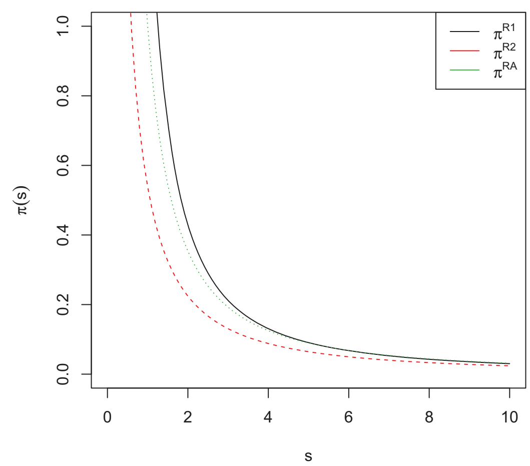

Figure 3.1: Comparsion between priors π^{R1}, π^{R2} and π^{RA}.

This approximation will not only lead to a proper posterior distribution, but also simplifies the posterior computation under other objectives priors in Theorems 1 and 2. It can be seen from Figure 3.1 that this approximated form falls between the two reference priors.

We also observe from expressions (3.1) and (3.2) that the joint posterior distribution of the unknown parameters (μ, s, σ) is not recognizable, so an efficient Gibbs sampler algorithm needs to be developed for generating posterior samples to make statistical inference.

Chapter 4

Posterior computation

The full conditional posterior distributions of the unknown model parameters are given by

$$
\begin{aligned}
&\mu \mid \sigma, s, \mathbf{u}, \mathbf{y} \propto 1, \;\; \max_i \{\log(y_i) - \sigma u_i^{1/s}\} < \mu < \min_i \{\log(y_i) + \sigma u_i^{1/s}\}, \\
&\sigma \mid \mu, s, \mathbf{u}, \mathbf{y} \propto \frac{1}{\sigma^{n+1}}, \;\; \sigma > \max_i \left\{ \frac{|\mu - \log(y_i)|}{u_i^{1/s}} \right\}, \\
&u_i \mid \mu, s, \sigma, \mathbf{y} \propto \exp(u_i), \;\; u_i > \left\{ \frac{|\log(y_i) - \mu|}{\sigma} \right\}^s, \quad i = 1, \cdots, n, \\
&s \mid \mu, s, \sigma, \mathbf{u}, \mathbf{y} \propto \frac{s^{n-1}}{\Gamma^n(1/s)}, \max_{i \in S^-} \{1, a_i\} < s < \min_{i \in S^+} a_i,
\end{aligned} \tag{4.1}
$$

where $S^- = \{i : \log(|\mu - \log(y_i)|/\sigma < 0\}$, $S^+ = \{i : \log(|\mu - \log(y_i)|/\sigma > 0\}$, and

$$a_i = \frac{\log(u_i)}{\log(|\mu - \log(x_i)|/\sigma}, \quad i = 1, \cdots, n.$$

Therefore, an efficient Gibbs sampler algorithm for generating posterior samples can be developed as follows.

i) Simulate μ from the uniform distribution in the interval $[\max_i\{\log(y_i) - \sigma u_i^{1/s}\},\ \min_i\{\log(y_i) + \sigma u_i^{1/s}\}]$.

ii) Simulate σ from the Pareto distribution with the scale parameter $\max_i\big\{|\mu - \log(y_i)|/u_i^{1/s}\big\}$ and the shape parameter n.

iii) Simulate u_i from the truncated exponential distribution with the rate parameter 1 by truncating on right side of $\big\{|\log(y_i) - \mu|/\sigma\big\}^s$ for $i = 1, \cdots, n$.

iv) Simulate s from the conditional posterior distribution in (4.1) using the acceptance-rejection method developed by [12].

It should noted that the full conditional posterior distribution of the parameter s is not of standard form; we here employ the acceptance-rejection method (Devroye, 1986) for posterior simulation. Our simulation studies in the next section indicate that the proposed sampling algorithm is quite efficient due to mixing and convergence under different simulation scenarios.

Chapter 5

Simulation studies

In this section, we use computational simulations to assess the performance of the Bayesian procedures under the various priors as well as the MLEs for the parameters of the logGN distribution. In these examples, we take the shape parameter $s = 1, 2, 3, 5$ and the sample size $n = 25, 50, 100$. Without loss of generality, the location parameter μ and the scale parameter σ are kept fixed at 0 and 1, respectively. All the results were based on $1,000$ repetitions. We report the average values of the posterior mean, the posterior median, and the squared root of mean squared error (SRMSE) of each estimator. We also report Bayesian frequentist coverage and the average length of Bayesian credible intervals for each parameter under the various priors studied in this paper.

Table 5.1
The averaged estimates of each parameter and the SRMSE(in the parenthesis) based on $1,000$ repetition for $n = 25$.

s	Parameter	π^{R1} Mean	π^{R1} Median	π^{R2} Mean	π^{R2} Median	π^{RA} Mean	π^{RA} Median	MLE
1	μ	−0.0064	−0.0071	0.0069	0.0066	−0.0034	−0.0029	−0.0099
		(0.2410)	(0.2393)	(0.2409)	(0.2383)	(0.2329)	(0.2311)	(0.2473)
	σ	1.4518	1.3972	1.5044	1.4492	1.4609	1.4056	1.0335
		(0.5612)	(0.5177)	(0.6085)	(0.5630)	(0.5697)	(0.5252)	(0.6165)
	s	1.5414	1.3430	1.6222	1.3993	1.5478	1.3507	1.3051
		(0.7738)	(0.4796)	(0.8913)	(0.5604)	(0.7743)	(0.4809)	(1.0264)
2	μ	0.0020	0.0018	0.0061	0.0063	−0.0029	−0.0030	−0.0042
		(0.1414)	(0.1422)	(0.1519)	(0.1528)	(0.1472)	(0.1484)	(0.1536)
	σ	0.9192	0.9056	0.9391	0.9286	0.9350	0.9235	0.9215
		(0.1989)	(0.2147)	(0.1978)	(0.2123)	(0.1915)	(0.2064)	(0.2929)
	s	2.2672	1.913	2.4907	1.9700	2.4048	1.9108	2.4849
		(1.0458)	(0.6621)	(1.3115)	(0.7691)	(1.2348)	(0.7401)	(1.5962)
3	μ	0.0054	0.0053	0.0064	0.0065	0.0012	0.0011	−0.0053
		(0.1225)	(0.1229)	(0.1240)	(0.1243)	(0.1175)	(0.1180)	(0.1266)
	σ	0.8562	0.8545	0.8915	0.8924	0.8830	0.8835	0.8967
		(0.2134)	(0.2232)	(0.1941)	(0.2020)	(0.1921)	(0.2002)	(0.2353)
	s	2.9233	2.2341	3.2717	2.4922	3.1573	2.3985	3.3182
		(1.3691)	(1.1596)	(1.6157)	(1.1451)	(1.5243)	(1.1500)	(1.8445)
5	μ	0.0038	0.0038	0.0064	0.0061	0.0016	0.0017	−0.0012
		(0.1043)	(0.1038)	(0.1109)	(0.1106)	(0.1076)	(0.1071)	(0.1115)
	σ	0.8776	0.8847	0.8422	0.8475	0.8615	0.8677	0.8573
		(0.1805)	(0.1799)	(0.2072)	(0.2089)	(0.1908)	(0.1914)	(0.2272)
	s	4.3633	3.1820	3.7222	2.7244	3.995	2.9404	4.1470
		(2.1025)	(2.2296)	(2.2070)	(2.5528)	(2.0862)	(2.3788)	(2.4149)

We ran the Gibbs sampler algorithm in Section 4 to obtain $50,000$ observations of Markov chains, where the first 5,000 samples are discarded as burn-in periods with a thinning of 10 draws. According to our examination, there is no proof of absence of merging in view of the run length control analytic([21]) and the union symptomatic test measurement (at an importance level of 5%) by [15].

Table 5.2

The averaged estimates of each parameter and the SRMSE(in the parenthesis) based on 1,000 repetition for $n = 50$.

s	Parameter	π^{R1}		π^{R2}		π^{RA}		
		Mean	Median	Mean	Median	Mean	Median	MLE
1	μ	−0.0018	−0.0018	0.0035	0.0036	−0.0017	−0.0017	0.0096
		(0.1520)	(0.1513)	(0.1631)	(0.1624)	(0.1564)	(0.1555)	(0.1680)
	σ	1.3095	1.2744	1.3269	1.2911	1.3385	1.3031	1.0253
		(0.3914)	(0.3636)	(0.3988)	(0.3693)	(0.4098)	(0.3807)	(0.4479)
	s	1.2796	1.2137	1.3050	1.2349	1.3097	1.2397	1.1173
		(0.3680)	(0.2877)	(0.3943)	(0.3098)	(0.3869)	(0.3068)	(0.5567)
2	μ	0.0036	0.0035	−0.0001	0.0000	0.0026	0.0024	−0.0047
		(0.1004)	(0.1007)	(0.1068)	(0.1071)	(0.1040)	(0.1045)	(0.1067)
	σ	0.9581	0.9560	0.9350	0.9312	0.9457	0.9425	0.9965
		(0.1651)	(0.1744)	(0.1759)	(0.1863)	(0.1683)	(0.1782)	(0.2112)
	s	2.2954	2.0434	2.1671	1.9330	2.1994	1.9659	2.5471
		(1.0894)	(0.7757)	(1.1216)	(0.8121)	(1.1192)	(0.7910)	(1.5899)
3	μ	−0.0014	−0.0012	0.0010	0.0009	−0.0018	−0.0017	0.0018
		(0.0885)	(0.0887)	(0.0874)	(0.0876)	(0.0850)	(0.0852)	(0.0844)
	σ	0.9186	0.9239	0.9314	0.9372	0.9315	0.9378	0.9851
		(0.1607)	(0.1632)	(0.1469)	(0.1486)	(0.1482)	(0.1498)	(0.1421)
	s	3.1762	2.7103	3.3509	2.8533	3.3440	2.8344	3.9661
		(1.7131)	(1.2122)	(1.7138)	(1.1906)	(1.7997)	(1.2118)	(2.5498)
5	μ	0.0029	0.0028	0.0016	0.0017	−0.0007	−0.0005	0.0018
		(0.0672)	(0.0670)	(0.0683)	(0.0681)	(0.0705)	(0.0703)	(0.0696)
	σ	0.9375	0.9455	0.9384	0.9459	0.9231	0.9306	0.9707
		(0.1186)	(0.1153)	(0.1167)	(0.1134)	(0.1339)	(0.1311)	(0.1014)
	s	5.5300	4.3712	5.5007	4.3688	5.2316	4.1774	6.0692
		(2.8941)	(2.0151)	(2.9836)	(2.0658)	(2.8792)	(2.1094)	(3.2712)

The first study is devoted to the comparison between the MLEs and the Bayesian estimations based on the two reference priors in Theorem 2 and the approximated one $\pi^{\mathrm{RA}}(s) \propto s^{-3/2}$. Tables 5.1, 5.2, and 5.3 provide the Baysian estimates and the MLEs with their squared root mean squared error for all three parameters μ, σ, and s of the three different priors $\pi^{\mathrm{R1}}, \pi^{\mathrm{R2}}$, and π^{RA} when $n = 25, 50$, and 100, respectively. From these tables, some features can be drawn as follows:

Table 5.3

The averaged estimates of each parameter and the SRMSE(in the parenthesis) based on 1,000 repetition for $n = 100$.

s	Parameter	π^{R1} Mean	π^{R1} Median	π^{R2} Mean	π^{R2} Median	π^{RA} Mean	π^{RA} Median	MLE
1	μ	−0.0041	−0.0038	0.0069	0.0066	−0.0034	−0.0029	−0.0008
		(0.1028)	(0.1026)	(0.2409)	(0.2383)	(0.2329)	(0.2311)	(0.1063)
	σ	1.2170	1.1945	1.5044	1.4492	1.4609	1.4056	1.0085
		(0.2690)	(0.2503)	(0.6085)	(0.5630)	(0.5697)	(0.5252)	(0.2972)
	s	1.1788	1.1459	1.6222	1.3993	1.5478	1.3507	1.0406
		(0.2091)	(0.1778)	(0.8913)	(0.5604)	(0.7743)	(0.4809)	(0.2372)
2	μ	0.0020	0.0018	0.0061	0.0063	−0.0029	−0.0030	0.0006
		(0.1414)	(0.1422)	(0.1519)	(0.1528)	(0.1472)	(0.1484)	(0.0731)
	σ	0.9192	0.9056	0.9391	0.9286	0.9350	0.9235	1.0055
		(0.1989)	(0.2147)	(0.1978)	(0.2123)	(0.1915)	(0.2064)	(0.1372)
	s	2.2672	1.8113	2.4907	1.9700	2.4048	1.9108	2.2298
		(1.0458)	(0.6621)	(1.3115)	(0.7691)	(1.2348)	(0.7401)	(0.7019)
3	μ	0.0054	0.0053	0.0064	0.0065	0.0012	0.0011	−0.0007
		(0.1225)	(0.1229)	(0.1240)	(0.1243)	(0.1175)	(0.1180)	(0.0599)
	σ	0.8562	0.8545	0.8915	0.8924	0.8830	0.8835	0.9983
		(0.2134)	(0.2232)	(0.1941)	(0.2020)	(0.1921)	(0.2002)	(0.0995)
	s	2.9233	2.2341	3.2717	2.4922	3.1573	2.3985	3.5329
		(1.3691)	(1.1596)	(1.6157)	(1.1451)	(1.5243)	(1.1500)	(1.5313)
5	μ	0.0038	0.0038	0.0064	0.0061	0.0016	0.0017	−0.0011
		(0.1043)	(0.1038)	(0.1109)	(0.1106)	(0.1076)	(0.1071)	(0.0477)
	σ	0.8776	0.8847	0.8422	0.8475	0.8615	0.8677	0.9937
		(0.1805)	(0.1799)	(0.2072)	(0.2089)	(0.1908)	(0.1914)	(0.0659)
	s	4.3633	3.1820	3.7222	2.7244	3.995	2.9404	6.0746
		(2.1025)	(2.2296)	(2.2070)	(2.5528)	(2.0862)	(2.3788)	(2.7785)

(1) Intuitively, when we increase the sample size, all the considered estimates become closer the true parameter value and their SRMSE decreases indicating the expanded in the exactness of the evaluating process.

(2) Compared to MLEs, the Bayesian estimates under the three considered priors seem to be more steady even when the sample size becomes large. However, all

the considered estimators perform well for estimating the parameter s when s is small; and they are all consistently less accurate for the large values of s.

(3) The posterior medians outperform the posterior means in most cases in terms of the estimation precision and the SRMEs. In addition, the Bayesian estimates under the approximated 'marginal' prior of s perform very well under different simulation scenarios.

Table 5.4
Comparison among Bayesian approaches under the different objective priors due to the frequentist coverage and the average length (in the parenthesis) of 95% credible interval of each parameter when $n = 25$.

s	π^{R1}			π^{R2}			π^{RA}		
	μ	σ	s	μ	σ	s	μ	σ	s
1	0.966	0.938	1.000	0.965	0.910	1.000	0.967	0.926	1.000
	(0.9851)	(1.4584)	(1.4731)	(1.0066)	(1.5287)	(1.8679)	(0.9957)	(1.4916)	(1.6165)
2	0.951	0.934	0.943	0.952	0.944	0.965	0.930	0.948	0.961
	(0.5667)	(0.8536)	(4.0365)	(0.5709)	(0.8735)	(4.5414)	(0.5687)	(0.8628)	(4.3309)
3	0.942	0.895	0.878	0.937	0.921	0.921	0.938	0.909	0.919
	(0.5010)	(0.7472)	(5.8994)	(0.4935)	(0.7435)	(6.9714)	(0.4912)	(0.7381)	(6.7876)
5	0.949	0.906	0.827	0.955	0.909	0.874	0.966	0.912	0.857
	(0.4497)	(0.6568)	(9.3916)	(0.4437)	(0.6469)	(10.3896)	(0.4436)	(0.6493)	(10.118)

The second study is devoted to the comparison of the Baysian frequentist coverage probabilities for the three parameters under the three different priors studied in this paper. Under the same simulation scenarios mentioned above, we examine the coverage probabilities of Bayesian 95% credible intervals and the average length for the three parameters. The results of the simulation study have been summarized in Tables 5.4, 5.5, 5.6. Even we have small sample size ($n = 25$), we can still observe from these tables that the frequentist coverage probabilities are very close to 0.95 for the three parameters; and thus we may conclude that the Baysian procedures under the considered priors enjoy good frequentist properties. As expected, the Bayesian procedure under the prior π^{R1} performs the ones under the other two priors. It clearly guarantees the validity of the result that π^{R1} is a first-order matching prior when we interest in s.

Table 5.5

Comparison among Bayesian approaches under the different objective priors due to the frequentist coverage and the average length (in the parenthesis) of 95% credible interval of each parameter when $n = 50$.

s	π^{R1}			π^{R2}			π^{RA}		
	μ	σ	s	μ	σ	s	μ	σ	s
1	0.973	0.926	1.000	0.963	0.914	0.999	0.969	0.917	0.999
	(0.6611)	(1.0554)	(0.8761)	(0.6621)	(1.0749)	(0.8977)	(0.6674)	(1.0796)	(0.9208)
2	0.941	0.929	0.934	0.938	0.925	0.931	0.933	0.920	0.919
	(0.3946)	(0.6929)	(2.989)	(0.3984)	(0.6998)	(3.0346)	(0.3978)	(0.6955)	(2.9948)
3	0.948	0.911	0.871	0.952	0.936	0.921	0.932	0.916	0.889
	(0.3362)	(0.5704)	(5.2888)	(0.3362)	(0.5647)	(5.9576)	(0.3353)	(0.5668)	(5.6251)
5	0.947	0.931	0.862	0.963	0.960	0.902	0.954	0.946	0.895
	(0.2861)	(0.4444)	(10.5004)	(0.2838)	(0.4292)	(11.6785)	(0.2828)	(0.4321)	(11.692)

Table 5.6

Comparison among Bayesian approaches under the different objective priors due to the frequentist coverage and the average length (in the parenthesis) of 95% credible interval of each parameter when $n = 100$.

s	π^{R1}			π^{R2}			π^{RA}		
	μ	σ	s	μ	σ	s	μ	σ	s
1	0.957	0.902	0.994	0.960	0.897	0.989	0.944	0.895	0.995
	(0.4456)	(0.7527)	(0.5236)	(0.4489)	(0.7545)	(0.5374)	(0.4481)	(0.7529)	(0.5297)
2	0.953	0.928	0.931	0.949	0.902	0.898	0.921	0.926	0.923
	(0.2799)	(0.5347)	(1.7827)	(0.2772)	(0.5317)	(1.8305)	(0.2790)	(0.5329)	(2.0278)
3	0.950	0.940	0.909	0.961	0.937	0.911	0.951	0.949	0.918
	(0.2318)	(0.4023)	(3.5488)	(0.2313)	(0.3982)	(3.6467)	(0.2310)	(0.3967)	(3.7328)
5	0.966	0.943	0.911	0.950	0.947	0.903	0.952	0.950	0.900
	(0.1878)	(0.2786)	(8.3203)	(0.1865)	(0.2747)	(8.473)	(0.1878)	(0.2794)	(8.508)

Chapter 6

Real data application

The aim of this section is to illustrate the application of the Bayesian procedures under the various priors by using the data originally studied by [23]. The data have been previously used in the literature and can be found in Table 2 of [18]. This data provide the life time of 59 test conductor of 400-micrometer length. The 59 specimens were all tested under the same temperature and current density until they all raced to fail at a specific high temperature and current density. [18] have shown that the logGN distribution provides a better data fitting than the one based on the lognormal distribution.

For each choice of the considered priors, we generate 22,000 posterior samples using the proposed sampling algorithm in Section 4, where the first 2,000 samples are

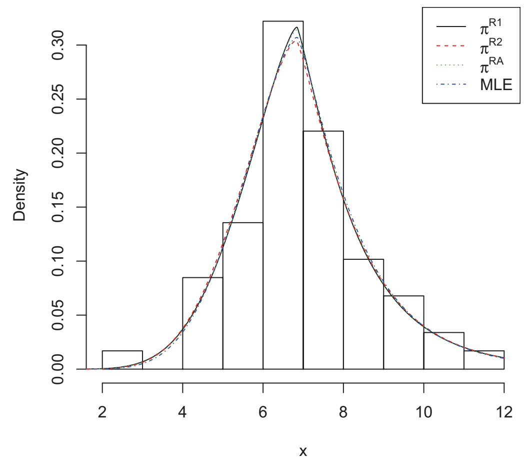

Figure 6.1: Fitted curves and data histogram.

discarded as burn-in periods with a thinning of 10 draws. The posterior summaries of each parameter are displayed in Table 6.1. We observe that the point estimates of each parameter are close to each other. Figure 6.1 shows the fit of the predictive densities of the logGN distribution evaluated at the different point estimates in Table 6.1. It can be seen from the figure that the estimated densities of the logGN distribution fits the data quite well and that they are are almost overlapping.

Table 6.1
Point estimate of each parameter for the data given in Table 2 of [18].

Parameter	π^{R1}		π^{R2}		π^{RA}		MLE
	Mean	Median	Mean	Median	Mean	Median	
μ	1.9236	1.9240	1.9223	1.9225	1.9240	1.9246	1.9260
σ	0.2551	0.2492	0.2667	0.2624	0.2590	0.2519	0.2600
s	1.3519	1.2858	1.4137	1.3494	1.3732	1.3088	1.3762

Chapter 7

Concluding remarks

we have studied Bayesian analysis of the logGN distribution under the various objective priors constructed using some formal rules. Specifically, we have derived different types of objective priors, including the Jeffreys-type priors, the reference priors based on the different group orderings of the parameters, and the first-order matching priors. It has been shown that some of the commonly used objective priors (such as the Jeffrey-rule prior) preclude the existence of proper posterior distributions due to their relative heavy tails for the large values of s. This result may be the opposite of a general sense that the Jeffreys prior almost always lead to a proper posterior distribution; see, for example, [2], [27], among others. Thus, special attention should be paid, especially when we adopt improper priors for the unknown parameters in practical situations.

We have shown that the two types of the reference priors in Theorem 2 result in proper posterior distributions. The problem is that which of the two reference priors should be recommenced in many practical situations. Because the prior π^{R1} is a first-order matching prior when s is the parameter of interest, we have a preference for this reference prior. Numerical simulations also show that the performance of the Bayesian approach under the prior π^{R1} is superior than the ones under the priors π^{R2} and π^{RA}.

One possible extension to our work is to extend the proposed Bayesian procedures to the censored data, which are commonly occurred from reliability tests. [24] recently considered Byasian analysis of the logGN distribution for the censored data using subjective priors for the unknown parameters. Here, subjective priors are normally gotten from the experimenter's learning about the conduct of arbitrary procedures under thought. Without former learning, using Beysian procedure to analyse the censored data is right now under scrutiny and will be accounted for in future work.

References

[1] Aitchison, J. and Brown, J. (1957). *The Lognormal Distribution.* Cambridge University Press.

[2] Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Anal.* **1**, 385–402.

[3] Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors. In *Bayesian statistics, 4 (Peñíscola, 1991).* New York: Oxford Univ. Press, pp. 35–60.

[4] Berger, J. O. and Bernardo, J. M. (1992). Ordered group reference priors with application to the multinomial problem. *Biometrika* **79**, 25–37.

[5] Bernardo, J.-M. (1979). Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc. Ser. B* **41**, 113–147. With discussion.

[6] Box, G. E. P. and Tiao, G. C. (1962). A further look at robustness via Bayes's theorem. *Biometrika* **49**, 419–432.

[7] CHEN, C. (2006). Tests of fit for the three-parameter lognormal distribution. *Comput. Statist. Data Anal.* **50**, 1418–1440.

[8] COX, D. R. and REID, N. (1987). Parameter orthogonality and approximate conditional inference. *J. Roy. Statist. Soc. Ser. B* **49**, 1–39. With a discussion.

[9] CROW, E. L. and SHIMIZU, K., eds. (1988). *Lognormal Distributions*, vol. 88 of *Statistics: Textbooks and Monographs*. Marcel Dekker, Inc., New York. Theory and applications.

[10] DATTA, G. S., GHOSH, M. and MUKERJEE, R. (2000). Some new results on probability matching priors. *Calcutta Statist. Assoc. Bull.* **50**, 179–192.

[11] DATTA, G. S. and SWEETING, T. J. (2005). Probability matching priors. In *Bayesian thinking: modeling and computation*, vol. 25 of *Handbook of Statist.* Elsevier/North-Holland, Amsterdam, pp. 91–114.

[12] DEVROYE, L. (1986). *Non-Uniform Random Variate Generation.* Springer-Verlag.

[13] FERREIRA, M. A. R. and SALAZAR, E. (2014). Bayesian reference analysis for exponential power regression models. *Journal of Statistical Distributions and Applications* , 1–12.

[14] FONSECA, T. C. O., FERREIRA, M. A. R. and MIGON, H. S. (2014).

'Objective Bayesian analysis for the Student-t regression model' [erratum to mr2521587]. *Biometrika* **101**, 252.

[15] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian statistics, 4 (Peñíscola, 1991).* Oxford Univ. Press, New York, pp. 169–193.

[16] Jeffreys, H. (1998). *Theory of probability.* Oxford Classic Texts in the Physical Sciences. The Clarendon Press, Oxford University Press, New York. Reprint of the 1983 edition.

[17] Liang, F., Liu, C. and Wang, N. (2007). A robust sequential Bayesian method for identification of differentially expressed genes. *Statist. Sinica* **17**, 571–597.

[18] Martín, J. and Pérez, C. J. (2009). Bayesian analysis of a generalized lognormal distribution. *Comput. Statist. Data Anal.* **53**, 1377–1387.

[19] Mukerjee, R. and Dey, D. K. (1993). Frequentist validity of posterior quantiles in the presence of a nuisance parameter: higher order asymptotics. *Biometrika* **80**, 499–505.

[20] Mukerjee, R. and Ghosh, M. (1997). Second-order probability matching priors. *Biometrika* **84**, 970–975.

[21] Raftery, A. E. and Lewis, S. M. (1992). [Practical Markov Chain Monte

Carlo]: Comment: One long run with diagnostics: Implementation strategies for Markov Chain Monte Carlo. *Statistical Science* **7**, 493–497.

[22] Salazar, E., Ferreira, M. A. R. and Migon, H. S. (2012). Objective Bayesian analysis for exponential power regression models. *Sankhya B* **74**, 107–125.

[23] Schafft, H., Staton, T., Mandel, J. and Shott, J. (1987). Reproducibility of electromigration measurements. *IEEE Transactions on Electron Devices (Impact Factor: 2.47)* **34(3)**, 673 – 681.

[24] Singh, B., Sharma, K. K., Rathi, S. and Singh, G. (2012). A generalized log-normal distribution and its goodness of fit to censored data. *Comput. Statist.* **27**, 51–67.

[25] Wang, M. and Lu, T. (2015). A matching prior for the shape parameter of the exponential power distribution. *Statist. Probab. Lett.* **97**, 150–154.

[26] West, M. (1984). Outlier models and prior distributions in Bayesian linear regression. *J. Roy. Statist. Soc. Ser. B* **46**, 431–439.

[27] Yang, R. and Berger, J. O. (1997). A catalog of noninformative priors. ISDS Discussion paper 97-42, Institute of Statistics and Decision Sciences, Duke University.

Appendix A

Proof

A.1 Proof of Theorem 1

Proof of Theorem 1: From the Fisher information of the logGN distribution in (2.1), we can get:

† For the independence Jeffreys prior $\pi^{\mathrm{J1}}(\theta)$:

$$\pi^{\mathrm{J1}}(\mu, \sigma, s) \;=\; \pi^{\mathrm{J1}}(\mu)\;\pi^{\mathrm{J1}}(\sigma, s),$$

where $\pi^{\mathrm{J1}}(\mu) \propto \sqrt{\det(H_{\mu\mu})} \propto 1$, and

$$\begin{aligned}\pi^{\mathrm{J1}}(\sigma, s) &\propto \sqrt{H_{\sigma\sigma}H_{ss} - H_{\sigma s}^2} \\ &= \sqrt{s^{-2}\sigma^{-2}[A^2+B] - A^2\sigma^{-2}s^{-2}} \\ &= \sigma^{-1}s^{-1}[B]^{1/2}.\end{aligned}$$

† The independence Jeffreys prior is as following:

$$\pi^{\mathrm{J2}}(\theta) = \pi^{\mathrm{J2}}(\mu)\pi^{\mathrm{J2}}(\sigma)\pi^{\mathrm{J2}}(s).$$

From the Fisher information matrix, we have $\pi^{\mathrm{J2}}(\mu) \propto 1$, $\pi^{\mathrm{J2}}(\sigma) \propto \sigma^{-1}$, and

$$\pi^{\mathrm{J2}}(s) \propto \; s^{-3/2}[A^2+B]^{1/2},$$

which provides that

$$\pi^{\mathrm{J2}}(\mu, \sigma, s) \; \propto \sigma^{-1}s^{-3/2}[A^2+B]^{1/2}.$$

† For Jeffreys-rule prior $\pi^J(\mu, \sigma, s)$:

$$\pi^J(\mu, \sigma, s) \; \propto \; \sqrt{\det[H(\theta)]} \; = \; \sqrt{H_{\sigma\sigma}H_{ss} - H_{\sigma s}^2}\sqrt{\det(H_{\mu\mu})},$$

where $\det(H_{\mu\mu}) = \frac{(s-1)s\Gamma(1-s^{-1})}{\sigma^2\Gamma(s^{-1})}$. Therefore, it follows that

$$\begin{aligned}
\pi^J(\mu, \sigma, s) \propto & \left[\frac{(s-1)s\Gamma(1-1/s)}{\sigma^2\Gamma(1/s)}\right]^{1/2} \pi^{\mathrm{J1}}(\sigma, s) \\
= & \sigma^{-1}\left[\frac{(s-1)s\Gamma(1-1/s)}{\Gamma(1/s)}\right]^{1/2} \pi^{\mathrm{J1}}(\sigma, s).
\end{aligned}$$

A.2 Proof of Theorem 2

Proof of Theorem 2: We derive the reference prior for the ordering (σ, μ, s) in part 1 and the ordering (s, μ, σ) in part 2. Since derivations of reference priors for other orderings are similar and are thus omitted for simplicity.

† For the ordering $\theta = (\sigma, \mu, s)$, the new Fisher information matrix changes to

$$\begin{aligned}
S(\theta) = & H^{-1}(\theta) \\
= & \begin{pmatrix} \frac{\sigma^2[A^2+B]}{\psi'(1+s^{-1})(1+s)-s} & 0 & \frac{As\sigma}{B} \\ 0 & \frac{\sigma^2\Gamma(s^{-1})}{(-1+s)s\Gamma(1-s^{-1})} & 0 \\ \frac{As\sigma}{B} & 0 & \frac{s^3}{B} \end{pmatrix}.
\end{aligned}$$

Thus, we obtain

$$S_1 = \frac{\sigma^2[A^2+B]}{\psi'(1+s^{-1})(1+s)-s},$$

$$S_2 = \begin{pmatrix} \frac{\sigma^2[A^2+B]}{\psi'(1+s^{-1})(1+s)-s} & 0 \\ 0 & \frac{\sigma^2\Gamma(s^{-1})}{(-1+s)s\Gamma(1-s^{-1})} \end{pmatrix},$$

and $S_3 = S(\theta)$. Moreover, if we let $H_j = S_j^{-1}$, we have

$$H_1 = \frac{\psi'(1+s^{-1})(1+s)-s}{\sigma^2[A^2+B]},$$

$$H_2 = \begin{pmatrix} \frac{\psi'(1+s^{-1})(1+s)-s}{\sigma^2[A^2+B]} & 0 \\ 0 & \frac{(s-1)s\Gamma(1-s^{-1})}{\sigma^2\Gamma(s^{-1})} \end{pmatrix},$$

and $H_3 = H(\theta)$. Let h_j be the $n_j \times n_j$ lower right corner of H_j. Then, it follows

$$h_1 = \frac{\psi'(1+s^{-1})(1+s)-s}{\sigma^2[A^2+B]},$$

$$h_2 = \frac{(s-1)s\Gamma(1-s^{-1})}{\sigma^2\Gamma(s^{-1})},$$

$$h_3 = [A^2+B]s^{-3}.$$

Following the procedures and notions in [13], we have

$$
\begin{aligned}
\pi_3^l(s \mid \sigma,\mu) =& \pi_3^l(\theta_{[-2]} \mid \theta_{[2]}) \\
=& \frac{|h_3(\theta)|^{1/2}\mathbf{I}_{\Theta^l_{(3)}}(\theta_{(3)})}{\int_{\Theta^l_{(3)}} |h_3(\theta)|^{1/2} d(\theta_{(3)})} \\
=& \frac{\left|[A^2+B]s^{-3}\right|^{1/2}\mathbf{I}_{[1,l]}(s)}{\int_1^l \left|[A^2+B]s^{-3}\right|^{1/2} ds} \\
=& [c_1(l)]^{-1}\ ([A^2+B]s^{-3})^{1/2}\ \mathbf{I}_{[1,l]}(s),
\end{aligned}
$$

where $c_1(l) = \int_1^l \left|[A^2+B]s^{-3}\right|^{1/2}\ ds$. Moreover,

$$
\begin{aligned}
\pi_2^l(\mu, s \mid \sigma) =& \pi_2^l(\theta_{[-1]} \mid \theta_{[1]}) \\
=& \frac{\pi_3^l(\theta_{[-2]}|\theta_{[2]})\ \exp(0.5\mathbb{E}_2^l[\log|h_2(\theta)|\big|\theta_{[2]}])\ \mathbf{I}_{\Theta^l_{(2)}}(\theta_{(2)})}{\int_{\Theta^l_{(2)}} \exp(0.5\mathbb{E}_2^l[\log|h_2(\theta)|\big|\theta_{[2]}])\ d(\theta_{(2)})},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{E}_2^l[\log|h_2(\theta)|\big|\theta_{[2]}] =& \int_{\Theta^l_{(3)}} \log|h_2(\theta)|\ \pi_3^l(\theta_{[-2]} \mid \theta_{[2]})\ d\theta_{[-2]} \\
=& \int_1^l \log\left(\frac{\Gamma(1-s^{-1})s(s-1)}{\Gamma(s^{-1})\sigma^2}\right)\ [c_1(l)]^{-1} \\
& ([A^2+B]s^{-3})^{1/2}\ \mathbf{I}_{[1,l]}(s)ds \\
=& -2k\log\sigma + c_2(l),
\end{aligned}
$$

with

$$
\begin{aligned}
c_2(l) =& [c_1(l)]^{-1} \ \int_1^l \log\left(\frac{\Gamma(1-s^{-1})s(s-1)}{\Gamma(s^{-1})}\right) \\
& ([A^2+B]s^{-3})^{1/2} \ \mathbf{I}_{[1,l]}(s)ds.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\pi_2^l(\mu, s \mid \sigma) =& \frac{\pi_3^l(s \mid \sigma, \mu)\exp[0.5(-2k\log\sigma + c_2(l))] \ \mathbf{I}_{[-l,l]^k}(\mu)}{\int_{[-l,l]^k}\exp[0.5(-2k\log\sigma + c_2(l))]\mathbf{I}_{[-l,l]^k}d\mu} \\
=& \pi_3^l(s|\sigma,\mu) \ (2l)^{-k} \ \mathbf{I}_{[-l,l]^k}(\mu).
\end{aligned}
$$

Further, we obtain

$$
\begin{aligned}
& \pi_1^l(\sigma,\mu,s)\pi_1^l(\theta_{[-0]}|\theta_{[0]}) \\
& \qquad = \frac{\pi_2^l(\theta_{[-1]}|\theta_{[1]}) \ \exp(0.5\mathbb{E}_1^l[\log|h_1(\theta)|\big|\theta_{[1]}]) \ \mathbf{I}_{\Theta^l_{(1)}}(\theta_{(1)})}{\int_{\Theta^l_{(1)}} \ \exp(0.5\mathbb{E}_1^l[\log|h_1(\theta)|\big|\theta_{[1]}]) \ d(\theta_{(1)})},
\end{aligned}
$$

with

$$\mathbb{E}_1^l[\log |\ h_1(\theta)|\big|\ \theta_{[1]}] = \int_{\Theta^l_{[-1]}} \log |h_1(\theta)|\ \pi_2^l(\theta_{[-1]}|\theta_{[1]})\ d\theta_{[-1]}$$

$$= \int_{[-l,l]^k} \int_1^l \log\left(\frac{\psi^{'}(1+s^{-1})(1+s)-s}{\sigma^2[A^2+B]}\right)$$

$$\pi_2^l(\mu, s|\sigma)\, ds\, d\mu$$

$$= -\,2\log\sigma + c_3(l),$$

where

$$c_3(l) = \int_{[-l,l]^k} \int_1^l \log\left(\frac{\psi^{'}(1+s^{-1})(1+s)-s}{[A^2+B]}\right)\pi_2^l(\mu, s \mid \sigma)\, ds\, d\mu,$$

does not depend on $\theta = (\sigma, \mu, s)$. Hence,

$$\pi_1^l(\sigma,\mu,s) = \frac{\pi_2^l(\theta_{[-1]}|\theta_{[1]})\ \exp[0.5(-2\log\sigma + c_3(l))]\ \mathbf{I}_{\Theta^l_{(1)}}(\theta_{(1)})}{\int_{\Theta^l_{(1)}} \exp[0.5(-2\log\sigma + c_3(l))]\ d(\theta_{(1)})}$$

$$= \frac{\pi_2^l(\mu, s|\sigma)\ \sigma^{-1}\ \mathbf{I}_{[l^{-1},l]}(\sigma)}{\int_{l^{-1}}^l \sigma^{-1}\ d\sigma}$$

$$= \frac{\pi_2^l(\mu, s|\sigma)\ \sigma^{-1}\ \mathbf{I}_{[l^{-1},l]}(\sigma)}{2\log l}.$$

Thus,

$$
\begin{aligned}
\pi_1^l(s,\mu,\sigma) =& \pi_3^l(s|\sigma,\mu)\ (2l)^{-k}\ \mathbf{I}_{[-l,l]^k}(\mu)\ \sigma^{-1}\ \mathbf{I}_{[l^{-1},l]}(\sigma)\ (2\log l)^{-1} \\
=& \sigma^{-1}\ ([A^2+B]s^{-3})^{1/2} \\
& \times c_1(l)(2l)^{-k}\ (2\log l)^{-1}\ \mathbf{I}_{[1,l]}(s)\ \mathbf{I}_{[-l,l]^k}(\mu)\ \mathbf{I}_{[l^{-1},l]}(\sigma).
\end{aligned}
$$

Now we take any point $\theta^* = (\sigma^*,\mu^*,s^*) \in [l^{-1},l] \times [-l,l]^k \times [1,l]$. Then, the reference prior for the ordering (σ,μ,s) is given by

$$
\begin{aligned}
\pi(\sigma,\mu,s) \propto& \lim_{l\to\infty} \frac{\pi_1^l(\sigma,\mu,s)}{\pi_1^l(\sigma^*,\mu^*,s^*)} \\
=& \sigma^{-1}\ s^{-3/2}\ [A^2+B]^{1/2}
\end{aligned}
$$

† For the ordering $\theta = (s, \mu, \sigma)$, we have the Fisher information matrix:

$$S(\theta) = H^{-1}(\theta)$$

$$= \begin{pmatrix} \frac{s^3}{B} & 0 & \frac{As\sigma}{B} \\ 0 & \frac{\sigma^2\Gamma(1/s)}{(-1+s)s\Gamma(1-1/s)} & 0 \\ \frac{As\sigma}{B} & 0 & \frac{\sigma^2(\psi'(1+s^{-1})+s(\psi'(1+s^{-1})+A^2-1))}{s(s(-1+\psi'(1+s^{-1}))+\psi'(1+s^{-1}))} \end{pmatrix}.$$

Thus,

$$S_1 = \frac{s^3}{B},$$

$$S_2 = \begin{pmatrix} \frac{s^3}{B} & 0 \\ 0 & \frac{\sigma^2\Gamma(1/s)}{(-1+s)s\Gamma(1-1/s)} \end{pmatrix},$$

and $S_3 = S(\theta)$. Moreover, let $H_j = S_j^{-1}$. Thus,

$$H_1 = \frac{B}{s^3},$$

$$H_2 = \begin{pmatrix} \frac{B}{s^3} & 0 \\ 0 & \frac{(s-1)s\Gamma(1-1/s)}{\sigma^2\Gamma(1/s)} \end{pmatrix},$$

and $H_3 = H(\theta)$.

$$
\begin{aligned}
h_1 &= \frac{B}{s^3}, \\
h_2 &= \frac{(s-1)s\Gamma(1-1/s)}{\sigma^2\Gamma(1/s)}, \\
h_3 &= \frac{s}{\sigma^2}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\pi_3^l(\sigma \mid s, \mu) =& \pi_3^l(\theta_{[-2]}|\theta_{[2]}) \\
=& \frac{|h_3(\theta)|^{1/2}\mathbf{I}_{\Theta_{(3)}^l}(\theta_{(3)})}{\int_{\Theta_{(3)}^l} |h_3(\theta)|^{1/2}\ d(\theta_{(3)})} \\
=& \frac{|s\sigma^{-2}|^{1/2}\mathbf{I}_{[l^{-1},l]}(\sigma)}{\int_{l^{-1}}^{l} |s\sigma^{-2}|^{1/2}\ d\sigma} \\
=& \sigma^{-1}\ (2\log l)^{-1}\mathbf{I}_{[l^{-1},l]}(\sigma).
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\pi_2^l(\mu, \sigma|s) &= \pi_2^l(\theta_{[-1]}|\theta_{[1]}) \\
&= \frac{\pi_3^l(\theta_{[-2]}|\theta_{[2]})\ \exp(0.5\mathbb{E}_2^l[\log|h_2(\theta)|\big|\theta_{[2]}])\ \mathbf{I}_{\Theta_{(2)}^l}(\theta_{(2)})}{\int_{\Theta_{(2)}^l}\ \exp(0.5\mathbb{E}_2^l[\log|h_2(\theta)|\big|\theta_{[2]}])\ d(\theta_{(2)})}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{E}_2^l[\log|h_2(\theta)|\big|\theta_{[2]}] =& \int_{\Theta^l_{(3)}} \log|h_2(\theta)|\ \pi_3^l(\theta_{[-2]}|\theta_{[2]})\ d\theta_{[-2]} \\
=& \int_{l^{-1}}^{l} \log\left(\frac{\Gamma(1-s^{-1})s(s-1)}{\Gamma(s^{-1})\sigma^2}\right)\ (2\sigma\log l)^{-1}\ \mathbf{I}_{[l^{-1},l]}(\sigma)\ d\sigma \\
=& c_1(l,s),
\end{aligned}
$$

which does not depend on μ. Hence,

$$
\begin{aligned}
\pi_2^l(\mu,\sigma\mid s) =& \frac{\pi_3^l(\sigma|s,\mu)\ \ \exp[0.5c_1(l,s)]\ \mathbf{I}_{[-l,l]^k}(\mu)}{\int_{[-l,l]^k}\ \exp[0.5c_1(l,s)]\ \mathbf{I}_{[-l,l]^k}\ d\mu} \\
=& \pi_3^l(\sigma|s,\mu)\ (2l)^{-k}\ \mathbf{I}_{[-l,l]^k}(\mu).
\end{aligned}
$$

Further,

$$
\begin{aligned}
\pi_1^l(s,\mu,\sigma) =& \pi_1^l(\theta_{[-0]}|\theta_{[0]}) \\
=& \frac{\pi_2^l(\theta_{[-1]}|\theta_{[1]})\exp(0.5\mathbb{E}_1^l[\log|h_1(\theta)|\big|\theta_{[1]}])\mathbf{I}_{\Theta^l_{(1)}}(\theta_{(1)})}{\int_{\Theta^l_{(1)}}\exp(0.5\mathbb{E}_1^l[\log|h_1(\theta)|\big|\theta_{[1]}])d(\theta_{(1)})},
\end{aligned}
$$

with

$$\mathbb{E}_1^l[\log|h_1(\theta)||\theta_{[1]}] = \int_{\Theta^l_{[-1]}} \log|h_1(\theta)|\pi_2^l(\theta_{[-1]}|\theta_{[1]})\ d\theta_{[-1]}$$

$$= \int_{[-l,l]^k}\int_{l^{-1}}^{l} \log(s^{-3}[B])\ \pi_2^l(\mu,\sigma|s)\ d\sigma\ d\mu$$

$$= \log(s^{-3}\ [B]).$$

Hence,

$$\pi_1^l(s,\mu,\sigma) = \frac{\pi_2^l(\theta_{[-1]}|\theta_{[1]})\ \exp(0.5\log(s^{-3}[B]))\ \mathbf{I}_{\Theta^l_{(1)}}(\theta_{(1)})}{\int_{\Theta^l_{(1)}} \exp(0.5\log(s^{-3}[B]))\ d(\theta_{(1)})}$$

$$= \frac{\pi_2^l(\theta_{[-1]}|\theta_{[1]})\ (s^{-3}[B])^{1/2}\ \mathbf{I}_{\Theta^l_{(1)}}(\theta_{(1)})}{\int_{\Theta^l_{(1)}} \exp(0.5\log(s^{-3}[B]))\ d(\theta_{(1)})}$$

$$= \pi_2^l(\theta_{[-1]}|\theta_{[1]})\ (s^{-3}\ [B])^{1/2}\ c_2(l)\ \mathbf{I}_{\Theta^l_{(1)}}(\theta_{(1)})$$

$$= \pi_2^l(\mu,\sigma|s)\ (s^{-3}\ [B])^{1/2}\ c_2(l)\ \mathbf{I}_{[1,l]}(s),$$

where $[c_2(l)]^{-1} = \int_1^l \exp(0.5\log(s^{-3}[B]))\ d(s)$. Thus,

$$\pi_1^l(s,\mu,\sigma) = (2\sigma\log l)^{-1}\ (s^{-3}[B])^{1/2}$$

$$\times\ c_2(l)(2l)^{-k}\ \mathbf{I}_{[1,l]}(s)\ \mathbf{I}_{[-l,l]^k}(\mu)\ \mathbf{I}_{[l^{-1},l]}(\sigma).$$

Finally,

$$\pi(s,\mu,\sigma) \propto \lim_{l\to\infty} \frac{\pi_1^l(s,\mu,\sigma)}{\pi_1^l(s^*,\mu^*,\sigma^*)}$$

$$= \sigma^{-1}\ s^{-3/2}\ [B]^{1/2}.$$

This completed the proof.

A.3 Proof of Lemma 1

† Consider the ordering $\theta = (\sigma,\mu,s)$. From Theorem 1, we obtain that the marginal prior of s is

$$\pi^{\mathrm{R1}}(s) \propto s^{-3/2}[A^2+B]^{1/2}.$$

which provides

$$[\pi^{\mathrm{R1}}(s)]^2 \ \propto\ s^{-3}[A^2+B].$$

Since $\Gamma(a) \approx e^{-a}a^{a-1/2}\sqrt{2\pi}$ (Abramowitz and Stegun, 1986, p. 257), we obtain that

$$
\begin{aligned}
&[1+\psi(1+1/s)]^2 \\
=&1+2\psi(1+1/s)+[\psi(1+1/s)]^2 \\
=&1+2\log(1+1/s)-(1+1/s)^{-1}+[\log(1+1/s)-[2(1+1/s)]^{-1}]^2.
\end{aligned}
$$

Since $\log(1+1/s) \to 0$ as $s \to \infty$ and $(1+1/s)^{-1} \to 1$ as $s \to \infty$,

$$[1+\psi(1+s^{-1})]^2 \;\approx\; \frac{1}{4}.$$

Also, $\psi'(a) \;\approx\; a^{-1}+(2a^2)^{-1}$ for large a. Thus, for small s, it follows

$$\psi'(1+1/s) \;\approx\; (1+1/s)^{-1}+[2(1+1/s)^2]^{-1} \;\approx\; \frac{3}{2}.$$

Hence, $[\pi^{\mathrm{R1}}(s)]^2 \;\approx\; s^{-3}[\frac{3}{2}+\frac{1}{4}-1] \;=\; \frac{3}{4}s^{-3} \;=\; O(s^{-3})$. Therefore,

$$\pi^{\mathrm{R1}}(s) \;=\; O(s^{-3/2}).$$

† Consider the ordering $\theta = (s,\mu,\sigma)$. From Theorem 1, we obtain that the marginal prior is

$$\pi^{\mathrm{R2}}(s) \;\propto\; s^{-3/2}[B]^{1/2}.$$

Then, $[\pi^{\mathrm{R2}}(s)]^2 \;\propto\; s^{-3}[B]$. Hence,

$$[\pi^{\mathrm{R2}}(s)]^2 \;\approx\; s^{-3}[1+2(1+1/s)^{-1}-1] \;=\; \frac{1}{2}s^{-3}(1+1/s)^{-1}.$$

Since $(1+1/s)^{-1} \to 1$ as $s \to \infty$, $[\pi^{\mathrm{R2}}(s)]^2 \;\approx\; \frac{1}{2}s^{-3} \;=\; O(s^{-3})$. Therefore,

$$\pi^{\mathrm{R2}}(s) \;=\; O(s^{-3/2}).$$

† For marginal independence Jeffreys prior $\pi^{I_1}(s)$. Since $\pi^{I_1}(s) \;\propto\; s^{-1}[(1+s^{-1}\psi'(1+s^{-1})-1]^{1/2}$, we obtain

$$[\pi^{I_1}(s)]^2 \;\propto\; s^{-2}[B].$$

From the previous proof, we know $[B] \;\approx\; \frac{1}{2}$. Hence,

$$[\pi^{I_1}(s)]^2 \;\approx\; \frac{1}{2}s^{-2} \;=\; O(s^{-2}).$$

Therefore,

$$\pi^{I_1}(s) \;=\; O(s^{-1}).$$

† For marginal independence Jeffreys prior $\pi^{I_2}(s)$. This is the same as the marginal reference prior of ordering $\theta = (\sigma, \mu, s)$, so its proof is omitted for simplicity.

† For marginal Jeffreys-rule prior $\pi^J(s)$. We know that

$$\pi^J(s) \;\propto\; \left[\frac{s(s-1)\Gamma(1-s^{-1})}{\Gamma(s^{-1})}\right]^{k/2} \pi^{I_1}(s).$$

Then,

$$[\pi^J(s)]^2 \;\propto\; \left[\frac{s(s-1)\Gamma(1-s^{-1})}{\Gamma(s^{-1})}\right]^{k} [\pi^{I_1}(s)]^2.$$

Thus, as $s \to \infty$, we have $\Gamma(s^{-1}) \approx s$. Also, $\Gamma(1-s^{-1}) \to \Gamma(1) = 1$ as $s \to \infty$.

Hence,

$$\frac{\Gamma(1-1/s)}{\Gamma(1/s)} \;\approx\; 1/s.$$

Since $[\pi^{I_1}(s)]^2 \;\approx\; \frac{1}{2}s^{-2}$,, it follows

$$[\pi^J(s)]^2 \;\approx\; [\frac{s(s-1)}{s^{-1}}]^k(\frac{1}{2}s^{-2}) \;\propto\; (s-1)^k(s^{-2}) = O(s^{k-2}).$$

Therefore,

$$\pi^J(s) = O(s^{(k-2)/2}).$$

A.4 Proof of Lemma 2

Proof of Lemma 2 First, we notice that σ can be integrated out analytically.

$$
\begin{aligned}
L^I(\mu, s; y) &= \int_o^\infty L(\mu, \sigma, s : y)\pi(\sigma) d\sigma \\
&= \int_o^\infty (2\sigma)^{-n}\ s^n\ [\Gamma(1/s)]^{-n}\ \exp\Bigg(-\sum_{i=1}^n \Bigg|\frac{\log(y_i) - \mu}{\sigma}\Bigg|^s\Bigg)\ \sigma^{-a}\ d\sigma \\
&\propto s^{-1}\left(\frac{s}{\Gamma(s^{-1})}\right)^n\ \Gamma\left(\frac{n+a-1}{s}\right)\left\{\sum_{i=1}^n \Big|\log(y_i) - \mu\Big|^s\right\}^{-\frac{n+a-1}{s}}.
\end{aligned}
$$

Then, with $\pi(\mu) \propto 1$, the integrated likelihood for s is

$$
\begin{aligned}
L^I(s; y) &= \int_{\mathbb{R}} L^I(\mu, s; y)\pi(\mu) d\mu \\
&\propto \frac{1}{s}\left\{\frac{s}{\Gamma(s^{-1})}\right\}^n \Gamma\left(\frac{n+a-1}{s}\right) \int_{\mathbb{R}} \left\{\sum_{i=1}^n \Big|\log(y_i) - \mu\Big|^s\right\}^{-\frac{n+a-1}{s}}\ d\mu.
\end{aligned}
$$

Following the proof of Lemma 2 ([22]), it can be easily showed that

$$
m_1(y)\ \le\ n^{\frac{n+a-1}{s}} \int_{\mathbb{R}} \left\{\sum_{i=1}^n \Big|\log(y_i) - \mu\Big|^s\right\}^{-\frac{n+a-1}{s}} d\mu\ \le\ m_2(y),
$$

where

$$
m_1(y) = \min\Big(\int (\max|\log(y_i) - \mu|)^{-(n+a-1)} d\mu, \int (\min|\log(y_i) - \mu|)^{-(n+a-1)} d\mu\Big)
$$

and

$$m_2(y) = \max\Big(\int (\max|\log(y_i) - \mu|)^{-(n+a-1)} d\mu, \int (\min|\log(y_i) - \mu|)^{-(n+a-1)} d\mu\Big),$$

which are independent of s. Then, we obtain

$$\int_{\mathbb{R}} \left\{ \sum_{i=1}^{n} \Big|\log(y_i) - \mu\Big|^s \right\}^{-\frac{n+a-1}{s}} d\mu = O(n^{-(n+a-1)/s}).$$

The result will allow us to analyse the tail behaviour of the integrated likelihood for s. We know that as $s \to \infty$, we have $\Gamma(s^{-1}) \approx s$. Therefore,

$$\begin{aligned} L^I(s; y) &\propto s^{-1} \left(\frac{s}{\Gamma(s^{-1})}\right)^n \Gamma\left(\frac{n+a-1}{s}\right) \int_{\mathbb{R}} \left\{ \sum_{i=1}^{n} \Big|\log(y_i) - \mu\Big|^s \right\}^{-\frac{n+a-1}{s}} d\mu \\ &= s^{-1}\, s^n\, \big[\Gamma(s^{-1})\big]^{-n}\, \Gamma\left(\frac{n+a-1}{s}\right) \int_{\mathbb{R}} \left\{ \sum_{i=1}^{n} \Big|\log(y_i) - \mu\Big|^s \right\}^{-\frac{n+a-1}{s}} d\mu \\ &\approx \frac{1}{n+a-1}\, O(n^{-(n+a-1)/s}) = O(1). \end{aligned}$$

This completed the proof.