Bayesian Hierarchical Scale Mixtures of Log-Normal Models for Inference in Reliability with Stochastic Constraint

Hea-Jung Kim

Department of Statistics, Dongguk University-Seoul, Pil-Dong 3Ga, Chung-Gu, Seoul 100-715, Korea;
kim3hj@dongguk.edu; Tel.: +82-2-2260-3221

Abstract: This paper develops Bayesian inference in reliability of a class of scale mixtures of log-normal failure time (SMLNFT) models with stochastic (or uncertain) constraint in their reliability measures. The class is comprehensive and includes existing failure time (FT) models (such as log-normal, log-Cauchy, and log-logistic FT models) as well as new models that are robust in terms of heavy-tailed FT observations. Since classical frequency approaches to reliability analysis based on the SMLNFT model with stochastic constraint are intractable, the Bayesian method is pursued utilizing a Markov chain Monte Carlo (MCMC) sampling based approach. This paper introduces a two-stage maximum entropy (MaxEnt) prior, which elicits a priori uncertain constraint and develops Bayesian hierarchical SMLNFT model by using the prior. The paper also proposes an MCMC method for Bayesian inference in the SMLNFT model reliability and calls attention to properties of the MaxEnt prior that are useful for method development. Finally, two data sets are used to illustrate how the proposed methodology works.

Keywords: Bayesian reliability analysis; Bayesian hierarchical model; MCMC method; scale mixtures of log-normal failure time model; stochastic constraint; two-stage MaxEnt prior

MSC: 62H05; 62F15; 94A17

1. Introduction

In some practical situations, reliability analysis must engage positive data. This includes, for example, cases requiring power plant system analysis, software reliability analysis, and hydrology (flows in a river) among others (see [1,2]). Any reliability analysis must be based on precisely defined model in order to estimate the reliability measure of interest from the available data and subsequently provide a logical basis for improving the reliability of a system. The exponential model is the most fundamental parametric model used to establish reliability. We refer to [3] for an excellent review of reliability analysis based on the exponential model. Complex failure time events, such as system failure events, appear to always produce a set of positively skewed observations and show, initially, an increase over time and, then, a decrease in hazard rate. In such cases, alternative to the exponential model, log-normal distribution (denoted by $\mathcal{LN}(\mu, \sigma^2)$) can be used to characterize and construct a more plausible model for assessing the failure time Y of the event. The density of $Y \sim \mathcal{LN}(\mu, \sigma^2)$ is

$$f(y|\mu, \sigma^2) = (y\,\sigma)^{-1}(2\pi)^{-1/2} \exp\left\{-\frac{(log(y)-\mu)^2}{2\sigma^2}\right\}, \ y > 0,$$

where $\mu = E[log(Y)]$ and $\sigma^2 = var(log(Y))$. In current statistics literature, the use of the log-normal distribution in reliability has become increasingly widespread. Ref. [4–7] used log-normal models

please confirm. for reliability and security analysis; Ref. [8–10] applied these models for survival analysis, loss reserving method, and hydrology, respectively. These studies were mainly concerned with the reliability of log-normal models with no constraints on the parameter space. There are also numerous studies in literature that deal with specific constrained parameter space problems in reliability analysis. See [11–13] for examples.

However, information on the constraint may sometimes be uncertain. In these cases, we have to analyze reliability of the log-normal model with a stochastic constraint on its parameter space, which causes restriction of reliability measures in their functional forms. For example, in a automobile system analysis, we may be interested in estimating MTTF (mean time to failure) of steering, brakes, or other mechanisms based on a log-normal model. Suppose it is known beforehand that such an automotive system exhibits MTTF $\in (a, b)$, but this information is uncertain. In spite of uncertain prior information, if we adopt a log-normal model whose parameter space is truncated to $\mu \in \mathcal{C}$ (i.e., MTTF $\in (a, b)$) for reliability analysis, we then have two problems. First, we pass over the uncertainty of the prior information associated with our inference so that we completely ignore possibility of $\mu \notin \mathcal{C}$ in the inference, even though empirical data strongly advocates this possibility. In this case, we may commit incorrect reliability analysis of the model. Second, we may lose data information, which strongly contradicts the uncertain information $\mu \in \mathcal{C}$. Furthermore, there are few valid robust models to assess poorly-distributed and fat-tailed failure time observations. A class of robust FT models can be obtained by the use of scale mixtures of log-normal models. The study about the robust model and its reliability estimation have not yet been tackled in the literature. These factors motivate the contents of the present paper.

The remainder of this paper is arranged as follows. Section 2 introduces a class of scale mixtures of log-normal failure time (SMLNFT) models by applying a scale mixture technique to a log-normal model. Subsequently, various class reliability measures are obtained. Section 3 provides a two-stage MaxEnt prior of μ by applying Boltzmann's maximum entropy theorem (see, e.g., [14,15]) to the two-stages of prior hierarchy frame according to [16]. We also introduce a scale (degree of prior belief) for stochastic constraint presumption $\mu \in \mathcal{C}$ that is accounted for by the MaxEnt prior. Section 4 develops the Bayesian hierarchical SMLNFT model by utilizing the two-stage MaxEnt prior hierarchy involving μ and a scale mixture hierarch of the SMLNFT model. This section also explores Bayesian inference in reliability for the proposed model utilizing an MCMC sampling based approach. This section also develops an MCMC sampling method based on the Gibbs sampler and the Metropolis–Hastings algorithm, and discusses Monte Carlo methods in posterior estimation of reliability measures. Section 5 illustrates the empirical performance of the proposed methodology based on a real and artificial data applications involving proposed SMLNFT models. Section 6 provides concluding remarks and discussion.

2. The Class of SMLNFT Models

Let failure time Y have a log-normal distribution with parameters $(\mu, \kappa(\eta)\sigma^2)$, denoted by $\mathcal{LN}(\mu, \kappa(\eta)\sigma^2)$, where $\kappa(\eta)$ is a suitably chosen positive weight function of a mixing variable $\eta \sim G(\eta)$ and $G(\eta)$ denotes the cumulative distribution function (cdf) of η. Subsequently, a simple location model for the failure time Y can be formulated in terms of $X \overset{d}{=} log(Y)$:

$$X = \mu + \varepsilon, \ \ \varepsilon \sim F, \tag{1}$$

where $F \in \mathcal{F}$ with

$$\mathcal{F} = \left\{ F : N\left(0, \kappa(\eta)\sigma^2\right), \ \eta \sim G(\eta) \text{ with } \kappa(\eta) > 0, \text{ and } \eta > 0 \right\} \tag{2}$$

so that the distribution of X is a scale mixtures of normal distributions, denoted by $X \sim \mathcal{SMN}(\mu, \sigma^2, \kappa, G)$. Refer to [17] for detailed properties of the scale mixtures of normal distributions.

The proposed model for the failure time Y can be defined as follows:

Definition 1. *Let the model (1) construct a plausible model for a random failure time Y. Then, the model for Y can then be referred to as an SMLNFT (scale mixtures of log-normal failure time) model. The distribution law of Y is a scale mixture of log-normal (SMLN) distributions with a weight function $\kappa(\eta)$ and the cdf G of a mixing variable η. This is written by $Y \sim \mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$.*

The density of $Y \sim \mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ is

$$f(y|\mu, \sigma^2) = \int_0^\infty (y\,\sigma)^{-1}(2\pi\kappa(\eta))^{-1/2} \exp\left\{ -\frac{(log(y)-\mu)^2}{2\kappa(\eta)\sigma^2} \right\} dG(\eta), \quad y > 0. \tag{3}$$

The SMLNFT model can vary depending on designation for the function $\kappa(\eta)$ and the distribution of η. In the special case where the distribution of η degenerates at $\kappa(\eta) = 1$, the SMLNFT model produces a log-normal failure time (LNFT) model. In cases where $\eta \sim Gamma(\nu/2, \nu/2)$ with $E[\eta] = 1$ and $\kappa(\eta) = 1/\eta$ are chosen, the SMLNFT model changes to log-t_ν failure time (Lt_νFT) or log-Cauchy failure time (LCFT $\equiv$ Lt_1FT) models, allowing for the regulation of model tail distribution by means of the degrees of freedom. The LCFT model has been particularly used for certain survival processes where significant outliers or extreme results may occur (see, e.g., [18]). We also see that the SMLNFT model approximately reduces to log-logistic failure time (LLFT) model (see [19]), provided that the choices are $\kappa(\eta) = 4\eta^2$ and $\eta^2 \sim IG(2.5, 1.233)$, where $IG(\alpha, \beta)$ is an inverse gamma distribution with a probability function $f(x) = \beta^\alpha x^{-(\alpha+1)} e^{\beta/x}/\Gamma(\alpha)$. This approximation is known to simplify the implementation of MCMC sampling. The LLFT model has been used in survival analysis as a parametric model for mortality rate, hydrology, and networked telerobots (see, e.g., [20]). Following the same procedure as in [17], we can construct new robust FT models through differing designations of $\kappa(\eta)$ and the distribution of η. Particularly, in the regulation of tail distribution in the FT model, a log-slash failure time (LSFT) model can be obtained by taking $\kappa(\eta) = \eta^{-1}$ and $\eta \sim Beta(1/2, 1)$—that is *Beta* distribution. Therefore, the proposed class of SMLNFT models defined by Equation (1) is flexible enough to include existing FT models (i.e., LNFT, LCFT, and LLFT) and also provide new robust failure time models (such as Lt_νFT and LSFT).

The densities of five $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ class members are shown in Figure 1 for different designations of the mixing distribution η and $\kappa(\eta)$. Figure 1 shows that the tail area of the log-normal density is thinner than that of the other densities. Thus, the use of $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ models is useful for flexible and robust reliability analysis of failure time data involving fat-tailed empirical distribution. Figure 1 also demonstrates that the class of $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ models is useful to describe those situations in which early failure or occurrences dominate the distribution.

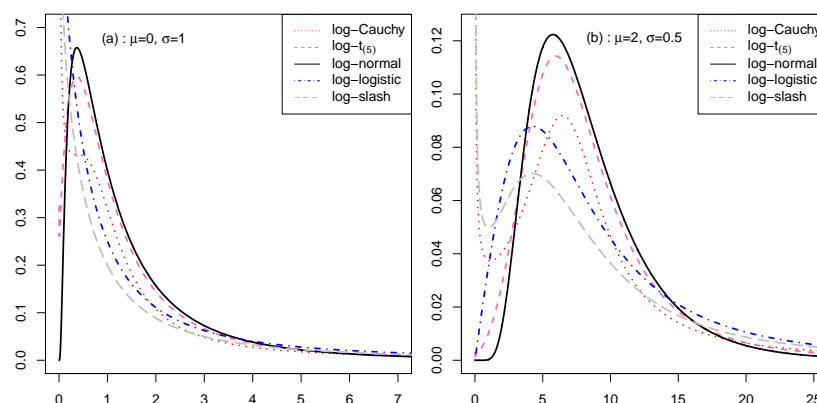**Figure 1.** Densities of $Y \sim \mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ for five different choices of the mixing variable distribution $G(\eta)$ and $\kappa(\eta)$: (**a**) $(\mu, \sigma) = (0, 1)$ case and (**b**) $(\mu, \sigma) = (2, 0.5)$ case.

Reliability and hazard rate of the class of SMLNFT models are directly related to those of the $\mathcal{SMN}(\mu, \sigma^2, \kappa, G)$ distribution. Let Y be failure time or down time formulated by the $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ model, and then respective MTTF (mean time to failure) and variance of Y be

$$E(Y) = E\left[\exp\left\{ \mu + \frac{\kappa(\eta)\sigma^2}{2} \right\} \right] \quad \text{and} \quad Var(Y) = E\left[\exp\left\{ 2\mu + \kappa(\eta)\sigma^2 \right\} \left(\exp\left\{ \kappa(\eta)\sigma^2 \right\} - 1 \right) \right]. \quad (4)$$

In addition, the reliability and hazard rates of the SMLNFT model are given by

$$R(y; \mu, \sigma^2) = 1 - E\left[\Phi\left(A(\eta) \right) \right] \tag{5}$$

and

$$h(y; \mu, \sigma^2) = \frac{E\left[\kappa(\eta)^{-1/2} \phi\left(A(\eta) \right) \right]}{\sigma \, y \, E\left[\left\{ 1 - \Phi\left(A(\eta) \right) \right\} \right]}, \tag{6}$$

where $A(\eta) = (\log(y) - \mu)/(\kappa(\eta)^{1/2}\sigma)$, $\eta \sim G(\eta)$, and ϕ and Φ are the density and the cdf of a standard normal variable, respectively. Exact expressions for Equations (5) and (6) based on the above five FT models are as follows:

(i) LNFT model:

$$R(y; \mu, \sigma^2) = 1 - \Phi\left(B \right) \quad \text{and} \quad h(y; \mu, \sigma^2) = \frac{\phi\left(B \right)}{\sigma \, y \, R(y; \mu, \sigma^2)},$$

where $B = (\log(y) - \mu)/\sigma$.

(ii) Lt_νFT model:

$$R(y; \mu, \sigma^2) = 1 - F_\nu\left(B \right) \quad \text{and} \quad h(y; \mu, \sigma^2) = \frac{f_\nu\left(B \right)}{\sigma \, y \, R(y; \mu, \sigma^2)},$$

where $F_\nu(\cdot)$ and $f_\nu(\cdot)$ denote the cdf and density function of a Student-t_ν distribution, respectively.

(iii) LCFT model:

$$R(y; \mu, \sigma^2) = \frac{1}{2} - \frac{\arctan\left(B \right)}{\pi} \quad \text{and} \quad h(y; \mu, \sigma^2) = \frac{1}{\sigma \, y\pi \left[1 + B^2 \right] R(y; \mu, \sigma^2)}.$$

(iv) LLFT model:

$$R(y; \mu, \sigma^2) = \frac{1}{1 + e^B} \quad \text{and} \quad h(y; \mu, \sigma^2) = \frac{e^B \, R(y; \mu, \sigma^2)}{\sigma \, y}.$$

(v) LSFT model:

$$R(y; \mu, \sigma^2) = \begin{cases} 1 - \Phi\left(B \right) + \left[\phi\left(0 \right) - \phi\left(B \right) \right] \big/ B & \text{for } B \neq 0, \\ 1/2 & \text{for } B = 0, \end{cases}$$

$$h(y; \mu, \sigma^2) = \begin{cases} \left[\phi\left(0 \right) - \phi\left(B \right) \right] \big/ \left\{ B^2 \sigma \, y \, R(y; \mu, \sigma^2) \right\} & \text{for } B \neq 0, \\ 2\left[\phi\left(0 \right) - \phi\left(B \right) \right] \big/ \left\{ B^2 \sigma \, y \right\} & \text{for } B = 0. \end{cases}$$

The reliable life T_γ, denoting the 100(1-γ)th percentile of the failure time distribution $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ of the SMLNFT model, may also be expressed as

$$T_\gamma = E\left[\exp\left\{ \kappa(\eta)^{1/2} \sigma \, \Phi^{-1}(1-\gamma) + \mu \right\} \right]. \tag{7}$$

The expectations made for the reliability measures (Equation (4) through Equation (7)) are given in respect to the distribution of η.

Figure 2 illustrates the reliability and hazard functions for each of five different $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ models (i.e., LNFT, LCFT, LLFT, Lt_5FT, and LSFT models) for the choice of $\mu = 2$ and $\sigma = 0.5$. It is apparent from the left panel of Figure 2 that all the reliability functions are monotone decreasing functions of y (time) and approaching zero in relation to sufficiently large values of times, but the shape of the functions are not the same. As expected from the densities of Figure 1, we see that each reliability function of heavy tailed models (i.e., LCFT, LLFT, Lt_5FT, and LSFT models) initially decreases more rapidly than that of the LNFT model; then, it decreases toward zero more slowly than that of the LNFT model. The right panel of Figure 2 describes hazard rates of the five models. The hazard rate of the LNFT model initially increases and then decreases toward zero as y (time) passes by. Hazard functions of the LLFT and Lt_5FT models resemble that of the LNFT model, aside from a rapid decrease toward zero after the initial increase. However, the right panel shows that the LCFT and LSFT models have different hazard functions from that of the LNFT model in that the initial increase of the hazard rate does not apply to these models.
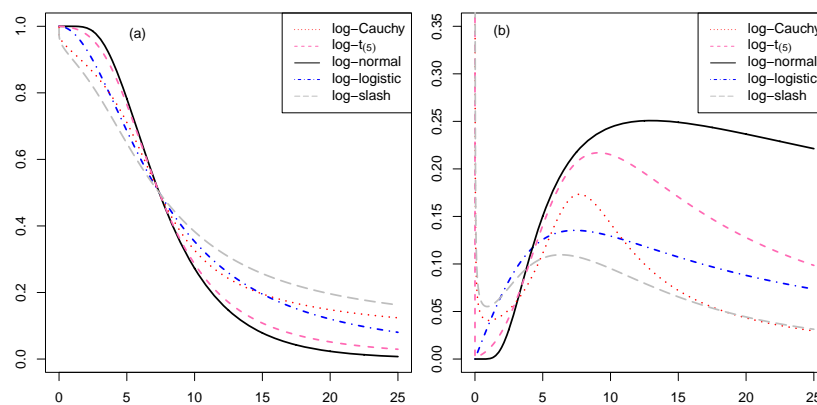

Figure 2. Reliability and hazard functions for five $\mathcal{SMLN}(\mu, \sigma^2, \kappa, G)$ models with $(\mu, \sigma) = (2, 0.5)$: (**a**) reliability functions, $R(y; \mu, \sigma^2)$; (**b**) hazard functions, $h(y; \mu, \sigma^2)$.

3. Two-Stage MaxEnt Prior

3.1. Stochastic Constraint on Reliability Measures

Let $y_1, \ldots, y_n$ be n complete (nontruncated/noncensored) failure times generated from the SMLNFT model in Equation (1), and let the reliability measures, defined by Equation (4) through Equation (7), have a stochastic functional constraint in terms of μ and σ^2, where the constraint does not depend on σ^2. Thus, a reliability measure, written by $B(\mu, \sigma^2)$, needs to be located in a restricted interval $B(\mu, \sigma^2) \in (a, b)$ (or equivalently $\mu \in \mathcal{C}$) with a degree of prior belief α, where $\mathcal{C} \subset \mathbb{R}$ is either an interval or a set of intervals. Nevertheless, observations from the SMLNFT model often do not provide strong evidence that information of the constraint, $\mu \in \mathcal{C}$, is true and therefore may appear to contradict the assumption of the model associated with the constraint.

In this situation, a Bayesian approach can be effectively adopted to model the stochastic (or uncertain) functional constraint on parameter μ. Bayesian reliability analysis of the model (1) begins

with specifying prior distribution $\pi(\mu, \sigma^2)$, which represents information concerning the parameters μ and σ^2 that are combined with the joint probability distribution of y_is to yield the posterior distribution:

$$p(\mu, \sigma^2 | y_1, \ldots, y_n) \propto \pi(\mu, \sigma^2) \prod_{i=1}^{n} f(y_i | \mu, \sigma^2),$$

where $f(y_i | \mu, \sigma^2)$ is the density given in Equation (3). When there are no constraints on the location parameter μ, then a common joint prior $\pi(\mu, \sigma^2)$ (e.g., Jeffreys prior or normal-inverse gamma prior) can be used, and posterior reliability inference can be performed without any difficulty. When we have sufficient evidence that the constraint condition on the model (1) is true (i.e., $\alpha = 1$), then a suitable restriction on the parameter space such as using a truncated prior distribution, e.g., $\pi(\mu, \sigma^2) I(\mu \in \mathcal{C})$, is expected. Here, $I(\cdot)$ denotes an indicator function. See, e.g., [21–24], for various applications of truncated prior distribution in Bayesian inference. However, it is often the case that prior information about the constraint is not certain to carry out Bayesian reliability analysis. In this case, it is expected that the uncertainty, i.e., $Pr(\mu \in \mathcal{C}) = \alpha$, is taken into account in eliciting a prior distribution of μ.

3.2. Two-Stage MaxEnt Prior

Assume that σ^2 is known and that it is possible to specify that partial information concerning the parameter $\mu \in \mathbb{R}$ is of the form

$$E[t_j(\mu)] = \int_{\mathbb{R}} t_j(\theta) \pi(\mu) d\mu = t_j, \quad j = 1, \ldots, k, \tag{8}$$

but nothing else about prior distribution $\pi(\mu)$. In this case, maximum entropy prior can be obtained by choosing $\pi(\theta)$ that maximizes the entropy

$$Ent(\pi) = - \int_{\mathbb{R}} \pi(\mu) \log \pi(\mu) d\mu,$$

in the presence of the partial information in the form of Label (8). Boltzmann's maximum entropy theorem (see, e.g., [14,15]) tells us that the density $\pi(\mu)$ that maximizes $Ent(\pi)$, subject to the constraints $E[t_j(\mu)] = t_j, j = 1, \ldots, k$, takes the k-parameter exponential family form

$$\pi(\mu) \propto \exp\left\{ \lambda_1 t_1(\mu) + \lambda_2 t_2(\mu) \cdots + \lambda_k t_k(\mu) \right\}, \quad \mu \in \mathbb{R}, \tag{9}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ can be determined, via the k-constraints, in terms of $t_1, \ldots, t_k$.

The theorem can be directly applied to obtain the MaxEnt prior $\pi(\mu)$ of the SMLNFT model, provided that prior information on the mean and variance of unconstrained μ can be specified, i.e., $E[\mu] = \theta_0$ and $E[(\mu - \theta_0)^2] = \sigma_0^2$. Supposing that we have additional information that $p(\mu \in \mathcal{C}) = \alpha$, $0 < \alpha < 1$, we then have the following lemma.

Lemma 1. *Suppose we can specify partial prior information about μ by $E[\mu] = \theta_0$, $E[(\mu - \theta_0)^2] = \sigma_0^2$, and $\mu \in \mathcal{C}$ with degree of belief α. Then, the MaxEnt prior density of μ is given by*

$$\pi_{MaxEnt}(\mu) = \phi(\mu; \theta_0, \sigma_0^2) \frac{\bar{\Phi}\left(\mathcal{C}; \theta_0 + \delta(\mu - \theta_0), \delta(1 - \delta)\sigma_0^2 \right)}{\bar{\Phi}(\mathcal{C}; \theta_0, \delta\sigma_0^2)}, \quad \mu \in \mathbb{R}, \tag{10}$$

where the value of δ ($\delta \in [0, 1]$) is determined by the following equation

$$\alpha = \bar{\Phi}_2\left(\mathcal{C} \times \mathcal{C}; \boldsymbol{\theta}_0, \Sigma_0 \right) / \bar{\Phi}(\mathcal{C}; \theta_0, \delta\sigma_0^2) \quad \text{with} \quad \boldsymbol{\theta}_0 = \begin{pmatrix} \theta_0 \\ \theta_0 \end{pmatrix} \quad \text{and} \quad \Sigma_0 = \sigma_0^2 \begin{pmatrix} \delta & \delta \\ \delta & 1 \end{pmatrix}. \tag{11}$$

Here, $\bar{\Phi}(\mathcal{C}; \theta_*, \sigma_*^2)$ is probability $p(W \in \mathcal{C})$ of $W \sim N(\theta_*, \sigma_*^2)$, and $\bar{\Phi}_2\left(\mathcal{C} \times \mathcal{C} : \boldsymbol{\theta}_0, \Sigma_0\right)$ denotes a joint probability $p(X_1 \in \mathcal{C}, X_2 \in \mathcal{C})$ of X_1 and X_2 whose joint distribution is a bivariate normal $N_2(\boldsymbol{\theta}_0, \Sigma_0)$.

Proof. The stochastic constraint $p(\mu \in \mathcal{C}) = \alpha$ can be expressed in terms of a moment as $E[I(\mu \in \mathcal{C})] = \alpha$. Take $t_1(\mu) = \mu$, $t_1 = \theta_0$, $t_2(\mu) = (\mu - \theta_0)^2$, $t_2 = \sigma_0^2$, $t_3(\mu) = I(\mu \in \mathcal{C})$, $t_3 = \alpha$, and $t_k(\mu) = 0$ for $k > 3$. Then, by setting $\lambda_1 = 0$ and $\lambda_2 = -1/(2\sigma_0^2)$, the MaxEnt prior of μ in Label (9) reduces to

$$\pi_{MaxEnt}(\mu) \quad \propto \quad \exp\left\{(\mu - \theta_0)^2/2\sigma_0^2\right\} \exp\left\{\lambda_3 t_3(\mu)\right\}, \quad \mu \in \mathbb{R}. \tag{12}$$

Now, the second exponential term on the right-hand side of Label (12) can be determined by using the stochastic constraint $E[I(\mu \in \mathcal{C})] = \alpha$: among all of the possible proper prior densities of the form (12), the choice of $\lambda_3 t_3(\mu) = \ln\left\{\bar{\Phi}\left(\mathcal{C}; \theta_0 + \delta(\mu - \theta_0), \delta(1 - \delta)\sigma_0^2\right)/\bar{\Phi}(\mathcal{C}; \theta_0, \delta\sigma_0^2)\right\}$ yields a proper prior density. Furthermore, this choice leads to Equation (11), which satisfies the stochastic constraint

$$E[I(\mu \in \mathcal{C})] = \int_{\mathcal{C}} \pi_{MaxEnt}(\mu) d\mu = \bar{\Phi}_2\left(\mathcal{C} \times \mathcal{C} : \boldsymbol{\theta}_0, \Sigma_0\right)/\bar{\Phi}(\mathcal{C}; \theta_0, \delta\sigma_0^2) = \alpha.$$

□

From Lemma 1, we see that $\delta = 0$ implies no functional stochastic constraint on μ of the SMLNFT model, while $\delta = 1$ denotes that the functional constraint on the model is certain so that $Pr(\mu \in \mathcal{C}) = 1$. We also see that $\pi_{MaxEnt}(\mu)$ reduces to the normal density $\pi_0(\mu) = \phi(\mu; \theta_0, \sigma_0^2)$ for the former case and the truncated normal density $\pi_1(\mu) = \phi(\mu; \theta_0, \sigma_0^2)/\bar{\Phi}(\mathcal{C}; \theta_0, \sigma_0^2)$ with the support $\mu \in \mathcal{C}$ for the latter case. This indicates that $\delta \in [0, 1]$ is directly related to the degree α of prior belief on the constraint $\mu \in \mathcal{C}$, where $\alpha = Pr(\mu \in \mathcal{C})$. The relationship between them is as follows.

Corollary 1. *The degree α of prior belief on the constraint $\mu \in \mathcal{C}$ accounted for by the MaxEnt prior $\pi_{MaxEnt}(\mu)$ is $\alpha = \bar{\Phi}_2\left(\mathcal{C} \times \mathcal{C} : \boldsymbol{\theta}_0, \Sigma_0\right)/\bar{\Phi}(\mathcal{C}; \theta_0, \delta\sigma_0^2)$, and its range is*

$$\bar{\Phi}(\mathcal{C}; \theta_0, \sigma_0^2) \le \alpha \le 1.$$

Proof. For the cases where $\delta = 0$ and $\delta = 1$, the degree of prior beliefs are $\int_{\mathcal{C}} \pi_{MaxEnt}(\mu) d\mu = \int_{\mathcal{C}} \pi_0(\mu) d\mu = \bar{\Phi}(\mathcal{C}; \theta_0, \sigma_0^2)$ and $\int_{\mathcal{C}} \pi_{MaxEnt}(\mu) d\mu = \int_{\mathcal{C}} \pi_1(\mu) = 1$, respectively. When $0 < \delta < 1$,

$$\bar{\Phi}(\mathcal{C}; \theta_0, \sigma_0^2) < \alpha < 1$$

because $\bar{\Phi}_2\left(\mathcal{C} \times \mathcal{C} : \boldsymbol{\theta}_0, \Sigma_0\right) > \bar{\Phi}(\mathcal{C}; \theta_0, \sigma_0^2)\bar{\Phi}(\mathcal{C}; \theta_0, \delta\sigma_0^2)$ by the Theorem of [25]. □

Corollary 1 indicates that $\pi_{MaxEnt}(\mu)$ is useful for eliciting priori uncertainty concerning constraint on μ by varying the value of δ. By applying the work of [26], we can organize $\pi_{MaxEnt}(\mu)$ within the frame of the two-stage prior hierarchy by [16].

Theorem 1. *The distribution of $\pi_{MaxEnt}(\mu)$ can be expressed as a two-stage prior hierarchy:*

$$\begin{aligned}\mu|\mu_0 &\sim N(\mu_0, (1 - \delta)\sigma_0^2), \\ \mu_0 &\sim TN_{\mathcal{C}}(\theta_0, \delta\sigma_0^2),\end{aligned}$$

where $\delta \in [0, 1]$ and $TN_{\mathcal{C}}(\theta_0, \delta\sigma_0^2)$ denote a truncated $N(\theta_0, \delta\sigma_0^2)$ distribution with the support $\mu_0 \in \mathcal{C}$.

Proof. From the hierarchy of the prior distributions, we see that

$$
\begin{aligned}
\pi(\mu) &= \int_{\mu_0 \in \mathcal{C}} p(\mu|\mu_0)p(\mu_0)d\mu_0 = \frac{\int_{\mu_0 \in \mathcal{C}} \phi(\mu;\mu_0(1-\delta)\sigma_0^2)\phi(\mu_0;\theta_0,\delta\sigma_0^2)d\mu_0}{Pr(\mu_0 \in \mathcal{C})} \\
&= \frac{\phi(\mu;\theta_0,\sigma^2)}{\bar{\Phi}(\mathcal{C};\theta_0,\delta\sigma^2)} \int_{\mu_0 \in \mathcal{C}} \phi\big(\mu_0;\theta_0 + \delta(\mu-\theta_0),\delta(1-\delta)\sigma_0^2\big)d\mu_0,
\end{aligned}
$$

which is equivalent to $\pi_{MaxEnt}(\mu)$. $\quad\square$

From now on, we shall call the MaxEnt prior $\pi_{MaxEnt}(\mu)$, expressed by the two-stage prior hierarchy in Theorem 1, as a "two-stage MaxEnt prior" of μ.

4. Bayesian Hierarchical SMLNFT Model

4.1. Bayesian Hierarchical Model

Let $(y_1,\ldots,y_n)$ be n complete failure times as generated from the SMLNFT model in Equation (1), and let η_i denote contribution of latent variable η in relation to each observation y_i. Suppose one of the reliability measures defined by Equation (4) through Equation (7) has a stochastic functional constraint in terms of μ, e.g., $B(\mu) \in (a,b)$ (or equivalently $\mu \in \mathcal{C}$), with the degree of prior belief α, and suppose $\mathcal{C} \subset \mathbb{R}$ is either an interval or a set of intervals. Then, from Equations (3) and (10), we obtain posterior distribution of μ and σ^2 given by

$$
\begin{aligned}
p(\mu,\sigma^2|y_1,\ldots,y_n) &\propto \prod_{i=1}^{n} \int_0^\infty \left[(y_i\,\sigma)^{-1}(2\pi\kappa(\eta_i))^{-1/2} \exp\left\{ -\frac{(log(y_i)-\mu)^2}{2\kappa(\eta_i)\sigma^2} \right\} \right] G(\eta_i)d\eta_i \\
&\times \frac{\phi(\mu;\theta_0,\sigma_0^2)\bar{\Phi}\big(\mathcal{C};\theta_0 + \delta(\mu-\theta_0),\delta(1-\delta)\sigma_0^2\big)}{\bar{\Phi}(\mathcal{C};\theta_0,\delta\sigma_0^2)} p(\sigma^2),
\end{aligned}
$$

where $p(\sigma^2)$ is a prior distribution of σ^2. This is a complex function for the Bayesian inference in reliability based on the SMLNFT model. Alternately, following hierarchical representation of the stochastically constrained SMLNFT Model is useful for simple Bayesian inference. Using the two-stage MaxEnt prior in Theorem 1 and assuming prior independence of μ and $\tau = 1/\sigma^2$, we formulate a Bayesian hierarchical model for the stochastically constrained SMLNFT model as follows:

$$
\begin{aligned}
y_i|(\mu,\tau,\eta_i) &\stackrel{d}{=} \exp\{x_i\}|(\mu,\tau,\eta_i),\ \ \tau = \sigma^{-2}\ \ i = 1,\ldots,n, \\
x_i|(\mu,\tau,\eta_i) &\stackrel{ind}{\sim} N(\mu,\kappa(\eta_i)\tau^{-1}), \\
\mu|\mu_0 &\sim N(\mu_0,(1-\delta)\sigma_0^2),\ \ 0 < \delta < 1, \\
\mu_0 &\sim TN(\theta_0,\delta\sigma_0^2)\ I\big((\mu_0 \in \mathcal{C}\big), \\
\tau &\sim Gamma(\nu_1,\nu_2), \\
\eta_i &\stackrel{ind}{\sim} G(\eta_i).
\end{aligned}
\tag{13}
$$

All hyper-parameters, $\{\theta_0,\sigma_0,\delta,\nu_1,\nu_2\}$, are assumed to be given from the prior information of previous studies or alternate sources. In particular, the value of δ is chosen to satisfy Equation (11) for given values of α, θ_0, and σ_0^2. In cases where the prior information is not available, a convenient strategy of avoiding improper posterior distribution is to use proper priors with fixed hyper-parameters as appropriate quantity to reflect the diffuseness of the priors (i.e., limiting non-informative priors).

4.2. The Gibbs Sampler

Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)^\top$ be the latent variables, and let $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ be logarithmic failure time data where $x_i = log(y_i)$. Based on the Bayesian Hierarchical SMLNFT model structure, the joint posterior distribution of $(\mu, \mu_0, \tau, \boldsymbol{\eta})$ given the observed data $\boldsymbol{x}$ is

$$
\begin{aligned}
p(\mu, \mu_0, \tau, \boldsymbol{\eta} | \boldsymbol{x}) \quad \propto \quad & \Big(\textstyle\prod_{i=1}^n \phi(x_i; \mu, \kappa(\eta_i)\tau^{-1}) \Big) \phi(\mu; \mu_0, (1-\delta)\sigma_0^2) \\
\times \quad & \frac{\phi(\mu_0; \theta_0, \delta\sigma_0^2) I(\mu_0 \in \mathcal{C})}{\Phi(\mathcal{C}; \theta_0, \delta\sigma_0^2)} f(\tau; \nu_1, \nu_2) \Big(\textstyle\prod_{i=1}^n g(\eta_i) \Big),
\end{aligned}
\tag{14}
$$

where $f(\tau; \nu_1, \nu_2)$ is the density of $Gamma(\nu_1, \nu_2)$ variate and $g(\eta_i)$ is the density of the mixing variable η_i. Note that the joint posterior distribution of the structure parameters, μ and τ, is

$$
p(\mu, \tau | \boldsymbol{x}) = \int_{\boldsymbol{\eta} > \boldsymbol{0}} \int_{\mathcal{C}} p(\mu, \mu_0, \tau, \boldsymbol{\eta} | \boldsymbol{x}) d\mu_0 d\boldsymbol{\eta}.
$$

This is not simplified in an analytic form of known density, and is thus intractable for posterior inference. Accordingly, we treat latent observations in $\boldsymbol{\eta}$ as hypothetical missing data, and augment the observed data set $\boldsymbol{x}$ with $\boldsymbol{\eta}$ in posterior analysis. Thus, we derive each conditional posterior distribution of μ, μ_0, τ and $\boldsymbol{\eta}$, for posterior inference based on Markov chain Monte Carlo (MCMC). All of the full conditional posterior distributions are as follows:

(1)   The full conditional distribution of μ is an univariate normal given by:

$$
[\mu | \mu_0, \tau, \boldsymbol{\eta}, \boldsymbol{x}] \quad \sim \quad N\Big(\frac{\tau^{-1}\mu_0 + \Delta \sum_{i=1}^n \kappa(\eta_i)^{-1} x_i}{\tau^{-1} + \Delta \sum_{i=1}^n \kappa(\eta_i)^{-1}}, \frac{1}{\Delta^{-1} + \tau \sum_{i=1}^n \kappa(\eta_i)^{-1}} \Big),
\tag{15}
$$

where $\Delta = (1-\delta)\sigma_0^2$.

(2)   The full conditional distribution of μ_0 is a truncated normal given by:

$$
[\mu_0 | \mu, \tau, \boldsymbol{\eta}, \boldsymbol{x}] \quad \sim \quad TN\Big(\frac{\Delta\theta_0 + \Delta_0 \mu}{\Delta + \Delta_0}, \frac{1}{\Delta^{-1} + \Delta_0^{-1}} \Big) I(\mu_0 \in \mathcal{C}),
\tag{16}
$$

where $\Delta_0 = \delta\sigma_0^2$.

(3)   The full conditional distribution of τ is a Gamma distribution:

$$
[\tau | \mu, \mu_0, \boldsymbol{\eta}, \boldsymbol{x}] \quad \sim \quad Gamma\Big(\nu_1 + \frac{n}{2}, \nu_2 + \frac{\sum_{i=1}^n \kappa(\eta_i)^{-1}(x_i - \mu)^2}{2} \Big).
\tag{17}
$$

(4)   The full conditional distributions of η_is are independent and their densities are

$$
p(\eta_i | \mu, \mu_0, \tau, \boldsymbol{x}) \propto g(\eta_i)\kappa(\eta_i)^{-1/2} \exp\Big\{ -\frac{\tau\kappa(\eta_i)^{-1}(x_i - \mu)^2}{2} \Big\}, \quad i = 1, \ldots, n,
\tag{18}
$$

where $g(\eta_i)$ is the density of a mixing variable.

The conditionals in Equation (15) through Equation (18) define the Gibbs sampler, provided that Equation (18) is a known density of standard form. Otherwise, we can define a Metropolis-within-Gibbs sampler that uses an M–H (Metropolis–Hastings) step to sample from the conditional in Equation (18). See [26], for the Metropolis-within-Gibbs sampler.

4.3. Markov Chain Monte Carlo Method

An MCMC method that works with the posterior Equation (14) is not complicated since Gibbs sampling of (μ, μ_0, τ) is routinely implemented based on each of their full conditionals outlined in Section 4.2. However, in posterior sampling of $\boldsymbol{\eta}$, an M–H sampling algorithm may be used when the

conditional posterior density Equation (18) does not have explicit form of known distribution. For the MCMC method, one should note the following points.

note 1: With given initial values of $(\mu, \mu_0, \tau, \boldsymbol{\eta})$, implementation of the Gibbs (or Metropolis-within-Gibbs) sampling algorithm consists of drawing repeatedly from distributions Equation (15) through Equation (18). The R package `tmvtnorm` and the R package `mvtnorm` can be used to sample from the conditionals and to calculate δ for a given α from Equation (11).

note 2: In cases using the LNFT model, η_i degenerates at $\kappa(\eta_i) = 1$. This means that the conditional distribution Equation (18) can be eliminated from the Gibbs sampler by setting $\kappa(\eta_i) = 1$ for the conditionals of μ and τ.

note 3: When the Lt_νFT model is used for reliability analysis, the last stage of the Bayesian hierarchical model in Equation (13) becomes $\eta_i \sim Gamma(\nu/2, \nu/2)$ with $\kappa(\eta_i) = \eta_i^{-1}$. Thus, the conditional distribution in Equation (18) yields

$$\left[\eta_i | \boldsymbol{\eta}_{\backslash \eta_i}, \mu, \mu_0, \tau\right] \sim Gamma\left(\frac{\nu+1}{2}, \frac{\nu + \tau \eta_i (x_i - \mu)^2}{2}\right),$$

where $\boldsymbol{\eta}_{\backslash \eta_i}$ denotes an $(n-1) \times 1$ vector whose elements are those of $\boldsymbol{\eta}$ except for η_i. Note that the LCFT model is a special case of the Lt_νFT model with $\nu = 1$. To allow ν to be determined within the model, one can specify one more prior stage for the Bayesian hierarchy in Equation (13). As suggested by [19,27], a uniform prior on $1/\nu$ ($0 < 1/\nu < 1$) can be considered. To limit model complexity, we consider only fixed ν so that the investigation of different Lt_νFT models is possible.

note 4: The conditional distribution Equation (18) for LSFT model is

$$\left[\eta_i | \boldsymbol{\eta}_{\backslash \eta_i}, \mu, \mu_0, \tau\right] \sim Gamma\left(1, \tau(x_i - \mu)^2/2\right) I(\eta_i \in (0,1)),$$

a truncated gamma whose support is $\eta_i \in (0, 1)$.

note 5: It is easily seen that the approximate conditional distribution η_i^2 for LLFT model is

$$\left[\eta_i^2 | \boldsymbol{\eta}_{\backslash \eta_i}, \mu, \mu_0, \tau\right] \sim IG\left(3, 1.233 + \tau(x_i - \mu)^2/8\right).$$

Then, a move to the proposal point η_i and posterior sampling of η_is can be made by using the M–H algorithm suggested by [19]. See, e.g., [28,29], a general procedure for M–H algorithm.

note 6: The convergence of an MCMC algorithm is an important issue for the correct estimation of the posterior distribution of interest. See [30] for an example of multiple convergence diagnosis and output analysis. When the Markov chain is converged, Rao–Blackwellization yields good estimates of μ and τ.

note 7: As measures of model comparison among SMLNFT models, a deviance information criterion (DIC) can be used. This measure can be calculated based on extensions of the MCMC method. See [31] and references therein for a review and comparisons of such extensions.

The Rao-Blackwellized Bayesian estimate of μ and $\tau = 1/\sigma^2$ from m post-convergence Markov chain values are given by the ergodic theorem as:

$$
\begin{aligned}
\hat{\mu} &= \frac{1}{m}\sum_{k=1}^{m} E[\mu | \mu_0^{(k)}, \tau^{(k)}, \boldsymbol{\eta}^{(k)}, \boldsymbol{x}] = \frac{1}{m}\sum_{k=1}^{m} \frac{\mu_0^{(k)}/\tau^{(k)} + \Delta^{(k)} \sum_{i=1}^{n} x_i/\kappa(\eta_i^{(k)})}{1/\tau^{(k)} + \Delta^{(k)} \sum_{i=1}^{n} 1/\kappa(\eta_i^{(k)})}, \\
\hat{\tau} &= \frac{1}{m}\sum_{k=1}^{m} E[\tau | \mu^{(k)}, \mu_0^{(k)}, \boldsymbol{\eta}^{(k)}, \boldsymbol{x}] = \frac{1}{m}\sum_{k=1}^{m} \frac{\nu_1 + n/2}{\nu_2 + \sum_{i=1}^{n}(x_i - \mu^{(k)})^2/(2\kappa(\eta_i^{(k)}))},
\end{aligned}
$$

and　　　$\hat{\sigma}^2 = \hat{\tau}^{-1}.$

As indicated by Equation (4) through Equation (7), it is evident that the reliability measures are functions of μ and σ^2 as denoted by $u(\mu, \sigma^2)$. Represent Bayesian estimate with

$$E[E[u(\mu, \sigma^2|\eta)]] = \int_0^\infty \left\{ \int_\Omega u(\mu, \sigma^2|\eta) p(\mu, \sigma^2|\boldsymbol{y}) d\mu d\sigma^2 \right\} g(\eta) d\eta,$$

where $p(\mu, \sigma^2|\boldsymbol{y})$ is joint posterior density of μ and σ^2 based on a fitted SMLNFT model, $g(\eta)$ is the density of η, and $\Omega = \{(\mu, \sigma^2); \mu \in \mathbf{R}, \sigma^2 > 0\}$ is the support of $p(\mu, \sigma^2|\boldsymbol{y})$. Next, a Bayesian Monte Carlo estimate of $u(\mu, \sigma^2)$ is given by

$$\hat{u}(\mu, \sigma^2) = \frac{1}{K} \sum_{k=1}^{K} \left(\frac{1}{m} \sum_{r=1}^{m} u(\mu^{(r)}, \sigma^{2(r)}|\eta^{(k)}) \right), \tag{19}$$

where $(\mu^{(r)}, \sigma^{2(r)})$, $r = 1, \ldots, m$, denote the rth Gibbs sample obtained from a fitted SMLNFT model, while $\eta^{(k)}$ is an independently generated kth sample value from the mixing distribution $\eta \sim G(\eta)$. When we have exact expression of $u(\mu, \sigma^2)$, then the Monte Calro integration, using generated $\eta^{(k)}$s, vanishes from Equation (19). For example, the estimate of posterior predictive reliability of the Lt_νFT model for specified y is

$$\hat{R}(y; \mu, \sigma^2) = 1 - \frac{1}{m} \sum_{r=1}^{m} F_\nu \left(\frac{log(y) - \mu^{(r)}}{\sigma^{(r)}} \right),$$

where $(\mu^{(r)}, \sigma^{2(r)})$, $r = 1, \ldots, K$, denote the rth Gibbs sample obtained from the fitted Lt_νFT model. Now, $(1 - \alpha) \times 100\%$ credible interval of the reliability at time y can be calculated by $[R_{1-\alpha/2}, R_{\alpha/2}]$, where R_q denotes the qth quantile value of $R^{(r)}$s with $R^{(r)} = 1 - F_\nu \left(\frac{log(y) - \mu^{(r)}}{\sigma^{(r)}} \right)$. When the LNFT model is assumed, Bayesian estimates of reliability measures can be simply obtained by setting $\kappa(\eta^{(k)}) = 1$ to Equation (18). That is, the reliability at time y is

$$\hat{R}(y; \mu, \sigma^2) = 1 - \frac{1}{m} \sum_{r=1}^{m} \Phi \left(\frac{log(y) - \mu^{(r)}}{\sigma^{(r)}} \right).$$

5. Numerical Illustrations

We provide two data applications to illustrate and demonstrate the performance of the proposed methodology. The first real data application demonstrates the performance of the proposed methodology (MCMC method based on a Bayesian hierarchical SMLNFT model) for inference in reliability with stochastic constraint. Proposed SMLNFT models were fitted to the second data set and compared in terms of their data-fitting performances, estimating parameters, and robustness to fat-tailed empirical distribution. For numerical implementations, we developed our program written in R, which is available from the author upon request.

5.1. Equipment Failure Data Example

This example considers failure time data (assessed in hours) obtained from a reliability study of assembly line equipment in operation in an automobile plant. The data, available in [32], consists of 43 observations of the failure time (Y). It provides summary statistics (mean and standard deviation (S.D.)) as listed in Table 1. We implemented the Shapiro–Wilks (S–W) and Kolmogorov–Smirnov (K–S) tests to determine the log-normality of Y (i.e., normality of $X = log(Y)$) based on the data. The test statistic values and corresponding p-values are also listed in Table 1.

Table 1. Summary statistics for the failure time data.

Variable	Mean	S.D.	S–W (*p*-Value)	K–S (*p*-Value)
Y	2201.488	2519.174	0.732 (<0.01)	0.231 (<0.01)
X	7.143	1.088	0.982 (0.760)	0.064 (>0.150)

As seen in the table, the formal tests do not reject the log-normality of Y, i.e., $Y \sim \mathcal{LN}(\mu, \sigma^2)$. Our objective of data analysis is to estimate reliability measures of the assembly line equipment subject to $Pr(\mu \in \mathcal{C}) = \alpha$ with $\mathcal{C} = (c, \infty)$, where c and α are known values. We can get the prior information about μ from past studies or an assembly line quality control report. This stochastic constraint on μ indicates that, with $\alpha \times 100\%$ degree of prior belief, the reliability of the LNFT model decreases slower than the empirical implication because $\bar{x} = 7.143 < c$. Following the proposal of this paper, the two-stage MaxEnt prior $\pi_{Maxent}(\mu)$ can be used to elicit the stochastic constraint on μ. Utilizing $\pi_{Maxent}(\mu)$, we can use the Bayesian hierarchical LNFT model (developed in Section 4) to carry out reliability analysis of the assembly line.

To see the adequacy of the Bayesian hierarchical LNFT model, we fitted the failure time data to five Bayesian hierarchical models (LNFT, LLFT, LCFT, Lt_5FT, and LSFT Bayesian hierarchical models) and compared them in terms of deviance information criterion (DIC), which is a measure of model comparison and adequacy. For each of the five models, we set up and ran a MCMC sampling algorithm (using Gibbs sampler) on the failure time data set and generated 100,000 Gibbs samples of model parameters from the conditional posteriors given in Section 4.2. For the MCMC algorithm, we used a burn-in period of 1000, a thin interval of 100, and the following choice of hyper-parameter values: $\theta_0 = 7$, $\sigma_0 = 1$, and $\nu_1 = \nu_2 = 0.001$. DIC values of the five models were calculated and they were 132.421, 135.704, 150.209, 135.698, and 145.971 for LNFT, LLFT, LCFT, Lt_5FT, and LSFT models, respectively. Thus, Bayesian hierarchical LNFT model achieves minimum DIC value among the five models. This, along with the formal log-normality test results in Table 1, provides support in favor of the LNFT model for fitting the data.

Table 2 shows some posterior summary statistics for the LNFT model. The statistics include posterior estimate, posterior standard deviation (S.D.), 0.025th quantile, 0.975th quantile, and MC error calculated from the Gibbs samples obtained by differing the values of c and δ. Mote Carlo (MC) errors in the table were obtained by setting $\delta = 0.9$. In estimating the Mote Carlo (MC) errors, we used the batch mean method with 50 batches. See, e.g., ([31], pp. 39–40) for the batch mean method. The degree of the prior belief (α) in the constraint ($\mu \in \mathcal{C}$) were calculated by use of Equation (11), and their values are also listed in Table 2.

Table 2. Posterior summaries for each model parameters and degree α of prior belief.

$\mathcal{C}$	Parameter	$\delta = 0$	$\delta = 0.5$	$\delta = 0.9$	$\delta = 1$	S.D.	2.5%	97.5%	MC Error
	μ	7.141	7.182	7.261	7.553	0.155	6.964	7.569	0.002
$(7.5, \infty)$	σ^2	1.181	1.185	1.196	1.327	0.295	0.812	1.930	0.004
	α	0.308	0.698	0.865	1.000	-	-	-	-
	μ	7.141	7.202	7.364	8.029	0.157	7.067	7.684	0.002
$(8.0, \infty)$	σ^2	1.181	1.188	1.228	1.941	0.313	0.823	2.036	0.004
	α	0.158	0.660	0.834	1.000	-	-	-	-
	μ	7.141	7.255	7.637	9.035	0.187	7.306	8.041	0.002
$(9.0, \infty)$	σ^2	1.181	1.198	1.408	4.734	0.401	0.901	2.450	0.005
	α	0.023	0.610	0.780	1.000	-	-	-	-

The small MC error values listed in Table 2 convince us of the convergence of the MCMC algorithm. The table shows that $\pi_{MaxEnt}(\mu)$ induces a shrinkage effect in the Bayesian estimation of μ with the uncertain interval constraint, i.e., $Pr(\mu \in \mathcal{C}) = \alpha$. This effect can be seen from the comparison of the

posterior estimates of μ obtained from the case of $\delta = 0$ with those from $\delta \in (0, 1]$. This comparison indicates the followings: (i) the Bayesian estimate of μ, based on the two-stage prior $\pi_{MaxEnt}(\mu)$, shrinks toward the interval $\mathcal{C} = (c, \infty)$. The magnitude of shrinkage effect induced by using the proposed prior becomes more evident as value of δ (or the degree α of belief in the interval constraint $\mu \in \mathcal{C}$) gets larger. Especially, in the case of $\mathcal{C} = (9, \infty)$ and $\delta = 0.9$, we see that the 95% credible interval for μ does not include $\hat{\mu}_{Bayes} = 7.141$ (estimate of μ with no constraint). This highlights the shrinkage effect induced by using $\pi_{MaxEnt}(\mu)$. (ii) When the priori constraint is certain, i.e., $\delta = \alpha = 1$, Table 2 shows that estimate of μ is located in the given constraint $\mathcal{C}$. This advocates the use of the Bayesian hierarchical LNFT model for reflecting the uncertain parameter constraint in reliability analysis.

Posterior predictive reliability measures of the LNFT model were also estimated by using the failure time data. Figure 3 depicts estimates of posterior predictive reliability and hazard rate functions for each case where the stochastic constraint is $\mu \in \mathcal{C} = (c, \infty)$, $c = 0.5, 0.7, 0.9$ for given $\delta = 0.9$. The figure also compares the functions with those obtained by setting $c = -\infty$, which denotes the case where there is no priori constraint on μ, i.e., the case of $\delta = 0$. Figure 3 clearly indicates that the two-stage MaxEnt prior appropriately reflects the priori stochastic constraint $\mu \in \mathcal{C} = (c, \infty)$ in estimating the posterior predictive predictive measures. The left panel (a) shows that reliability of the constrained LNFT model decreases slower than that of the unconstrained model (with $c = -\infty$), and this phenomenon is more evident as the constrained interval $\mathcal{C}$ locates far from the empirical mean value $\bar{x} = 7.143$. This, in turn, affects shape of hazard rate functions in the right panel (b). For $c > \bar{x}$, priori constrained LNFT models tend to produce lower hazard rates than the unconstrained LNFT model at all time points, but the pattern of their hazard rate functions is similar. We also see that the hazard rate becomes lower as the value of c gets larger. This coincides with the implication of the reliability functions in the panel (a).
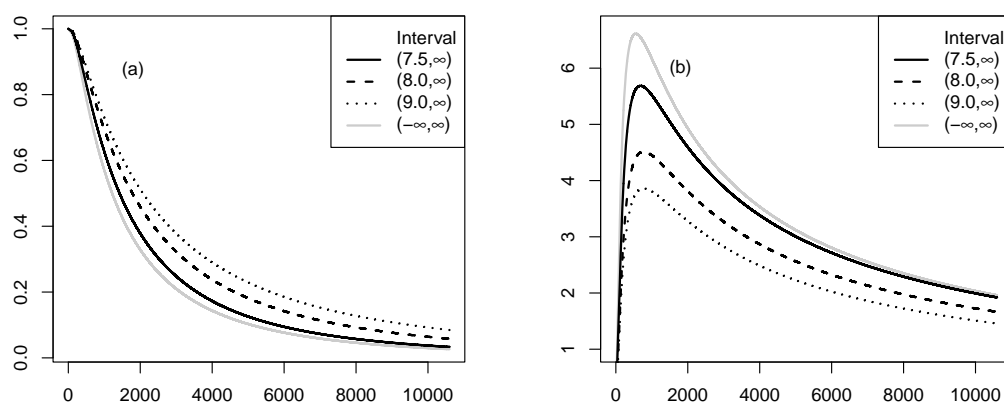

Figure 3. Estimated posterior predictive reliability measures for $\delta = 0.9$: the label for the y-axis of (**a**) is reliability and that of (**b**) is hazard rate. The label for the x-axis is failure time and the value for tick marks in the y-axis of (**b**) is multiplied by 10^4.

5.2. Artificial Data Example

A study using artificial data is done to evaluate the performance of proposed Bayesian hierarchical estimation methodology based on the class of SMLNFT models. For the study, we have considered a Lt_5FT model with location $\mu = 2$ and scale parameter $\sigma^2 = 0.25$. By use of the model, we simulated a sample of size 300 in complete failure times. Five models (LNFT, LCFT, Lt_5FT, LSFT, and LLFT models) were fitted to the same simulated data. Using the data, we ran the MCMC sampling algorithm based on each Bayesian hierarchical model with no constraint (i.e., $\delta = 0$) and generated 50,000 random samples from the conditional posteriors by using a burn-in period of 5000 and a thin interval of 10. For constructing each hierarchical model, we used the following choice of hyper-parameter values,

$\theta_0 = 0$, $\sigma_0 = 3$, and $\nu_1 = \nu_2 = 0.001$, which reflect the diffuseness of the priors (i.e., limiting non-informative priors) of μ and τ.

　　Table 3 provides posterior summaries from the posterior samples generated by using the MCMC algorithm developed for each of the five models. The small MC error values listed in Table 3 indicate that we have calculated the parameter estimates with precision and the MCMC algorithm has converged to its target distributions as well. DIC values of all models are also presented in the table. We observe that the Lt_5FT model has the lower DIC value, while the other models' results are considerably higher DIC values than the corresponding one under the Lt_5FT model. This demonstrates model selection function of the methodology. Concerning the posterior estimates, the estimate of σ^2 under the Lt_5FT model is accurate in comparison to that under the other models. To be more specific, the scale parameter σ^2 ranges from 0.025 to 0.365 with 95% posterior probability for the Lt_5FT model, while, for the other models, none of the credible intervals include true value $\sigma^2 = 0.25$ with the same posterior probability. On the other hand, Table 3 shows that the MCMC algorithm applied to each of the five models correctly estimates true location parameter value $\mu = 2$. This implies that the proposed Lt_5FT model is robust to fat tailed failure time observations affecting the scale parameter.

Table 3. Posterior summaries for parameters of five models.

Model	Parameter	Mean	S.D.	MC Error	2.5%	Median	97.5%	DIC
LNFT	μ	2.043	0.147	<0.001	1.755	2.043	2.332	861.602
	σ^2	4.329	0.438	0.002	3.554	4.300	5.274	-
LCFT	μ	1.948	0.043	<0.001	1.863	1.948	2.035	463.242
	σ^2	0.146	0.026	<0.001	0.101	0.146	0.204	-
Lt_5FT	μ	1.948	0.043	<0.001	1.862	1.948	2.033	394.339
	σ^2	0.289	0.036	<0.001	0.225	0.286	0.365	-
LSFT	μ	1.992	0.118	<0.001	1.761	1.992	2.227	867.325
	σ^2	0.557	0.098	<0.001	0.389	0.548	0.776	-
LLFT	μ	2.025	0.132	<0.001	1.767	2.025	2.282	842.847
	σ^2	1.363	0.176	<0.001	1.056	1.350	1.745	-

6. Conclusions

　　This paper has provided a methodology for Bayesian inference in reliability of SMLNFT models in cases where the interval constraint $\mu \in \mathcal{C}$ must be incorporated due to uncertainty. For failure time modeling with stochastic restriction, we proposed a Bayesian hierarchical model involving a two-stage MaxEnt prior distribution of μ based on Boltzmann's maximum entropy theorem. The two-stage MaxEnt prior reflects uncertainty about the prior constraint, that is, the Bayesian hierarchical SMLNFT models are subject to reliability restriction with uncertainty. Furthermore, we provide an MCMC method for assessing inference in the reliability of SMLNFT models based on their Bayesian hierarchical models. The effectiveness of our methodology is demonstrated by conducting two data applications.

　　We find that the proposed class of SMLNFT models is flexible enough to account for the behavior of reliability, unlike the log-normal FT model. We also find a connection between the degree of constraint presumption and the hyper-parameters of the two-stage MaxEnt prior, indicating adequacy of the MaxEnt prior in eliciting the priori uncertain constraint (see Corollary 1). Finally, we find that the proposed Bayesian hierarchical methodology enables the development of a simple MCMC algorithm to assess the posterior inference in reliability of the SMLNFT models with a stochastic constraint. Our proposed methodology can be easily extended to cases where the failure time data is incomplete, as a result, for example, of truncated or censored data, as well as to other FT models, including exponential, Weibull, and Gamma FT models. In addition, the class of SMLNFT models could accommodate multivariate FT models at higher dimension without much difficulty.

The methodological process proposed here in relation to Bayesian estimation can be extended to the multivariate models with a multivariate stochastic constraint.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Stefanescu, C.; Turnbull, B.W. Multivariate frailty models for exchangeable survival data with covariates. *Technometrics* **2006**, *48*, 411–417.
2. Kavam, P.H.; Pena, E.A. Estimating load-sharing properties in a dynamic reliability system. *J. Am. Stat. Assoc.* **2005**, *100*, 262–272.
3. Sun, J. *The Statistical Analysis of Interval-Censored Failure Time Data*; Spring: New York, NY, USA, 2002.
4. Yue, S. The bivariate lognormal distribution for describing joint statistical properties of multivariate storm event. *Environmetrics* **2002**, *13*, 811–819.
5. Yerel, S.; Konuk, A. Bivariate lognormal distribution model of cutoff grade impurities: A case study of magnesitr ore deposit. *Sci. Res. Essay* **2009**, *4*, 1500–1504.
6. Lee, C.-F.; Finnerty, J.; Lee, J.; Lee, A.C.; Wort, D. *Security Analysis, Portpolio Management, and Financial Derivatives*; World Scientific Publishing Company: Singapore, 2012.
7. Halliwell, L.J. The lognormal random multivariate. In *Casualty Actuarial Society E-Forum*; Spring, 2015. Avaiable online: www.casact.org/pubs/forrum/15spforum/Halliwell.pdf (accessed on 13 June 2017).
8. Weinke, A. *Frailty Models in Survival Analysis*; Chapman and Hall: Baco Raton, FL, USA, 2010.
9. De Alba, E. Claims reserving when there iare negative values in the runoff triangle: Bayesian analysis using the three-parameter log-normal distribution. *N. Am. Axtuarial J.* **2006**, *10*, 28–38.
10. Yue, S. The bivariate lognormal distribution to model a multivariate flood episode. *Hydrol. Process.* **2000**, *14*, 2575–2588.
11. Elshqeirat, B.; Soh, S.; Rai, S.; Lazarescu, M. Dynamic programming for minimal cost topology with reliability constraint. *J. Adv. Comput. Netw.* **2013**, *1*, 286–290.
12. Lin, Y.-K. Reliability evaluation for an information network with node failure under cost constraint. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2007**, *37*, 180–188.
13. Koide, T.; Shinmori, S.; Ishii, H. Topological optimization with a network reliability constraint. *Discret. Appl. Math.* **2001**, *115*, 135–149.
14. Cercignani, C. *The Boltzman Equation and Its Applications*; Springer: Berlin, Germany, 1988.
15. Leonard, T.; Hsu, J.S.J. *Bayesian Methods: An Analysis for Statisticians and Interdisciplinary Researchers*; Cambridge University Press: Cambridge, UK, 1999.
16. O'Hagan, A.; Leonard, T. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **1976**, *63*, 201–203.
17. Lachos, V.H.; Labra, F.V.; Bolfarine, H.; Ghosh, P. Multivariate measurement error models based on scale mixtures of the skew-normal distribution. *Statistics* **2010**, *44*, 541–556.
18. Lindsey, J.K. *Statistical Analysis of Stochastic Process in Time*; Cambridge University Press: Cambridge, UK, 2004.
19. Chen, M.H.; Dey, D.K. Bayesian modeling of correlated binary responses via scale mixture of multivariate normal link functions. *Sankhyá Indian J. Stat.* **1998**, *60*, 322–343.
20. Gago-Benitez, A.; Fernádo-Madrigal, J.-A.; Cruz-Martin, A. Log-logistic modeling of sensory flow delays in networked telerobots. *IEEE Sens. J.* **2013**, *13*, 2294–2953.
21. Kim, H.J. A best linear threshold classification with scale mixture of skew normal populations. *Comput. Stat.* **2015**, *30*, 1–28.
22. Steiger, J. When constraints interact: A caution about reference variables, identification constraints, and scale dependencies in structural equation modeling. *Psychol. Methods* **2002**, *7*, 210–227.
23. Lopes, H.F.; West, M. Bayesian model assessment in factor analysis. *Stat. Sin.* **2004**, *14*, 41–67.
24. Loken, E. Identification constraints and inference in factor models. *Struct. Equ. Model.* **2005**, *12*, 232–244.

25. Gupta, S.D. A note on some inequalities for multivariate distribution. *Bull. Calcutta Stat. Assoc.* **1969**, *18*, 179–180.
26. Kim, H.J. A two-stage maximum entropy prior of location parameter with a stochastic multivariate interval constraint and its properties. *Entropy* **2016**, *18*, 1–20.
27. Sisson, S. Trans-dimensional Markov chains: A decade of progress and future perspectives. *J. Am. Stat. Assoc.* **2005**, *100*, 1077–1089.
28. Koop, G; Poirier, D.J.; Tobias, J.L. *Bayesian Econometric Methods*; Cambridge University Press: Cambridge, UK, 2007.
29. Chib, S.; Greenberg, E. Understanding the Metropolis-Hastings algorithm. *Am. Stat.* **1995**, *49*, 327–335.
30. Heidelberger, P.; Welch, P. Simulation run length control in the presence of an initial transient. *Oper. Res.* **1992**, *31*, 1109–1144.
31. Ntzoufras, I. *Bayesian Modeling Using WinBUGS*; Wiley: New York, NY, USA, 2009.
32. Dagpunar, J.S. *Simulation and Monte Carlo*; Wiley: New York, NY, USA, 2007.