stats

MDPI

Article

# Statistical Inference for Progressive Stress Accelerated Life Testing with Birnbaum-Saunders Distribution

**Naijun Sha** †

Department of Mathematical Sciences, University of Texas at El Paso, El Paso, TX 79968, USA; nsha@utep.edu; Tel.: +1-915-747-6844

† Current address: 500 W University Ave., El Paso, TX 79968, USA.

Received: 11 November 2018; Accepted: 10 December 2018; Published: 12 December 2018

**Abstract:** As a result of the two-parameter Birnbaum–Saunders (BS) distribution being successful in modelling fatigue failure times, several extensions of this model have been explored from different aspects. In this article, we consider a progressive stress accelerated life testing for the BS model to introduce a generalized Birnbaum–Saunders (we call it Type-II GBS) distribution on the lifetime of products in the test. We outline some interesting properties of this highly flexible distribution, present the Fisher's information in the maximum likelihood estimation method, and propose a new Bayesian approach for inference. Simulation studies are carried out to assess the performance of the methods under various settings of parameter values and sample sizes. Real data are analyzed for illustrative purposes to demonstrate the efficiency and accuracy of the proposed Bayesian method over the likelihood-based procedure.

**Keywords:** progressive accelerated life testing; generalized Birnbaum–Saunders distribution; Bayesian inference; Fisher's information; MCMC sampling

## 1. Introduction

The Birnbaum–Saunders (BS) model is based on a physical argument of cumulative damage that produces fatigue in materials, which is exerted by cyclical stress. The failure follows from the development and growth of a dominant crack in the material. Considering the basic characteristics of the fatigue process, Ref. [1] derived the BS$(\alpha, \beta)$ distribution function of the failure time $T$ expressed by:

$$F(t) = \Phi\left(\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{1/2} - \left(\frac{\beta}{t}\right)^{1/2}\right]\right), \; t > 0, \tag{1}$$

where $\alpha > 0, \beta > 0$ are shape and scale parameters and $\Phi(\cdot)$ is the distribution function of the standard normal variate. The BS distribution can be widely applied to describe fatigue life and lifetimes in general. Its field of application has been extended beyond the original context of material fatigue, and the model becomes the most versatile within the popular distributions for failure times. Over the years, various approaches of parameter inference, generalizations, and applications of the distribution have been introduced and developed by many authors (see [2–8], just to name a few).

In recent years, more research work has lied in the study of accelerated life testing (ALT) with the BS distribution. In industrial experiments, it is often very costly and time consuming to obtain information about the lifetime of highly reliable products under normal experimental conditions. To collect failure data rapidly and improve the efficiency of the experiment, people commonly apply ALT where products or materials are subjected to elevated stress conditions compared to those normally applied in practice. These stresses often include temperature, voltage, vibration, cycling rate, pressure, etc., which can be applied in mainly three ways: constant stress, step stress, and progressive stress.

Many authors have contributed extensively to the development and parameter inference on ALT for various lifetime distributions; see, for example, [9–13]. An ALT on the BS distribution was considered in [14], who developed the model under the inverse power law accelerated form and explored the inference procedure based on the lifetime data collected under several elevated stress levels. The work in [15] presented an inference approach for the BS distribution in step-stress accelerated life testing (SSALT) with the Type-II censoring scheme. In practice, to make a more efficient test plan for collecting the lifetime of highly reliable products, people usually resort to an ALT by a progressive stress, and we will consider such a model for the BS distribution.

## 2. Progressive Stress Accelerated Life Testing

We briefly introduce the SSALT first and then extend it to the progressive stress model for the BS distribution.

### *2.1. Step-Stress Test*

Suppose that, for a particular pattern of SSALT with $I$ steps in total, step $i$ runs at stress level $V_i$, starts at time $t_{i-1}$, and runs to time $t_i (t_0 = 0), i = 1, 2, ..., I$. The life distribution for specimens at the stress level $V_i$ is assumed as BS$(\alpha, \beta_i)$, whose distribution function $F_i(t)$ in (1), that is, $F_i(t) = \Phi([(u_i(t))^{1/2} - (u_i(t))^{-1/2}]/\alpha)$ with $u_i(t) = t/\beta_i$, where $\alpha$ is the common shape parameter, and the stress $V_i$ only influences the scale parameter $\beta_i$ through the inverse power law, i.e., $\beta_i = (V_0/V_i)^p$ with a standard level of stress $V_0 > 0$ and the power $p > 0$ being only associated with the characteristics of products. Accordingly, based on the principle of cumulative exposure (CE) [16], the population cumulative fraction of specimens failing in step $i$ is:

$$F(t) = F_i(t - t_{i-1} + s_{i-1}),\ t_{i-1} \le t \le t_i, i = 1, 2, ..., I \tag{2}$$

where the equivalent start time $s_{i-1}$ at step $i$ satisfies:

$$F_i(s_{i-1}) = F_{i-1}(t_{i-1} - t_{i-2} + s_{i-2}). \tag{3}$$

It brings about $[u_i(s_{i-1})]^{1/2} - [u_i(s_{i-1})]^{-1/2} = [u_{i-1}(t_{i-1} - t_{i-2} + s_{i-2})]^{1/2} - [u_{i-1}(t_{i-1} - t_{i-2} + s_{i-2})]^{-1/2}$, leading to the following recursive expression from $u_i(s_{i-1}) = u_{i-1}(t_{i-1} - t_{i-2} + s_{i-2})$ by the fact that $g(u) = u^{1/2} - u^{-1/2}$ is an increasing function in $u > 0$:

$$\frac{s_{i-1}}{\beta_i} = \frac{t_{i-1} - t_{i-2}}{\beta_{i-1}} + \frac{s_{i-2}}{\beta_{i-1}}. \tag{4}$$

By using the recursive equation above with the time duration being $\Delta_i = t_i - t_{i-1}$ at the $i^{\text{th}}$ step, the population fraction having failed in (2) over the exposure time $t_I = \Delta_1 + \Delta_2 + \cdots + \Delta_I$ after $I$ steps is:

$$F(t_I) = F_i(t_I - t_{I-1} + s_{I-1}) = \Phi\left(\frac{1}{\alpha}\left[(u(t_I))^{1/2} - (u(t_I))^{-1/2}\right]\right), \tag{5}$$

where the "cumulative exposure" $u(t_I)$ is the sum of the "scaled" time experienced at each step given by:

$$u(t_I) = (\Delta_1/\beta_1) + (\Delta_2/\beta_2) + \cdots + (\Delta_I/\beta_I). \tag{6}$$

As all time lengths $\Delta_i \to 0$, the step-stress levels $V_i$ become a progressive or varying stress $V(t)$ continuously over time in ALT, and as a result, the scale parameter $\beta(V) = (V_0/V)^p$ is a function of time, namely $\beta(t)$, and the corresponding cumulative exposure $u(t)$ (or damage), which appears as a sum in (6) for a step-stress testing, becomes the integral $u(t) = \int_0^t 1/\beta(x)dx$. Specifically, here, we consider a ramp stress $V(t) = Rt$ with $R > 0$ being the rate of rise of stress, and then, the scale

parameter becomes $\beta(t) = [V_0/V(t)]^p = V_0^p/(R^p t^p)$. The corresponding cumulative exposure time of products at time $t$ is:

$$u(t) = \int_0^t \frac{dx}{\beta(x)} = \int_0^t \frac{R^p x^p}{V_0^p} dx = \frac{R^p t^{p+1}}{(p+1)V_0^p}, \tag{7}$$

and the lifetime distribution of products at the progressive stress $V(t)$ becomes:

$$F(t) = \Phi\left(\frac{1}{\alpha}\left[(u(t))^{1/2} - (u(t))^{-1/2}\right]\right) = \Phi\left(\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^m - \left(\frac{\beta}{t}\right)^m\right]\right), \tag{8}$$

where:

$$m = (p+1)/2,\ \beta = [(p+1)V_0^p/R^p]^{1/(p+1)}. \tag{9}$$

The distribution in (8) is a generalization of the Birnbaum–Saunders distribution (GBS) first proposed in [17] by allowing the exponent (1/2 on the BS) to take on other values. Recently, considering the original BS distribution derived from a homogeneous Poisson process, [18] obtained the same GBS depending on a non-homogeneous Poisson process. To distinguish from the GBS developed in [19], we refer to the distribution as Type-II GBS, denoted as GBS-II$(m, \alpha, \beta)$, whose density function is given by:

$$f(t) = \frac{m}{\alpha t}\left[\left(\frac{t}{\beta}\right)^m + \left(\frac{\beta}{t}\right)^m\right]\phi\left(\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^m - \left(\frac{\beta}{t}\right)^m\right]\right),\ t > 0, \tag{10}$$

where $m > 0$ is another shape-type parameter and $\phi(\cdot)$ is the density function of the standard normal distribution. As the BS distribution, the transformation $Z = ((T/\beta)^m - (\beta/T)^m)/\alpha$ is a standard normal variate, and thus, it is easy to generate the GBS-II random variates from the standard normal variables. In the following, we briefly summarize the main properties of the distribution.

### 2.2. Properties of GBS-II

The three-parameter GBS-II distribution in (10) is a flexible family of distributions, and the shape of the density widely varies with different values of the parameters. Specifically (the detailed proof is long, and so omitted here): (i) when $0 < m \leq 1$, the density is unimodal or an upside-down bathtub; (ii) when $m > 1$ and $\alpha^2 \leq m/(m-1)$, the density is also unimodal (upside-down); (iii) if $m > 1$ and $\alpha^2 > m/(m-1)$, then the density is either unimodal or bimodal. Figure 1 shows various graphs of the density function for different values of $m$ and $\alpha$ with the scale parameter $\beta$ fixed at unity. Additionally, Figure 2 presents the failure rate function given by $\lambda(t) = f(t)/(1 - F(t))$ with various values of $m$ and $\alpha$, showing that it could have an increasing, decreasing, bathtub, and upside-down bathtub shape. Hence, it seems that the distribution is flexible enough to model various situations of product life.

The GBS-II has some interesting properties as the BS distribution. For example, $\beta$ remains the median for the distribution. The reciprocal property is also preserved by that if $T \sim$ GBS-II$(m, \alpha, \beta)$, then $T^{-1} \sim$ GBS-II$(m, \alpha, \beta^{-1})$. In fact, the GBS-II generally describes the distribution family of power transformations for a BS random variable from the following: if $T \sim$ BS$(\alpha, \beta)$, then for any nonzero real-valued constant $r$, $T^r \sim$ GBS-II$(0.5|r|^{-1}, \alpha, \beta^r)$, where $|\cdot|$ represents the absolute value function. Conversely, given $T \sim$ GBS-II$(m, \alpha, \beta)$, then $T^{2m} \sim$ BS$(\alpha, \beta^{2m})$. In addition, similar to the BS distribution, which can be written as an equal mixture of an inverse normal distribution and a distribution of the reciprocal of an inverse normal random variable [20], the GBS-II distribution can be also expressed as a mixture of power inverse normal-type distributions [21].

In the aspect of numeric characteristics for GBS-II, generally, there is no analytic form of moments except for some special cases. For example, from the fact that $T^m = \beta^m \left(\alpha Z + \sqrt{\alpha^2 Z^2 + 4}\right)/2$ with a standard normal variate $Z$, one may easily obtain the expressions of the moments $\mathrm{E}(T^{km})$ with an even number $k$, such as $\mathrm{E}(T^{2m}) = \beta^{2m}(1 + \alpha^2/2)$, $\mathrm{E}(T^{4m}) = \beta^{4m}(1 + 2\alpha^2 + 11\alpha^4/4)$, etc. The general moment expression was obtained in [22], who used the relationship between the GBS-II and three-parameter sinh-normal distribution (see [23]):

$$\mathrm{E}(T^r) = \frac{\beta^r \exp(\alpha^{-2})}{\alpha\sqrt{2\pi}} \left[K_{(r/m+1)/2}(\alpha^{-2}) + K_{(r/m-1)/2}(\alpha^{-2})\right], \tag{11}$$

where $r$ is a real number and $K_\omega(x)$ is an order $\omega$ modified Bessel function of the third kind, which can be expressed in an integral form as $K_\omega(x) = \int_0^\infty \exp\{-x\cosh(s)\}\cosh(\omega s)ds$ with the hyperbolic cosine function $\cosh(x) = [\exp(x) + \exp(-x)]/2$. Numerous software packages can be used to evaluate $K_\omega(x)$ for specific values of $\omega$ and $x$ for calculating the moments such as the mean and variance. The finite moments guarantee the rationality of the moment-based estimation methods.

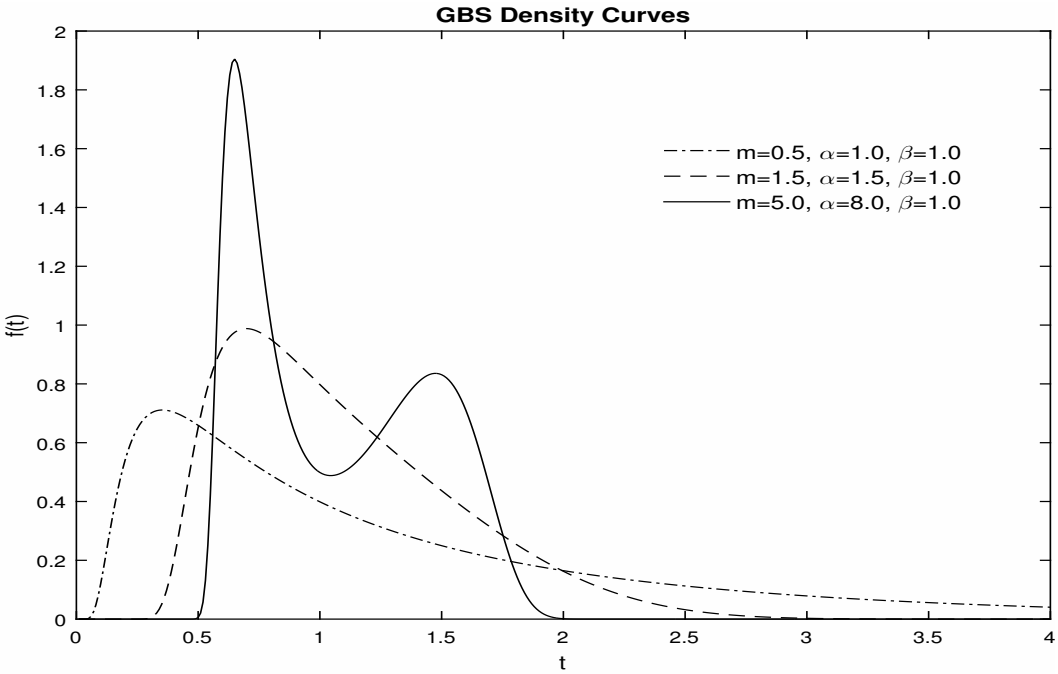


**Figure 1.** The generalized Birnbaum–Saunder-II$(m, \alpha, \beta)$ density curves for various parameter values.

So far, little research has been seen to study the parameter estimation and inference for the GBS-II distribution. The work in [21] discussed the maximum likelihood (ML) estimation of parameters and provided interval estimation based on an "observed" Fisher's information. To contribute a novel research of precise inference, in this article, we further investigate the ML method to obtain an analytic expression of Fisher's information and propose a new Bayesian approach for parameter inference. The rest of the article is arranged as follows. Section 3 presents the methodology of the estimation procedure. Subsequently, we carry out simulation studies to investigate the performance of proposed methods in Section 4. For illustrative purpose, two real datasets are analyzed in Section 5, followed by some concluding remarks in Section 6.

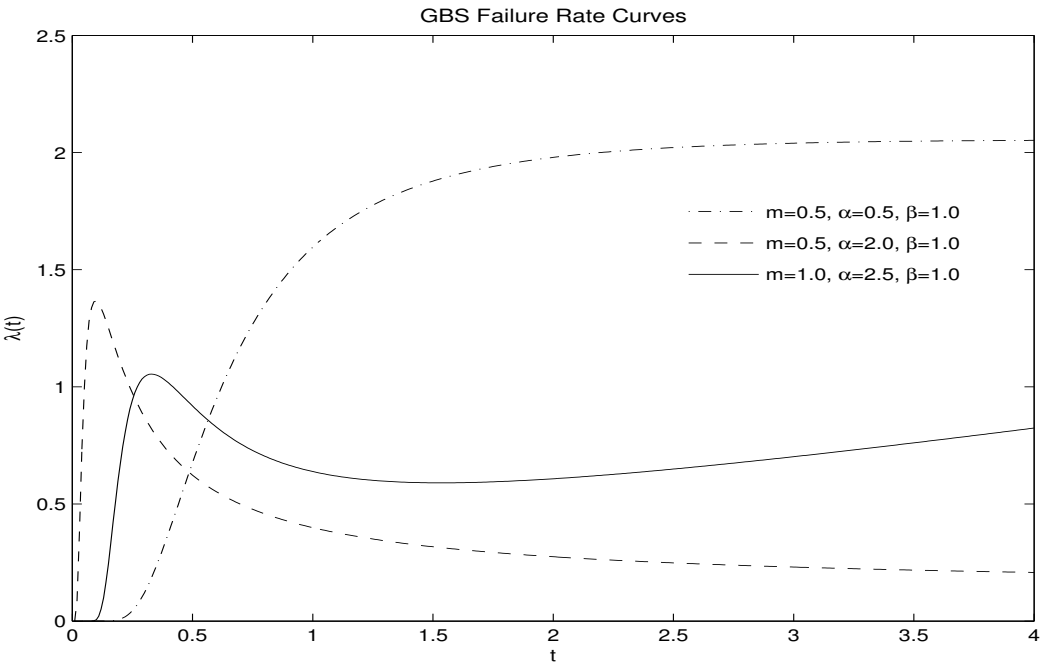


**Figure 2.** GBS-II$(m, \alpha, \beta)$ failure rate curves for various parameter values.

## 3. Estimations

Let $\boldsymbol{t} = (t_1, t_2, ..., t_n)$ be $n$ observational data from the GBS-II distribution. We first consider the likelihood-based approach in the following.

### *3.1. Likelihood-Based Method*

To make the notation simple, let $\varepsilon(t) = (t/\beta)^m - (\beta/t)^m, \delta(t) = (t/\beta)^m + (\beta/t)^m$. Then, the likelihood and log-likelihood functions, up to a constant, are given by:

$$L(m, \alpha, \beta|\boldsymbol{t}) = m^n \alpha^{-n} \prod_{i=1}^{n} (\delta(t_i)) \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^{n} \varepsilon^2(t_i)\right\}, \tag{12}$$

$$\ell = \ell(m, \alpha, \beta|\boldsymbol{t}) = n\log(m) - n\log(\alpha) + \sum_{i=1}^{n} \log(\delta(t_i)) - \frac{1}{2\alpha^2} \sum_{i=1}^{n} \varepsilon^2(t_i), \tag{13}$$

and we have the first partial derivatives w.r.t the parameters:

$$\frac{\partial \ell}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^3} \sum_{i=1}^{n} \varepsilon^2(t_i), \ \frac{\partial \ell}{\partial \beta} = -\frac{m}{\beta}\left[\sum_{i=1}^{n} \frac{\varepsilon(t_i)}{\delta(t_i)} - \frac{1}{\alpha^2} \sum_{i=1}^{n} \varepsilon(t_i)\delta(t_i)\right], \tag{14}$$

$$\frac{\partial \ell}{\partial m} = \frac{n}{m} + \sum_{i=1}^{n} \log(\beta^{-1} t_i) \frac{\varepsilon(t_i)}{\delta(t_i)} - \frac{1}{\alpha^2} \sum_{i=1}^{n} \log(\beta^{-1} t_i) \varepsilon(t_i)\delta(t_i). \tag{15}$$

Due to the complexity of the expression above, a numerical method has to be applied to obtain maximum likelihood estimates (MLEs) $\hat{m}, \hat{\alpha}, \hat{\beta}$ by solving the equations $\partial\ell/\partial m = 0, \partial\ell/\partial\alpha = 0$ and $\partial\ell/\partial\beta = 0$ simultaneously. The work in [21] used the "observed" Fisher's information matrix to obtain a large sample-based interval estimation for the parameters. However, our study shows that a "partially" analytic form of it can be obtained, so that more precise inference can be made. Let the Fisher's information matrix be the following form:

$$I(m,\alpha,\beta) = \begin{pmatrix} u_{mm} & u_{m\alpha} & u_{m\beta} \\ u_{\alpha m} & u_{\alpha\alpha} & u_{\alpha\beta} \\ u_{\beta m} & u_{\beta\alpha} & u_{\beta\beta} \end{pmatrix}, \tag{16}$$

whose elements are the negative expectation of the second partial derivatives of parameters for the log-likelihood function in (13), such as $u_{\alpha\alpha} = -\mathrm{E}(\partial^2\ell/\partial\alpha^2), u_{\alpha\beta} = -\mathrm{E}(\partial^2\ell/\partial\alpha\partial\beta)$, etc. These elements are given by (the detailed derivations are provided in Appendix A):

$$u_{\alpha\alpha} = \frac{2n}{\alpha^2},\ u_{\alpha\beta} = u_{\beta\alpha} = u_{\beta m} = u_{m\beta} = 0, \tag{17}$$

$$u_{\alpha m} = u_{m\alpha} = -\frac{2n}{\alpha^2 m}\mathrm{E}\left[Zg(Z)\sqrt{\alpha^2 Z^2+4}\right], \tag{18}$$

$$u_{\beta\beta} = \frac{2nm^2}{\alpha^2\beta^2}\left[\alpha^2 - 2\alpha h(\alpha) + 2\right], \tag{19}$$

$$u_{mm} = \frac{n}{m^2} - \frac{4n}{m^2}\mathrm{E}\left(\frac{g^2(Z)}{\alpha^2 Z^2+4}\right) + \frac{2n}{m^2\alpha^2}\mathrm{E}\left[g^2(Z)(\alpha^2 Z^2+2)\right], \tag{20}$$

where $Z$ is the standard normal variate, $g(z) = \log[(\alpha z + \sqrt{\alpha^2 z^2+4})/2]$ and $h(\alpha) = \sqrt{\pi/2} \times e^{2/\alpha^2}[1 - \Phi(2/\alpha)]$. Therefore, the asymptotic normal distribution of MLEs has the following covariance matrix:

$$I^{-1}(m,\alpha,\beta) = \begin{pmatrix} u_{\alpha\alpha}/c & -u_{m\alpha}/c & 0 \\ -u_{\alpha m}/c & u_{mm}/c & 0 \\ 0 & 0 & 1/u_{\beta\beta} \end{pmatrix}, \tag{21}$$

with $c = u_{mm}u_{\alpha\alpha} - u_{m\alpha}^2$. Notice from (21) that the MLEs $\hat{m}$ and $\hat{\beta}, \hat{\alpha}$, and $\hat{\beta}$ are asymptotically independent. It is also clear that when $m = 1/2$, the GBS-II reduces to the BS distribution whose Fisher's information is the lower $2 \times 2$ matrix in (16) with $u_{\beta\beta} = n\left[\alpha^2/2 - \alpha h(\alpha) + 1\right]/(\alpha\beta)^2$, the same as the one provided in [4].

By the fact of positive parameters and the asymptotic normality of MLEs, we may use log transformation to obtain approximate confidence intervals (CIs) for the parameters [24]. In particular, for parameter $m$ and its MLE $\hat{m}$, we have the approximate normal distribution $\log(\hat{m}) \sim \mathrm{N}(\log(m), \mathrm{Var}(\log(\hat{m})))$, where the variance can be approximated by the delta method as $\widehat{\mathrm{Var}(\log(\hat{m}))} = \widehat{\mathrm{Var}(\hat{m})}/\hat{m}^2 = \hat{u}_{\alpha\alpha}/(\hat{c}\hat{m}^2)$ with $\hat{u}_{\alpha\alpha}, \hat{c}$ being the values of $u_{\alpha\alpha}$ and $c$, respectively, evaluated at MLEs $\hat{m}, \hat{\alpha}, \hat{\beta}$. Then, a $(1-\gamma)100\%$ CI for $m$ is then given by:

$$\left[\hat{m} \times \exp\left\{-\frac{z_{\gamma/2}\sqrt{\mathrm{Var}(\hat{m})}}{\hat{m}}\right\},\ \hat{m} \times \exp\left\{\frac{z_{\gamma/2}\sqrt{\mathrm{Var}(\hat{m})}}{\hat{m}}\right\}\right] \tag{22}$$

where $z_{\gamma/2}$ is the upper $100 \times \gamma/2^{\text{th}}$ percentile of the standard normal distribution. The normal-approximated CIs for $\alpha$ and $\beta$ can be constructed in the same way to have the form in (22) with $\hat{m}$ being replaced by $\hat{\alpha}$ and $\hat{\beta}$, respectively, and $\widehat{\mathrm{Var}(\hat{\alpha})} = \hat{u}_{mm}/\hat{c}, \widehat{\mathrm{Var}(\hat{\beta})} = 1/\hat{u}_{\beta\beta}$. It should be noted that these intervals can exhibit less accurate coverage for small samples.

### 3.2. Bayesian Inference

Since there are no tractable forms of MLEs and CIs, this is not very applicable and efficient by the use of the ML estimation method. We propose a Bayesian inference approach as an alternative. From the model development of the GBS-II distribution in [18], we notice that the parameter $m$ is independent of the other two parameters $\alpha$ and $\beta$, which have a relation $\alpha^2 \propto \beta$, and so, we propose a joint prior $\pi(m,\alpha,\beta) = \pi(m)\pi(\alpha|\beta)\pi(\beta)$ with the conditional prior mean $E(\alpha^2|\beta) \propto \beta$. From the

likelihood function in (12), it is easily discernible that for $\alpha^2$, an inverse gamma is a conjugate prior for the conditional likelihood $L(\alpha|m,\beta,\boldsymbol{t})$. Therefore, we specify that $\alpha^2|\beta \sim \mathrm{IG}(a_0/2, a_1\beta/2)$, and thus, the prior density of $\alpha|\beta$ is:

$$\pi(\alpha|\beta) \propto \alpha^{-(a_0+1)} \exp\left\{-\frac{a_1\beta}{2\alpha^2}\right\} \tag{23}$$

with the hyperparameters $a_0 > 4$ and $a_1 > 0$ to ensure the existence of the variance. It is also clear that there is no conjugate prior for $m$ and $\beta$. However, we may consider a prior distribution that has a similar functional form as their conditional likelihood functions. In this case, we pick both priors of $m$ and $\beta$ to be a gamma distribution,

$$m \sim \mathrm{Gamma}\left(\frac{d_0}{2}, \frac{d_1}{2}\right),\ \beta \sim \mathrm{Gamma}\left(\frac{b_0}{2}, \frac{b_1}{2}\right), \tag{24}$$

with $b_0, b_1, d_0, d_1 > 0$. The hyperparameter values can be specified based on the following factors: (i) Since $\beta$ is the median of the GBS-II distribution, we can refer to the sample median as the estimate $\tilde{\beta} = \mathrm{median}(t_1, t_2, ..., t_n)$ in our attempt to specify $b_0$ and $b_1$. Further, by the fact that $E(Z) = 0, E(Z^2) = 1$ from $Z = [(T/\beta)^m - (\beta/T)^m]/\alpha \sim N(0,1)$ and by the principle of method of moments, we establish two equations by equating the first two sample and theoretical moments ($\beta$ is estimated by $\tilde{\beta}$), $\sum_{i=1}^n [(t_i/\tilde{\beta})^m - (\tilde{\beta}/t_i)^m] = 0$ and $\sum_{i=1}^n [(t_i/\tilde{\beta})^m - (\tilde{\beta}/t_i)^m]^2 = n\alpha^2$, to obtain the roots $\tilde{m}$ and $\tilde{\alpha}$ for specifying $a_0, a_1, d_0$ and $d_1$. (ii) The MLEs $\hat{m}, \hat{\alpha}, \hat{\beta}$ can also be used to determine the hyperparameters. The non-informative distribution with $a_i = 0, b_i = 0, d_i = 0, i = 0, 1$ can be chosen if no prior knowledge is available. The joint posterior distribution of the parameters $(m, \alpha, \beta)$ with the sample data $\boldsymbol{t}$ is given by:

$$\pi(m,\alpha,\beta|\boldsymbol{t}) \propto L(m,\alpha,\beta|\boldsymbol{t}) \times \pi(m)\pi(\alpha|\beta)\pi(\beta). \tag{25}$$

It follows that the full conditional posteriors are:

$$\pi(m|\alpha,\beta,\boldsymbol{t}) \quad \propto m^{n+\frac{d_0}{2}-1} \prod_{i=1}^n (\delta(t_i)) \exp\left\{-\frac{1}{2}\left(d_1 m + \frac{\sum_{i=1}^n \varepsilon^2(t_i)}{\alpha^2}\right)\right\}, \tag{26}$$

$$\alpha^2|(m,\beta,\boldsymbol{t}) \quad \sim \mathrm{IG}\left(\tfrac{\nu_0}{2}, \tfrac{\nu_1}{2}\right),\ \nu_0 = n + a_0,\ \nu_1 = a_1\beta + \textstyle\sum_{i=1}^n \varepsilon^2(t_i), \tag{27}$$

$$\pi(\beta|m,\alpha,\boldsymbol{t}) \quad \propto \beta^{\frac{b_0}{2}-1} \prod_{i=1}^n (\delta(t_i)) \exp\left\{-\frac{1}{2}\left(b_1\beta + \frac{\nu_1}{\alpha^2}\right)\right\}. \tag{28}$$

We implement a Markov chain Monte Carlo (MCMC) algorithm, specifically here a Gibbs sampling procedure (see, for example, Casella and George [25]), to draw posterior samples from their full conditional posterior distributions. First, we take MLEs of $m, \alpha$ and $\beta$ as the initial values to make the algorithm converge more quickly and then repeat the following steps $M$ times, among which, given the values at the $k^{\text{th}}$ iteration, the $(k+1)^{\text{th}}$ iteration is as follows:

i. Draw $m_{k+1}$ from $\pi(m|\alpha_k, \beta_k, \boldsymbol{t})$ in (26) using a Metropolis–Hastings (MH) procedure (see, for example, Chib and Greenberg [26]). We first propose $m_p \sim \mathrm{Gamma}(c_m m_k, c_m)$, where the mean of the proposal is $m_k$ and $c_m$ is a tuning parameter to make the algorithm efficient, and then take $m_{k+1} = m_p$ with probability:

$$\lambda_m = \min\left\{1, \frac{\pi(m_p|\alpha_k,\beta_k,\boldsymbol{t}) \times q_m(m_k|m_p)}{\pi(m_k|\alpha_k,\beta_k,\boldsymbol{t}) \times q_m(m_p|m_k)}\right\}, \tag{29}$$

with the Gamma proposal density function $q_m(\cdot)$, and so:

$$\frac{q_m(m_k|m_p)}{q_m(m_p|m_k)} = \frac{\Gamma(c_m m_k) c_m^{c_m m_k}}{\Gamma(c_m m_p) c_m^{c_m m_p}} \times \frac{m_k^{c_m m_p - 1} \exp\{-c_m m_k\}}{m_p^{c_m m_k - 1} \exp\{-c_m m_p\}}. \tag{30}$$

ii. Draw $\alpha_{k+1}^2 \sim \text{IG}\,(\nu_0/2, \nu_1/2)$ with updated $\nu_0$ and $\nu_1$ in (27).

iii. Draw $\beta_{k+1}$ from $\pi(\beta|m_{k+1}, \alpha_{k+1}, \boldsymbol{t})$ in (28) using an MH procedure. We propose $\beta_p$ from a lognormal distribution centered at the previous value, i.e., $\log \beta_p \sim N\left(\log \beta_k, c_\beta^2 \sigma_k^2\right)$, where $c_\beta > 0$ is a tuning parameter. For the purpose of sampling efficiency, we intend to specify a proposal distribution, which closely resembles the conditional posterior of $\beta$. This consideration prompts us to specify the variance term $\sigma_k^2$, evaluated at updated values $(m_{k+1}, \alpha_{k+1}, \beta_k)$, as the reciprocal of Fisher information of the conditional posterior of $\log \beta$, whose log-likelihood is $\log \pi(\beta|m_{k+1}, \alpha_{k+1}, \boldsymbol{t}) + \log |J|$. The Jacobian term $J = \beta$ is needed here as we make a log transformation on $\beta$.

$$\begin{aligned} \sigma_k^2 &= \left[-E\left\{\frac{\partial^2(\log \pi(\beta|m_{k+1},\alpha_{k+1},\boldsymbol{t})+\log|J|)}{\partial(\log\beta)^2}\right\}\right]^{-1}\Bigg|_{m=m_{k+1},\alpha=\alpha_{k+1},\beta=\beta_k} \\ &= \left[-\beta_k^2 E\left(\frac{\partial^2 \ell}{\partial \beta^2}\right) + \frac{a_1\beta_k}{2\alpha_{k+1}^2} + \frac{b_1}{2\beta_k}\right]^{-1} = \left[\beta_k^2 u_{\beta_k\beta_k} + \frac{a_1\beta_k}{2\alpha_{k+1}^2} + \frac{b_1}{2\beta_k}\right]^{-1}, \end{aligned} \tag{31}$$

where $u_{\beta_k\beta_k} = 2nm_{k+1}^2/(\alpha_{k+1}^2\beta_k^2)[\alpha_{k+1}^2 - 2\alpha_{k+1}h(\alpha_{k+1}) + 2]$ from Equation (19). Finally, take $\beta_{k+1} = \beta_p$ with probability:

$$\lambda_\beta = \min\left\{1, \frac{\pi(\beta_p|m_{k+1}, \alpha_{k+1}, \boldsymbol{t}) \times q_\beta(\beta_k|\beta_p)}{\pi(\beta_k|m_{k+1}, \alpha_{k+1}, \boldsymbol{t}) \times q_\beta(\beta_p|\beta_k)}\right\}, \tag{32}$$

with the proposal density $q_\beta(\cdot)$, and so $q_\beta(\beta_k|\beta_p)/q_\beta(\beta_p|\beta_k) = \beta_p/\beta_k$.

The posterior inference of the parameters $m, \alpha$ and $\beta$ is made by their posterior samples $(m_k, \alpha_k, \beta_k), k = 1, 2, ..., M$, such as posterior mean, credible interval (CI), etc.

## 4. Simulation Study

We conduct a simulation study to assess the performance of parameter estimation by ML and the Bayesian methods, where we fix the scale parameter $\beta = 1$ and take four settings of the other two as $(m, \alpha) = (0.25, 0.50), (0.50, 0.50), (1.00, 1.00), (1.50, 2.00)$. We generate 1000 datasets for each of these parameter settings with three sample sizes $n = 20, 30, 50$. For the Bayesian analysis, we choose the hyperparameter values $a_0 = a_1 = b_0 = b_1 = d_0 = d_1 = 0$, such that the prior distributions are rather "flat" or "less informative" to reflect little prior knowledge about the parameters. With each simulated data, we find that the rule of thumb $c_m = 15, c_\beta = 2.4$ as the tuning parameter values are adequate to ensure the acceptance rates to hover around 35–40%, and we run five MCMC chains with fairly different initial values and each with a burn-in period of 2000 followed by 8000 iterations. The scale reduction factor estimate $\sqrt{\widehat{RF}} = \sqrt{Var(\theta)/W}$ is used to monitor the convergence of MCMC simulations [27], where $\theta$ is the estimand of the parameter of interest, $Var(\theta) = (N-1)W/N + B/N$, with the iteration number $N$ for each chain and the between- and within-sequence variances $B$ and $W$. The scale factors for the sequences of $m, \alpha$, and $\beta$ are within 1.00–1.02 for all five MCMC chains, indicating their convergence. The remaining 8000 samples are used to compute the average biases, mean squared error (MSE) of the estimates, average lengths (AL) of the 95% credible intervals (CI), and coverage probability (CP) for the parameters. The results are shown in Table 1 along with these estimates from the ML method for comparison purposes. The main features are summarized as follows: (i) as expected, the bias of estimates, MSE, and AL of 95% CI decrease, and CP is closer to the nominal level as sample size $n$ increases for all cases; (ii) the estimation of all parameters from the Bayesian method is much better than from the ML approach in terms of smaller biases and MSEs, narrower CIs and higher CPs. The MLE $\hat{\alpha}$ dose not perform well for the small to moderate sample sizes ($n = 20, 30$). (iii) comparatively, both methods produce much more accurate estimation of $\beta$, less precise for $m$, and least for $\alpha$. With the larger sample size ($n = 50$), the performances of the $\beta$ estimates are similar

for both methods; (iv) it seems that the estimate of $\alpha$ has a smaller MSE under a smaller true value of $\alpha$, whereas the estimate of $m$ has a smaller MSE under a larger true $m$. Overall, the Bayesian inference method outperforms the ML approach for all parameter settings, especially when the sample size is small.

**Table 1.** GBS-II estimation results for simulated data. AL, average length; CP, coverage probability.

| $n$ | | ML Method | | | | Bayesian | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Bias** | **MSE** | **AL** | **CP (%)** | **Bias** | **MSE** | **AL** | **CP (%)** |
| | | | | | $m = 0.25, \alpha = 0.5, \beta = 1.0$ | | | | |
| 20 | $m$ | 0.1229 | 0.2861 | 1.2258 | 92.45 | 0.1166 | 0.1889 | 0.7493 | 93.11 |
| | $\alpha$ | 0.1375 | 0.1578 | 0.4718 | 92.27 | 0.1302 | 0.1145 | 0.4523 | 93.19 |
| | $\beta$ | 0.0302 | 0.0736 | 0.2905 | 93.87 | 0.0216 | 0.0622 | 0.2322 | 94.69 |
| 30 | | 0.1092 | 0.2363 | 1.1052 | 93.17 | 0.1007 | 0.1567 | 0.5329 | 94.82 |
| | | 0.1236 | 0.1353 | 0.4546 | 93.21 | 0.1149 | 0.1072 | 0.4076 | 94.65 |
| | | 0.0244 | 0.0668 | 0.2118 | 94.29 | 0.0210 | 0.0516 | 0.1209 | 95.08 |
| 50 | | 0.1035 | 0.1339 | 0.9777 | 94.43 | 0.0916 | 0.1113 | 0.2244 | 94.94 |
| | | 0.1121 | 0.1151 | 0.4033 | 94.67 | 0.1108 | 0.0968 | 0.3323 | 95.15 |
| | | −0.0112 | 0.0240 | 0.1352 | 95.15 | 0.0036 | 0.0116 | 0.1054 | 95.20 |
| | | | | | $m = 0.5, \alpha = 0.5, \beta = 1.0$ | | | | |
| 20 | $m$ | 0.1104 | 0.1304 | 1.0141 | 93.27 | 0.1072 | 0.1184 | 0.7345 | 93.87 |
| | $\alpha$ | 0.1229 | 0.1073 | 0.3694 | 93.51 | 0.1115 | 0.1016 | 0.3541 | 93.50 |
| | $\beta$ | 0.0314 | 0.0940 | 0.3181 | 93.85 | −0.0214 | 0.0720 | 0.2299 | 94.71 |
| 30 | | 0.1051 | 0.1231 | 0.9386 | 94.52 | 0.1015 | 0.1125 | 0.5687 | 94.84 |
| | | 0.1181 | 0.1063 | 0.3457 | 94.40 | 0.1047 | 0.0905 | 0.3186 | 94.77 |
| | | 0.0158 | 0.0674 | 0.2422 | 94.88 | 0.0083 | 0.0613 | 0.1646 | 95.14 |
| 50 | | 0.0809 | 0.1150 | 0.5503 | 94.48 | 0.0202 | 0.0822 | 0.3789 | 95.18 |
| | | 0.0492 | 0.0738 | 0.2128 | 94.53 | 0.0327 | 0.0578 | 0.1832 | 95.22 |
| | | 0.0080 | 0.0345 | 0.1957 | 95.12 | 0.0023 | 0.0207 | 0.0819 | 95.29 |
| | | | | | $m = 1.0, \alpha = 1.0, \beta = 1.0$ | | | | |
| 20 | $m$ | 0.1792 | 0.1357 | 0.9778 | 91.94 | 0.1620 | 0.1212 | 0.7107 | 93.02 |
| | $\alpha$ | 0.2242 | 0.2433 | 0.6218 | 92.19 | 0.2032 | 0.1493 | 0.4865 | 92.89 |
| | $\beta$ | 0.0371 | 0.1067 | 0.4728 | 93.56 | 0.0272 | 0.0831 | 0.2191 | 94.05 |
| 30 | | 0.1628 | 0.1230 | 0.9018 | 93.52 | 0.1464 | 0.1108 | 0.6109 | 94.28 |
| | | 0.1854 | 0.2038 | 0.5896 | 93.88 | 0.1494 | 0.1247 | 0.4033 | 94.29 |
| | | −0.0232 | 0.0728 | 0.3480 | 94.26 | −0.0105 | 0.0624 | 0.1122 | 94.89 |
| 50 | | 0.1155 | 0.1103 | 0.5135 | 93.49 | 0.0974 | 0.1042 | 0.4278 | 94.58 |
| | | 0.1355 | 0.1665 | 0.4071 | 94.31 | 0.1134 | 0.1046 | 0.3083 | 94.82 |
| | | 0.0138 | 0.0416 | 0.2598 | 95.03 | 0.0094 | 0.0310 | 0.0844 | 95.26 |
| | | | | | $m = 1.5, \alpha = 2.0, \beta = 1.0$ | | | | |
| 20 | $m$ | 0.2832 | 0.1303 | 0.9650 | 92.38 | 0.2176 | 0.1163 | 0.7058 | 93.41 |
| | $\alpha$ | 0.3482 | 0.2777 | 0.8674 | 93.20 | 0.2887 | 0.2214 | 0.7195 | 93.14 |
| | $\beta$ | −0.0526 | 0.1432 | 0.5216 | 93.92 | 0.0411 | 0.0933 | 0.2852 | 94.17 |
| 30 | | 0.2583 | 0.1298 | 0.8960 | 93.35 | 0.1734 | 0.1091 | 0.6779 | 94.11 |
| | | 0.3009 | 0.2356 | 0.7744 | 93.83 | −0.1939 | 0.2042 | 0.6438 | 94.71 |
| | | −0.0423 | 0.1061 | 0.4238 | 94.33 | 0.0320 | 0.0720 | 0.1661 | 94.86 |
| 50 | | 0.1598 | 0.1091 | 0.7047 | 94.38 | 0.1145 | 0.0928 | 0.5682 | 95.03 |
| | | −0.2064 | 0.1836 | 0.5209 | 94.73 | 0.1303 | 0.1555 | 0.4105 | 95.17 |
| | | −0.0209 | 0.0674 | 0.2675 | 94.92 | 0.0106 | 0.0411 | 0.0784 | 95.30 |

## 5. Real Data Analysis

To further illustrate the usefulness of our method in parameter inference for the GBS-II, we consider a real data example given by [28] about the breakdown time in an accelerated test that employed a pair of parallel disk electrodes immersed in an insulating oil. Voltage $V$ across the pair was increased linearly with time $t$ at a specified rate $R$, and breakdown time was recorded at a one square inch electrode. The data are presented in Table 2, consisting of 60 measured

breakdown times (seconds). Fitting the GBS-II distribution with the Bayesian method, we choose the hyperparameters $b_0$ and $b_1$ such that the prior mean of $\beta$ is close to the sample median 4.3 of the data, $a_0$ and $a_1$ such that the conditional prior mean $E(\alpha^2|\beta = 4.3)$ is close to the MLE $\widehat{\alpha}^2 = 1.3656$, and the CV (coefficient of variation) of the gamma or inverse gamma priors is close to the CV of standard uniform distribution $(1/\sqrt{3})$. Hence, we have the hyperparameter values as follows: $a_0 = 10, a_1 = 2.54, b_0 = 0.67, b_1 = 0.16, d_0 = d_1 = 1$. We run a chain of 20,000 iterations with a burn-in period of 5000. To reduce the correlation among the samples, every fifth sample of the remaining 15,000 samples is used for posterior inference. The results are tabulated in Table 3, where, due to the relatively large sample size ($n = 60$), the point estimates obtained by both ML and Bayesian methods are close to each other. However, the 95% CIs of the Bayesian method are narrower than the ones of the ML approach, especially for the intervals of $m$ and $\alpha$. Additionally, the Chi-squared goodness of fit statistic and BIC values of model fitting by the Bayesian method are smaller than these by the ML method, indicating the greater accuracy in the Bayesian analysis. Based on the estimates of $m$ and $\beta$ and the given standard level of stress $V_0 = 42.30$, together with the relation in (9), the power $p$ in the inverse power law and the rate of rise of voltage $R$ are computed, respectively from ML and Bayesian approaches, as $\hat{p} = 8.9455, 8.4045$ and $\hat{R} = 11.0734, 10.1067$. Our results (especially from the Bayesian) are close to the estimates obtained by [28], who fitted a Weibull distribution for these data. For graphical comparison, Figure 3 shows the histogram of the data and fitted GBS-II density curves estimated by both methods.

**Table 2.** Oil breakdown time (seconds) in an accelerated test employed in an insulating oil.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.4 | 3.4 | 3.4 | 3.5 | 3.5 | 3.5 | 3.6 | 3.8 | 3.8 | 3.8 | 3.8 | 3.9 | 3.9 | 3.9 | 4.0 |
| 4.0 | 4.0 | 4.0 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.1 | 4.2 | 4.2 | 4.2 | 4.2 | 4.2 | 4.3 |
| 4.3 | 4.3 | 4.3 | 4.3 | 4.4 | 4.4 | 4.4 | 4.4 | 4.4 | 4.4 | 4.4 | 4.5 | 4.5 | 4.6 | 4.6 |
| 4.6 | 4.6 | 4.6 | 4.7 | 4.7 | 4.7 | 4.7 | 4.7 | 4.8 | 4.9 | 4.9 | 4.9 | 5.0 | 5.1 | 5.2 |

**Table 3.** Breakdown time: estimation results

| **Method** | $m$ **(95% CI)** | $\alpha$ **(95% CI)** | $\beta$ **(95% CI)** | $\chi^2$**, BIC** |
|---|---|---|---|---|
| ML method | 4.9728 (2.0185, 12.2511) | 1.1686 (0.3807, 3.5874) | 4.2058 (3.7681, 4.7816) | 16.19, 125.81 |
| Bayesian | 4.6523 (3.1521, 6.2339) | 1.2391 (0.9279, 2.3936) | 4.5614 (4.4250, 4.6945) | 8.82, 121.62 |

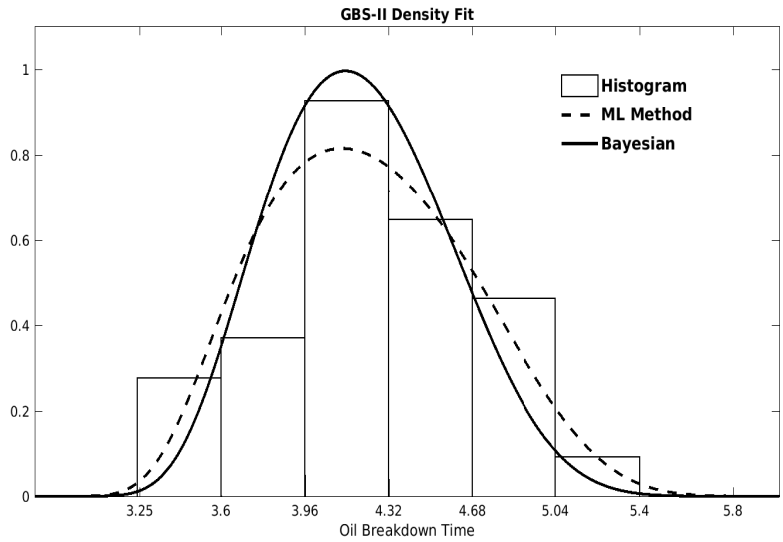


**Figure 3.** Failure data: histogram and fitted density curves by ML and Bayesian estimation methods.

The second real data were presented in [29] on active repair times (in hours) for an airborne communications transceiver. To illustrate the estimation performance on a small sample size, we randomly select 20 repair times out of the total of 46 observations to have the following data: 0.3, 0.5, 0.6, 0.6, 0.7, 0.7, 0.8, 1.0, 1.3, 1.5, 1.5, 2.0, 2.2, 2.5, 4.0, 4.7, 5.0, 7.5, 8.8, 22.0. Modeling the data by the GBS-II distribution, we adopt the same procedure as discussed to specify the hyperparameter values by using the sample median 1.5 of the data and MLE $\widehat{\alpha}^2 = 2.83$. In summary, we choose the following hyperparameter values: $a_0 = 10, a_1 = 15.07, b_0 = 0.67, b_1 = 1, d_0 = d_1 = 1$. An MCMC chain of 20,000 iterations with a burn-in period of 5000 produces the estimation results of the Bayesian approach, together with the results by the ML method, tabulated in Table 4. For these data with a relatively small sample size ($n = 20$), the produced 95% CIs are much narrower, as well as having much smaller Chi-squared goodness of fit statistic and BIC values of model fitting by the Bayesian method. Finally, for illustration, Figure 4 shows that the fitted GBS-II density curve estimated by the Bayesian method is better fit to the histogram than the one by the ML estimation. These outcomes demonstrate that the proposed Bayesian method produces much more accurate inference under the small sample size.

**Table 4.** Repair time: estimation results.

| Method | $m$ (95% CI) | $\alpha$ (95% CI) | $\beta$ (95% CI) | $\chi^2$, BIC |
|---|---|---|---|---|
| ML Method | 0.8326 (0.2170, 1.8444) | 1.6813 (0.3743, 7.5531) | 2.6093 (1.3854, 3.2163) | 97.30, 103.37 |
| Bayesian | 0.4717 (0.3229, 0.6202) | 1.0291 (0.7826, 1.5031) | 1.7718 (1.0633, 2.9523) | 12.98, 92.36 |

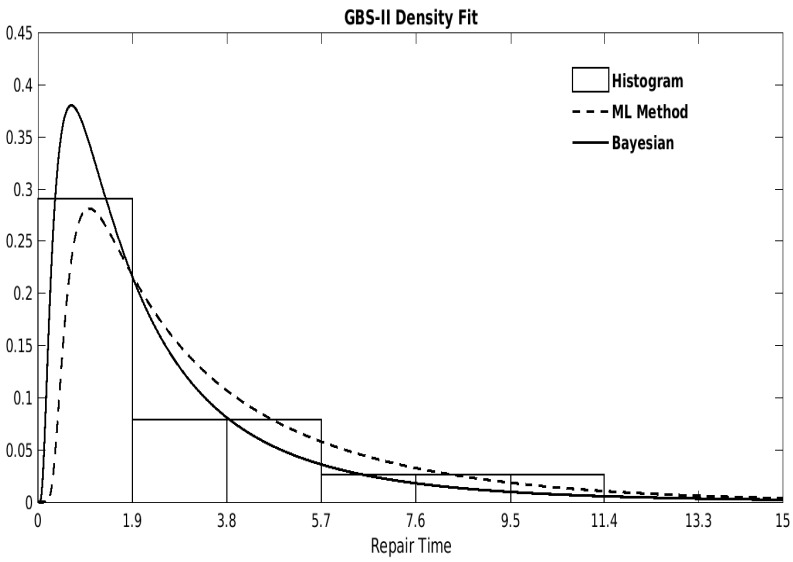


**Figure 4.** Repair data: histogram and fitted density curves by ML and Bayesian estimation methods.

## 6. Conclusions Remarks

We presented a Bayesian inference approach of parameter estimation in progressive stress accelerated life testing with the Birnbaum–Saunders (BS) model, which induced the Type-II generalized Birnbaum–Saunders (GBS-II) distribution followed by the lifetime of products in the testing. We summarized the properties of GBS-II and studied its Fisher's information used in the likelihood-based inference method. The simulation study demonstrated that the Bayesian method outperformed the traditional likelihood-based approach, especially its efficient and impressive performance under small sample sizes. We have also illustrated, with two real datasets, that our Bayesian method can be readily applied for efficient, reliable, and precise inference.

**Funding:** This research was funded by NSF CMMI-0654417 and NIMHD-2G12MD007592.

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

## Appendix A

We present a detailed derivation of the Fisher information matrix for the GBS-II$(m, \alpha, \beta)$. From the fact that:

$$\frac{\partial \varepsilon(t)}{\partial \beta} = -m\beta^{-1}\delta(t), \ \frac{\partial \varepsilon(t)}{\partial m} = \log(\beta^{-1}t)\delta(t), \tag{A1}$$

$$\frac{\partial \delta(t)}{\partial \beta} = -m\beta^{-1}\varepsilon(t), \ \frac{\partial \delta(t)}{\partial m} = \log(\beta^{-1}t)\varepsilon(t), \tag{A2}$$

$$\frac{\partial(\varepsilon(t)\delta(t))}{\partial \beta} = -\frac{m}{\beta}\left(\varepsilon^2(t)+\delta^2(t)\right), \ \frac{\partial(\varepsilon(t)\delta(t))}{\partial m} = \log(\beta^{-1}t)\left(\varepsilon^2(t)+\delta^2(t)\right), \tag{A3}$$

$$\frac{\partial\left(\frac{\varepsilon(t)}{\delta(t)}\right)}{\partial \beta} = -\frac{m}{\beta}\left(1-\frac{\varepsilon^2(t)}{\delta^2(t)}\right) = -\frac{4m}{\beta\delta^2(t)}, \tag{A4}$$

$$\frac{\partial\left(\frac{\varepsilon(t)}{\delta(t)}\right)}{\partial m} = \log(\beta^{-1}t)\left(1-\frac{\varepsilon^2(t)}{\delta^2(t)}\right) = \frac{4\log(\beta^{-1}t)}{\delta^2(t)}, \tag{A5}$$

and then the second partial derivatives of parameters for the log-likelihood function in (13) are in the following:

$$\frac{\partial^2 \ell}{\partial \alpha^2} = \frac{n}{\alpha^2} - \frac{3}{\alpha^4}\sum_{i=1}^{n}\varepsilon^2(t_i), \ \frac{\partial^2 l}{\partial\alpha\partial\beta} = -\frac{2m}{\alpha^3\beta}\sum_{i=1}^{n}\varepsilon(t_i)\delta(t_i), \tag{A6}$$

$$\frac{\partial^2 \ell}{\partial\alpha\partial m} = \frac{2}{\alpha^3}\sum_{i=1}^{n}\log(\beta^{-1}t_i)\varepsilon(t_i)\delta(t_i), \tag{A7}$$

$$\begin{aligned}\frac{\partial^2 \ell}{\partial \beta^2} &= \frac{m}{\beta^2}\left[\sum_{i=1}^{n}\frac{\varepsilon(t_i)}{\delta(t_i)} - \frac{1}{\alpha^2}\sum_{i=1}^{n}\varepsilon(t_i)\delta(t_i)\right] \\ &+ \frac{m^2}{\beta^2}\left[4\sum_{i=1}^{n}\frac{1}{\delta^2(t_i)} - \frac{1}{\alpha^2}\sum_{i=1}^{n}(\varepsilon^2(t_i)+\delta^2(t_i))\right],\end{aligned} \tag{A8}$$

$$\begin{aligned}\frac{\partial^2 \ell}{\partial\beta\partial m} &= -\frac{1}{\beta}\left[\sum_{i=1}^{n}\frac{\varepsilon(t_i)}{\delta(t_i)} - \frac{1}{\alpha^2}\sum_{i=1}^{n}\varepsilon(t_i)\delta(t_i)\right] \\ &- \frac{m}{\beta}\left[4\sum_{i=1}^{n}\frac{\log(\beta^{-1}t_i)}{\delta^2(t_i)} - \frac{1}{\alpha^2}\sum_{i=1}^{n}\log(\beta^{-1}t_i)(\varepsilon^2(t_i)+\delta^2(t_i))\right],\end{aligned} \tag{A9}$$

$$\frac{\partial^2 \ell}{\partial m^2} = -\frac{n}{m^2} + 4\sum_{i=1}^{n}\frac{\log^2(\beta^{-1}t_i)}{\delta^2(t_i)} - \frac{1}{\alpha^2}\sum_{i=1}^{n}\log^2(\beta^{-1}t_i)(\varepsilon^2(t_i)+\delta^2(t_i)). \tag{A10}$$

Let $z = \varepsilon(t)/\alpha$, and so, $\delta(t) = \sqrt{\alpha^2 z^2+4}$ and $\log(t/\beta) = \log\left[(\alpha z+\sqrt{\alpha^2z^2+4})/2\right]/m = g(z)/m$. By the fact that $Z = \varepsilon(T)/\alpha \sim \mathrm{N}(0,1)$ and $g(z) = \log\left[(\alpha z+\sqrt{\alpha^2z^2+4})/2\right]$ is an odd function of $z$, we have:

$$\mathrm{E}(\varepsilon^2(T)) = \alpha^2\mathrm{E}(Z^2) = \alpha^2, \ \mathrm{E}(\delta^2(T)) = \mathrm{E}(\alpha^2Z^2+4) = \alpha^2+4, \tag{A11}$$

$$\mathrm{E}(\varepsilon(T)\delta(T)) = \alpha\mathrm{E}(Z\sqrt{\alpha^2Z^2+4}) = 0, \ \mathrm{E}\left(\frac{\varepsilon(T)}{\delta(T)}\right) = \alpha\mathrm{E}\left(\frac{Z}{\sqrt{\alpha^2Z^2+4}}\right) = 0, \tag{A12}$$

$$\mathrm{E}\left(\frac{1}{\delta^2(T)}\right) = \mathrm{E}\left(\frac{1}{\alpha^2 Z^2+4}\right) = \frac{h(\alpha)}{\alpha},\ \mathrm{E}\left[\frac{\log(\beta^{-1}T)}{\delta^2(T)}\right] = \frac{1}{m}\mathrm{E}\left(\frac{g(Z)}{\alpha^2 Z^2+4}\right) = 0, \tag{A13}$$

$$\mathrm{E}\left(\frac{\log^2(\beta^{-1}T)}{\delta^2(T)}\right) = \frac{1}{m^2}\mathrm{E}\left(\frac{g^2(Z)}{\alpha^2 Z^2+4}\right), \tag{A14}$$

$$\mathrm{E}\left[\log(\beta^{-1}T)\varepsilon(T)\delta(T)\right] = \frac{\alpha}{m}\mathrm{E}\left[Zg(Z)\sqrt{\alpha^2 Z^2+4}\right], \tag{A15}$$

$$\mathrm{E}\left[\log(\beta^{-1}T)(\varepsilon^2(T)+\delta^2(T))\right] = \frac{2}{m}\mathrm{E}[g(Z)(\alpha^2 Z^2+2)] = 0, \tag{A16}$$

$$\mathrm{E}\left[\log^2(\beta^{-1}T)(\varepsilon^2(T)+\delta^2(T))\right] = \frac{2}{m^2}\mathrm{E}[g^2(Z)(\alpha^2 Z^2+2)], \tag{A17}$$

where $\mathrm{E}(1/(\alpha^2 Z^2+4)) = h(\alpha)/\alpha$ with $h(\alpha) = \sqrt{\pi/2}e^{2/\alpha^2}(1-\Phi(2/\alpha))$ being provided in [4]. Then, the elements of Fisher's information in (16) are:

$$u_{mm} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial m^2}\right) = \frac{n}{m^2} - \frac{4n}{m^2}\mathrm{E}\left(\frac{g^2(Z)}{\alpha^2 Z^2+4}\right) + \frac{2n}{m^2\alpha^2}E\left[g^2(Z)(\alpha^2 Z^2+2)\right], \tag{A18}$$

$$u_{\alpha m} = u_{m\alpha} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\alpha\partial m}\right) = -\frac{2n}{\alpha^2 m}\mathrm{E}\left[Zg(Z)\sqrt{\alpha^2 Z^2+4}\right], \tag{A19}$$

$$u_{\beta m} = u_{m\beta} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\beta\partial m}\right) = 0, \tag{A20}$$

$$u_{\alpha\alpha} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\alpha^2}\right) = \frac{2n}{\alpha^2},\ u_{\alpha\beta} = u_{\beta\alpha} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\alpha\partial\beta}\right) = 0, \tag{A21}$$

$$u_{\beta\beta} = -\mathrm{E}\left(\frac{\partial^2\ell}{\partial\beta^2}\right) = \frac{2m^2 n}{\alpha^2\beta^2}\left[\alpha^2 - 2\alpha h(\alpha) + 2\right]. \tag{A22}$$

Therefore, the Fisher's information matrix and its inverse have the following forms:

$$I(m,\alpha,\beta) = \begin{pmatrix} u_{mm} & u_{m\alpha} & 0 \\ u_{\alpha m} & u_{\alpha\alpha} & 0 \\ 0 & 0 & u_{\beta\beta} \end{pmatrix},\ I^{-1}(m,\alpha,\beta) = \begin{pmatrix} u_{\alpha\alpha}/c & -u_{m\alpha}/c & 0 \\ -u_{\alpha m}/c & u_{mm}/c & 0 \\ 0 & 0 & 1/u_{\beta\beta} \end{pmatrix} \tag{A23}$$

with $c = u_{mm}u_{\alpha\alpha} - u_{m\alpha}^2$. When $\alpha$ is small, some approximations can be implemented for the following terms:

$$\begin{aligned}
\mathrm{E}(Zg(Z)\sqrt{\alpha^2 Z^2+4}) &= \int_{-\infty}^{\infty} zg(z)\sqrt{\alpha^2 z^2+4}\phi(z)dz = -\int_{-\infty}^{\infty} g(z)\sqrt{\alpha^2 z^2+4}d\phi(z) \\
&= \int_{-\infty}^{\infty}\phi(z)d(g(z)\sqrt{\alpha^2 z^2+4}) = \int_{-\infty}^{\infty}\phi(z)\left[\alpha + \frac{\alpha^2 z}{\sqrt{\alpha^2 z^2+4}}g(z)\right]dz \\
&= \alpha + \frac{\alpha}{2}\int_{-\infty}^{\infty} z\phi(z)dg^2(z) = \alpha - \frac{\alpha}{2}\int_{-\infty}^{\infty} g^2(z)(1-z^2)\phi(z)dz \\
&= \alpha - \frac{\alpha}{2}\mathrm{E}\left[\frac{\alpha^2 Z^2}{4}\left(1 - \frac{\alpha^2 Z^2}{12} + \frac{\alpha^4 Z^4}{90} + O(\alpha^6)\right)(1-Z^2)\right] \\
&= \alpha - \frac{\alpha^3}{8}\left[\mathrm{E}(Z^2) - \left(1+\frac{\alpha^2}{12}\right)\mathrm{E}(Z^4) + \left(\frac{\alpha^2}{12}+\frac{\alpha^4}{90}\right)\mathrm{E}(Z^6) + O(\alpha^6)\right] \\
&= \frac{\alpha}{8}(8+2\alpha^2+\alpha^4) + O(\alpha^7),
\end{aligned} \tag{A24}$$

$$
\begin{aligned}
\mathrm{E}\left(\frac{g^2(Z)}{\alpha^2 Z^2+4}\right) &= \mathrm{E}\left(\frac{\frac{\alpha^2 Z^2}{4}\left[1-\frac{\alpha^2 Z^2}{12}+O(\alpha^4)\right]}{\alpha^2 Z^2+4}\right) \\
&= \frac{1}{48}\mathrm{E}\left(\frac{-(\alpha^2 Z^2+4)^2+20(\alpha^2 Z^2+4)-64}{\alpha^2 Z^2+4}\right)+O(\alpha^4) \\
&= \frac{1}{48}\left[-(\alpha^2\mathrm{E}(Z^2)+4)+20-64\mathrm{E}\left(\frac{1}{\alpha^2 Z^2+4}\right)\right]+O(\alpha^4) \\
&= \frac{1}{48}\left(16-\alpha^2-\frac{64h(\alpha)}{\alpha}\right)+O(\alpha^4), \qquad \text{(A25)} \\
\mathrm{E}(g^2(Z)(\alpha^2 Z^2+2)) &= \mathrm{E}\left[\frac{\alpha^2 Z^2}{4}\left(1-\frac{\alpha^2 Z^2}{2}+O(\alpha^4)\right)(\alpha^2 Z^2+2)\right] \\
&= \frac{1}{4}\left[2\alpha^2\mathrm{E}(Z^2)+\frac{5\alpha^4}{6}\mathrm{E}(Z^4)+O(\alpha^6)\right]=\frac{\alpha^2}{8}(4+5\alpha^2)+O(\alpha^6). \qquad \text{(A26)}
\end{aligned}
$$

## References


1. Birnbaum, Z.W.; Saunders, S.C. A new family of life distributions. *J. Appl. Probab.* **1969**, *6*, 319–327. [CrossRef]
2. Dupuis, D.J.; Mills, J.E. Robust estimation of the Birnbaum–Saunders distribution. *IEEE Trans. Reliab.* **1998**, *47*, 88–95. [CrossRef]
3. Engelhardt, M.; Bain, L.J.; Wright, F.T. Inferences on the parameters of the Birnbaum–Saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics* **1981**, *23*, 251–256. [CrossRef]
4. Lemonte, A.J.; Cribari-Neto, F.; Vasconcellos, K.L.P. Improved statistical inference for the two-parameter Birnbaum–Saunders distribution. *Comput. Stat. Data Anal.* **2007**, *51*, 4656–4681. [CrossRef]
5. Ng, H.K.T.; Kundub, D.; Balakrishnan, N. Modified moment estimation for the two-parameter Birnbaum–Saunders distribution. *Comput. Stat. Data Anal.* **2003**, *43*, 283–298. [CrossRef]
6. Sha, N.; Ng, T.L. Bayesian inference for Birnbaum–Saunders distribution and its generalization. *J. Stat. Comput. Simul.* **2017**, *87*, 2411–2429. [CrossRef]
7. Wang, M.; Sun, X.; Park, C. Bayesian analysis of Birnbaum–Saunders distribution via the generalized ratio-of-uniforms method. *Comput. Stat.* **2016**, *31*, 207–225. [CrossRef]
8. Xu, A.; Tang, Y. Bayesian analysis of Birnbaum–Saunders distribution with partial information. *Comput. Stat. Data Anal.* **2011**, *55*, 2324–2333. [CrossRef]
9. Bagdonavicius, V.B.; Gerville-Reache, L.; Nikulin, M. Parametric Inference for step-stress models. *IEEE Trans. Reliab.* **2002**, *51*, 27–31. [CrossRef]
10. Bai, D.S.; Chun, Y.R. Optimum simple step-stress accelerated life tests with competing causes of failure. *IEEE Trans. Reliab.* **1991**, *40*, 622–627. [CrossRef]
11. Khamis, I.H. Optimum M-step step-stress test with $K$ stress variables. *Commun. Stat. Simul. Comput.* **1997**, *26*, 1301–1313. [CrossRef]
12. Sha, N.; Pan, R. Bayesian analysis for step-stress accelerated life testing using weibull proportional hazard model. *Stat. Pap.* **2014**, *55*, 715–726. [CrossRef]
13. Srivastava, P.W.; Shukla, R. A log-logistic step-stress model. *IEEE Trans. Reliab.* **2008**, *57*, 431–434. [CrossRef]
14. Owen, W.J.; Padgett, W.J. A Birnbaum–Saunders accelerated life model. *IEEE Trans. Reliab.* **2000**, *49*, 224–229. [CrossRef]
15. Sun, T.; Shi, Y. Estimation for Birnbaum–Saunders distribution in simple step stress-accelerated life test with Type-II censoring. *Commun. Stat. Simul. Comput.* **2016**, *45*, 880–901. [CrossRef]
16. Nelson, W. Accelerated life testing step-stress models and data analysis. *IEEE Trans. Reliab.* **1980**, *29*, 103–108. [CrossRef]
17. Díaz-García, J.A.; Domínguez-Molina, J.R. Some generalizations of Birnbaum–Saunders and sinh-normal distributions. *Int. Math. Forum* **2006**, *1*, 1709–1727. [CrossRef]
18. Fierro, R.; Leiva, V.; Ruggeri, F.; Sanhuezad, A. On a Birnbaum–Saunders distribution arising from a non-homogeneous Poisson process. *Stat. Probab. Lett.* **2013**, *83*, 1233–1239. [CrossRef]

19. Owen, W.J. A new three-parameter extension to the Birnbaum–Saunders distribution. *IEEE Trans. Reliab.* **2006**, *55*, 475–479. [CrossRef]
20. Desmond, A.F. On the relationship between two fatigue-life models. *IEEE Trans. Reliab.* **1986**, *35*, 167–169. [CrossRef]
21. Owen, W.J.; Ng, H.K.T. Revisit of relationships and models for the Birnbaum–Saunders and inverse-Gaussian distribution. *J. Stat. Distrib. Appl.* **2015**, *2*. [CrossRef]
22. Rieck, J.R. A moment-generating function with application to the Birnbaum–Saunders distribution. *Commun. Stat. Theory Methods* **1999**, *28*, 2213–2222. [CrossRef]
23. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*; John Wiley & Sons: New York, NY, USA, 1991; Volume 2.
24. Meeker, W.Q.; Escobar, L.A. *Statistical Methods for Reliability Data*; John Wiley & Sons: New York, NY, USA, 1998.
25. Casella, G.; George, E.I. Explaining the Gibbs sampler. *Am. Stat.* **1992**, *46*, 167–174. [CrossRef]
26. Chib, S.; Greenberg, E. Understanding the Metropolis-Hastings algorithm. *Am. Stat.* **1995**, *49*, 327–335. [CrossRef]
27. Gelman, A.; Carlin, J.B.; Stern, H.S.; Rubin, D.B. *Bayesian Data Analysis*, 2nd ed.; Chapman & Hall: London, UK, 2004.
28. Nelson, W. *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*; John Wiley & Sons, Inc.: New York, NY, USA, 2004; pp. 507–509.
29. Balakrishnan, N.; Leiva, V.; Sanhueza, A.; Cabrera, E. Mixture inverse Gaussian distribution and its transformations, moments and applications. *Statistics* **2009**, *43*, 91–104. [CrossRef]