Model Misspecification of Log-Normal and Birnbaum-Saunders Distributions

Suparna Basu *& Debasis Kundu †

Abstract

Model misspecification can be a serious issue in any lifetime data analysis. Assumption of the correct model is very important particularly in prediction for future observations and for estimating the tail probabilities of any lifetime distribution. In this paper we have considered the model misspecification of the log-normal and Birnbaum-Saunders distributions. These two distributions have striking similarities both in terms of the probability density functions and hazard functions, in certain range of parameter values, which makes it extremely difficult to detect the correct model. Hence, to identify the correct model, first the conventional method of ratio of maximized likelihood approach has been tried. The necessary theoretical results have been derived. Based on extensive simulations it has been observed that for certain range of parameter values the model misspecification can be quite high even for very large sample sizes. Some of the other methods like the ratio of maximized product of spacings and minimized Kolmogorov-Smirnov distance, have also been explored. But none of the method performs uniformly better than the other over the entire range of the parameter space. Some counter intuitive results have also been obtained. Hence, we finally conclude that it is extremely difficult to choose the correct model in case of log-normal and Birnbaum-Saunders distribution for certain range of parameter values. Finally, we have suggested a modified ratio of maximized likelihood approach, and the performances are quite satisfactory.

Key Words and Phrases: Model choice; lifetime data analysis; likelihood ratio; probability of correct selection.

AMS Subject Classifications: 62F10, 62F03, 62H12.

*Department of Statistics, MMV, Banaras Hindu University, Varanasi, UP, India

†Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

Choosing the *correct* model is very important in any lifetime data analysis. Although, most of the models behave quite well in the *middle* portion of the data, their tail behavior can be significantly different. Model misspecification can be a serious issue particularly for future prediction or for estimating the extreme event. The problem becomes more severe if the two models are non-nested, and they have the same number of parameters, see for example the monograph by Pereira and Pereira [1] in this respect. An extensive amount of work has been done in model misspecification in case of Weibull and log-normal distributions. Both the distributions have been extensively used in lifetime data analysis. Both of them have two parameters each, they are positively skewed and they are non-nested in the sense of Ghosh and Subramanyam [2] metric or the Kullback-Liebler divergence measure as used by Pesaran [3]. Interestingly, if the shape parameter of the Weibull distribution is greater than one, then the shape of the probability density functions (PDFs) of the two distributions are very similar, although their hazard functions are quite different.

Dumonceaux and Antle [4] first considered the problem of model misspefication of Weibull and log-normal distributions, as a testing of hypothesis problem, and provided the critical values for different level of significance in testing one vs. the other. Kundu and Manglick [5] treated this as a model misspecification problem, and provided the minimum sample size needed to specify the correct model. Basu et al. [6] showed how the model misspefication affects in determining the strength for brittle materials. Basavalingappa et al. [7] explored the difficulty of estimating the probability of rare events, which is less than one in one billion in case of integrated circuit (IC), by using misspecified model. Pasari [8] also mentioned that how the wrong model affects in predicting the rare events like high magnitude earthquake in the Himalayan regions. Dey and Kundu [9] investigated the effect of model misspecification in case of Type-II censoring. A detailed comparison of model misspefication of Weibull and log-normal distributions can be seen in Kim and Yum [10].

The aim of this paper is to consider the model misspecification analysis between log-normal and Birnbaum-Saunders distribution. The two-parameter log-normal distribution has the following probability density function for $\theta > 0$ and $\sigma > 0$;

$$f_{LN}(x; \sigma, \theta) = \frac{1}{\sqrt{2\pi}x\sigma} \exp\left(-\frac{(\ln x - \ln\theta)^2}{2\sigma^2}\right); \qquad x > 0, \tag{1}$$

and 0, otherwise. The two-parameter log-normal distribution is one of the most common positively skewed distributions which has been quite extensively used in lifetime data analysis. It has several desirable properties. Due to presence of the shape parameter σ and the scale parameter θ, it is a very flexible distribution which can be used quite conveniently for analyzing various types of data sets. The probability density function (PDF) of a log-normal distribution is always unimodal, and it also has an unimodal hazard function.

Birnbaum and Saunders [11, 12] introduced a two-parameter positively skewed lifetime distribution to model the fatigue life of metals subject to periodic stress. The two-parameter Birnbaum-Saunders distribution has the following probability density function for $\alpha > 0$ and $\beta > 0$;

$$f_{BS}(x; \alpha, \beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\left[\left(\frac{\beta}{x}\right)^{1/2} + \left(\frac{\beta}{x}\right)^{3/2}\right] \exp\left[-\frac{1}{2\alpha^2}\left(\frac{x}{\beta} + \frac{\beta}{x} - 2\right)\right]; \qquad x > 0, \tag{2}$$

and 0, otherwise. It has some interesting physical interpretations also. Recently, the Birnbaum-Saunders distribution has received a considerable amount of attention in the statistical literature because of its flexibility. It also has two parameters, i.e. shape (α) and scale (β). Due to presence of both the shape and scale parameters it has been used quite effectively to analyze different types of lifetime data. Both the probability density function and the hazard functions of a two-parameter Birnbaum-Saunders distribution are always unimodal. There is a striking similarity between a log-normal and a Birnbaum-Saunders distribution for a wide range of parameter values. Interestingly, if the shape parameters are small, although the probability density functions of the two distributions are very close to each other, the mode of the hazard functions can be significantly different. It has also been observed by Lemonte [13] that if the shape parameter of the Birnbaum-Saunders distribution is small then calculation of the mode of the hazard function is a numerically challenging problem, where as in case of log-normal distribution, there is no such issue. Hence, it is quite important to choose the *correct* model.

The objective of this study is to focus on the effect of model misspecification in case of log-normal and Birnbaum-Saunders distributions. That is, if the data are actually coming from a log-normal distribution, and it has been misspecified as a Birnbaum-Saunders distribution or vice versa, then what will be the effect on estimation on the tail probabilities or the mode of the hazard function. First we have explored the most traditional method based on the maximized likelihood approach. We have addressed the effects of model misspecification following the results of White [14], and have obtained the necessary analytical results for this purpose. It has been shown analytically that as $n \to \infty$, the model misspefication approaches zero, as expected. But it may not be felt even for large sample sizes in ceratin cases.

We have performed extensive simulations to verify the asymptotic results, but it is observed that for certain range of parameter values, even with very large sample sizes (say 10,000), the model misspecification is away from zero. Due to this reason, we have explored some of the other methods like minimization of the Kolmogorov-Smirnov distance approach or maximized product of spacings methods. Based on extensive simulations it has been observed that none of the methods outperforms the other over the entire range of parameter values. In fact some counter intuitive results have also been obtained. Hence, based on all these findings we come to a conclusion that it is extremely difficult to identify the correct model in case of log-normal and Birnbaum-Saunders distributions for certain range of parameter values. In those range of parameter values, where it has been observed that the probability of correct selection becomes less than 0.5, we have proposed a modified ratio of maximized log-likelihood approach, and it has a satisfactory performance.

The rest of the paper is organized as follows. In Section 2 we provide some motivations of the problem. The problem formulation, and the traditional method are provided in Section 3. The effect of model misspecification and the necessary theoretical results are provided in Section 4. Simulation results are presented in Section 5. A modified approach is proposed in Section 6. Finally, we conclude the paper in Section 7.

2 Motivation

In this section we will try to motivate the readers about the importance of the problem and the difficulties involved in it, based on two real life data sets.

Data Set 1:

Guinea pigs are well known to have high susceptibility to human tuberculosis. Hence, it is quite common to choose these species to study the effect of tubercle bacilli on its survival times. The following data set represents the survival times of guinea pigs injected with different doses of tuber bacilli. Here, the main concern is with the animals in the same cage and that are under the same regimen. The regimen number is the common logarithm of the number of bacilli units in 0.5 ml. of challenge solution, i.e. the regimen 6.6 corresponds to 4.0×10^6 bacilli units per 0.5 ml. ($\log(4.0 \times 10^6) = 6.6$). The following data set has 72 observations corresponding to regimen 6.6, and it was originally available at Bjerkedal [15].

12, 15, 22, 24, 24, 32, 32, 33, 34, 38, 38, 43, 44, 48, 52, 53, 54, 54, 55, 56, 57, 58, 58, 59, 60, 60, 60, 60, 61, 62, 63, 65, 65, 67, 68, 70, 70, 72, 73, 75, 76, 76, 81, 83, 84, 85, 87, 91, 95,

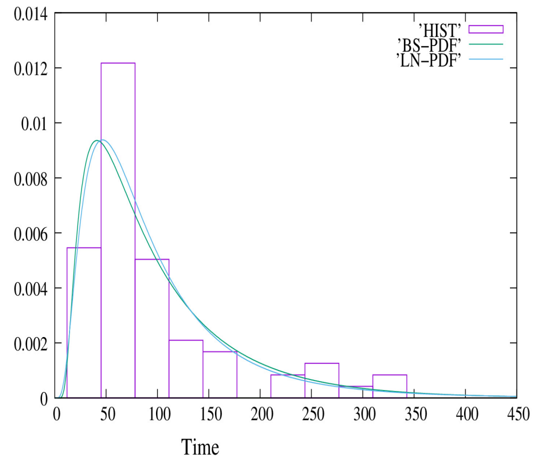

Figure 1: The probability density functions of the fitted Birnbaum-Saunders and Log-normal distributions and the histogram of the Data Set 1.

96, 98, 99, 109, 110, 121, 127, 129, 131, 143, 146, 146, 175, 175, 211, 233, 258, 258, 263, 297, 341, 341, 376

Gupta, Kannan and Raychaudhuri [16] analyzed this data set based on log-normal distribution, and Kundu, Kannan and Balakrishnan [17] analyzed it based on Birnbaum-Saunders distribution. In this case the maximum likelihood estimates of the log-normal parameters are: $\widehat{\sigma} = 0.7104$ and $\widehat{\theta} = 77.0394$, whereas for Birnbaum-Saunders distribution they are: $\widehat{\alpha} = 0.7654$ and $\widehat{\beta} = 77.3728$. In Figure 1 we have provided the histogram and the fitted probability density functions of the Birnbaum-Saunders and log-normal distributions. The empirical survival function and the fitted survival functions are provided in Figure 2. It is very clear that both the models fit the data quite well and their probability density functions and the survival functions are quite close to each other.

Now let us look at the fitted hazard functions of both the distributions, and they have been plotted in Figure 3. It is known that the hazard functions of both the log-normal and Birnbaum-Saunders distributions are unimodal. In fact, there are some disease where the mortality reaches a peak after some finite time and then decline slowly. In such a situation,

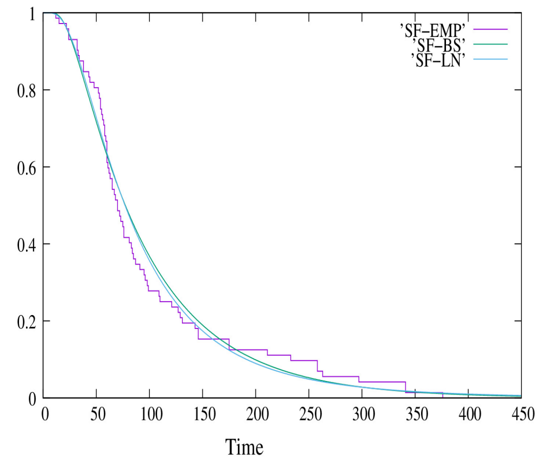

Figure 2: The survival functions of the fitted Birnbaum-Saunders and Log-normal distributions and the empirical survival function of the Data Set 1.

a distribution with an unimodal hazard function can be used quite effectively. Moreover, estimating the peak of the hazard function is an important issues in such a situation, see for example Gupta, Kannan and Raychaudhuri [16]. In case of Birnbaum-Saunders distribution, the peak of the hazard function is at 88.2, where as in case of log-normal distribution it is at 92.7. Further, if we compare the 99-th percentile points of the fitted Birnbaum-Saunders and log-normal distributions, they are 384.5 and 403.9, respectively. Hence, it is important to identify the *correct* model.

Data Set 2: It was originally provided by Birnbaum-Saunders [12] on the fatigue life of 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second (cps). The data set consists of 101 observations with maximum stress cycle 31,000 psi. The data are presented below:

70 90 96 97 99 100 103 104 104 105 107 108 108 108 109 109 112 112 113 114 114 114 116 119 120 120 120 121 121 123 124 124 124 124 124 128 128 129 129 130 130 130 131 131 131 131 131 132 132 132 133 134 134 134 134 134 136 136 137 138 138 138 139 139 141 141 142 142 142 142 142 142 144 144 145 146 148 148 149 151 151 152 155 156 157 157 157 157

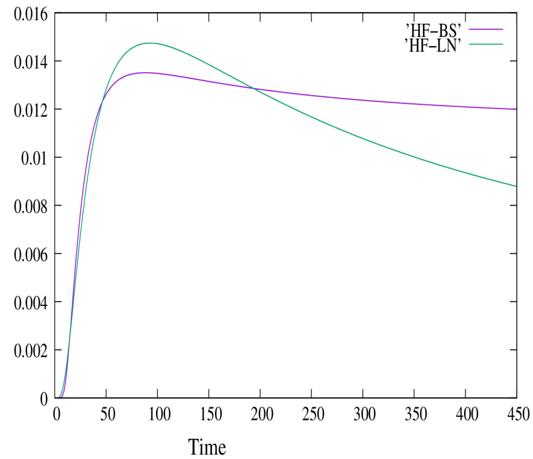

Figure 3: The hazard functions of the fitted Birnbaum-Saunders and Log-normal distributions of the Data Set 1.

158 159 162 163 163 164 166 166 168 170 174 196 212

Birnbaum and Saunders [12] and later Ng, Kundu and Balakrishnan [18] analyzed this data set based on the assumption that the data are coming from a two-parameter Birnbaum-Saunders distribution. The maximum likelihood estimates of α and β of the Birnbaum-Saunders distribution, as in (2), are $\widehat{\alpha} = 0.1704$ and $\widehat{\beta} = 131.8188$. Based on the assumption that the data are coming from a log-normal distribution, the maximum likelihood estimates of θ and σ of the log-normal parameters, as in (1), are $\widehat{\theta} = 131.8629$ and $\widehat{\sigma} = 0.1695$. We have plotted the probability density functions of the fitted Birnbaum-Saunders and log-normal distributions in Figure 4 and the corresponding survival functions in Figure 5. From Figures 4 and 5 it is clear that the two fitted probability density functions and the survival functions completely overlap with each other. This example shows that the two distributions behave very similarly. Hence, it is very important to study the effect of model misspecification on estimating the properties of a product's lifetime distribution if we fit a wrong model to a given data set.

Let us look at some of the other characteristics of the two fitted distributions. The

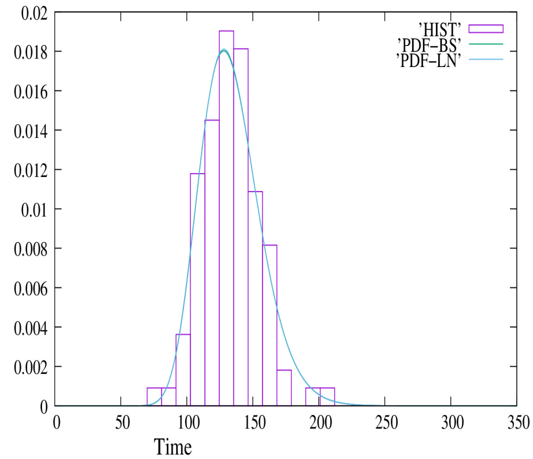

Figure 4: The probability density functions of the fitted Birnbaum-Saunders and Log-normal distributions and the histogram of the Data Set 2.

mean, variance, coefficient of variation (CV) , skewness (CS), kurtosis and 99-th percentile points for Birnbaum-Saunders (BS) and log-normal (LN) distributions are provided in Table 1. The percentile points of the fitted Birnbaum-Saunders and Log-normal distributions of the Data Set 2 are provided in Figure 6. As expected all the characteristics of the two fitted distributions are very close to each other. Although, the same cannot be said in case of the two hazard functions. In this case it is observed that the mode of the fitted hazard function for the log-normal distribution is at 298.1111, where as in case of Birnbaum-Saunders distribution the mode of the fitted distribution is greater than 2000.00. In this case it is numerically challenging to estimate the exact mode of the hazard function. This clearly indicates the effect of the model misspefication and how important it is to choose the *correct* model.

Table 1: Different characteristics of the two fitted distributions to Data Set 2.

Model	Mean	Variance	CV	Skewness	Kurtosis	99-th percentile
BS	133.7326	522.8503	0.1710	0.5104	3.4329	195.4
LN	133.7708	521.5737	0.1707	0.5172	3.3608	195.6

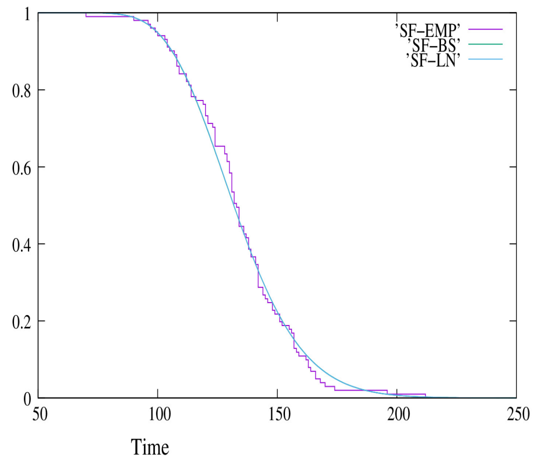

Figure 5: The survival functions of the fitted Birnbaum-Saunders and Log-normal distributions of Data Set 2.

3 Problem Formulation and Model Choice

In this section first we state the problem and then provide the model misspecification analysis. It should be emphasized that we treat this problem as a model misspecification problem not as a testing of hypothesis problem. The problem can be stated as follows: Suppose we have a random sample of size n say $\mathcal{D} = \{x_1, \ldots, x_n\}$ and it is known that it is coming from a log-normal or from a Birnbaum-Saunders distribution. We want to decide which one is the parent distribution and what is the corresponding probability of correct selection. Note that it is a different problem than the following testing problems namely

$$H_0 : \text{The data is from LN} \quad vs. \quad H_1 : \text{The data is from BS}$$

or

$$H_0 : \text{The data is from BS} \quad vs. \quad H_1 : \text{The data is from LN}.$$

The standard method in this case is to use ratio of the maximized likelihood approach. Choose that distribution which has higher maximized likelihood value. It has been used quite extensively in the literature, see for example Bain and Englehardt [19], Fearn and

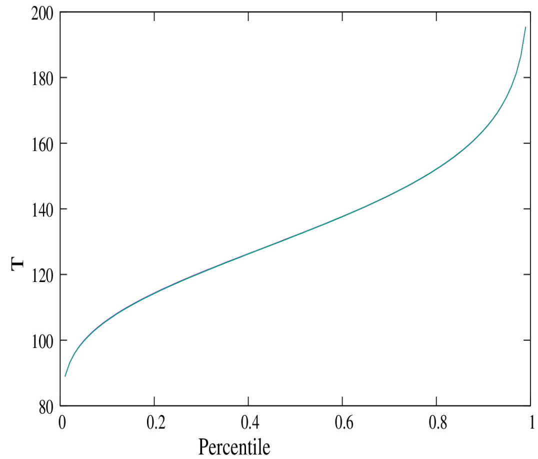

Figure 6: The percentile points of the fitted Birnbaum-Saunders and Log-normal distributions of Data Set 2.

Nebenzahl [20], Kundu and Manglick [5], the recent monograph by Pereira and Pereira [1] and see the references cited therein. It has been used by Marshall, Meza and Olkin [21] and Dey and Kundu [22] also, to choose the *best* among the three possible models. The main reason about the popularity of this method is that it is intuitively quite appealing and it has theoretical support. The method can be described as follows.

A Birnbaum-Saunders distribution with parameters α and β as defined in (2), will be denoted by BS(α, β), and a log-normal distribution with parameters θ and σ as defined in (1) will be denoted by LN(σ, θ). The corresponding likelihood function will be denoted by

$$L_{BS}(\alpha, \beta) = \prod_{i=1}^{n} f_{BS}(x_i; \alpha, \beta) \quad \text{and} \quad L_{LN}(\sigma, \theta) = \prod_{i=1}^{n} f_{LN}(x_i; \sigma, \theta),$$

respectively. The ratio of maximized likelihood is defined as

$$R = \frac{L_{BS}(\widehat{\alpha}, \widehat{\beta})}{L_{LN}(\widehat{\sigma}, \widehat{\theta})}. \tag{3}$$

Here $(\widehat{\alpha}, \widehat{\beta})$ and $(\widehat{\sigma}, \widehat{\theta})$ are the maximum likelihood estimators of (α, β) and (σ, θ), respectively. The maximum likelihood estimators of σ and θ can be obtained in explicit forms, whereas for (α, β), they have to be obtained by solving a non-linear equation. The details can be obtained in Balakrishnan and Kundu [23]. Now choose the Birnbaum-Saunders distribution as the parent distribution if

$$R > 1 \quad \Leftrightarrow \quad T = \ln R > 0, \tag{4}$$

and the log-normal distribution, otherwise.

To understand the effect of misspecification, it is important to study the distribution of T. Due to its complicated nature it is difficult to obtain the exact distribution of T, hence, we will study its asymptotic distribution. Let us introduce the following notations. If U is distributed as $\text{BS}(\alpha, \beta)$ ($\text{LN}(\theta, \sigma)$), then for any (measurable) function $g(\cdot)$, $E_{BS}(g(U)$ ($E_{LN}(g(U))$ and $V_{BS}(g(U)$ ($V_{LN}(g(U))$), denote the mean and variance of $g(U)$, if they exist. Similarly, the covariance of $g(U)$ and $h(U)$ is also defined, and it should be clear from the context. The cumulative distribution function, $F_{BS}(t)$, of $\text{BS}(\alpha, \beta)$ will be

$$F_{BS}(t; \alpha, \beta) = \Phi \left[\frac{1}{\alpha} \left\{ \left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right\} \right]; \quad 0 < t < \infty, \tag{5}$$

where $\Phi(t)$ is the cumulative distribution function of a standard normal distribution. Similarly, the cumulative distribution function of a $\text{LN}(\sigma, \theta)$ will be

$$F_{LN}(t; \sigma, \theta) = \Phi \left(\frac{\ln t - \ln \theta}{\sigma} \right); \quad 0 < t < \infty. \tag{6}$$

Along the same line the hazard function of $\text{BS}(\alpha, \beta)$ will be

$$h_{BS}(t; \alpha, \beta) = \frac{f_{BS}(t; \alpha, \beta)}{1 - F_{BS}(t; \alpha, \beta)}; \quad 0 < t < \infty, \tag{7}$$

and

$$h_{LN}(t; \sigma, \theta) = \frac{f_{LN}(t; \sigma, \theta)}{1 - F_{LN}(t; \sigma, \theta)}; \quad 0 < t < \infty. \tag{8}$$

It is known that both $h_{BS}(t; \alpha, \beta)$ and $h_{LN}(t; \sigma, \theta)$ are unimodal functions in $0 < t < \infty$, for all parameter values. The mode of the hazard function of the $\text{BS}(\alpha, \beta)$ and $\text{LN}(\sigma, \theta)$ cannot be obtained explicitly. They have to be obtained by solving $h'_{BS}(t; \alpha, \beta) = 0$ and $h'_{LN}(t; \sigma, \theta) =$

0, respectively, and both are unique. From now on the mode of the hazard function of the BS(α,β) and LN(σ,θ) will be denoted by $m_{BS}(\alpha,\beta)$ and $m_{LN}(\sigma,\theta)$, respectively. It can be easily seen that $m_{BS}(\alpha,\beta) = \beta m_{BS}(\alpha)$ and $m_{LN}(\sigma,\theta) = \theta m_{LN}(\sigma)$, where $m_{BS}(\alpha) = m_{BS}(\alpha,1)$ and $m_{LN}(\sigma) = m_{LN}(\sigma,1)$.

The p-th percentile point of a BS(α,β) is

$$\gamma_p(\alpha,\beta) = \frac{\beta}{4}\left[\alpha z_p + \sqrt{(\alpha z_p)^2+4}\right]^2 = \beta u_p(\alpha) \tag{9}$$

and for a LN(σ,θ) it is

$$\delta_p(\sigma,\theta) = \theta e^{\sigma z_p} = \theta v_p(\sigma). \tag{10}$$

Here, $z_p = \Phi^{-1}(1-p)$ is the p-th percentile point of a standard normal distribution. Further, $u_p(\alpha)$ and $v_p(\sigma)$ are the p-th percentile points of BS$(\alpha,1)$ and LN$(\sigma,1)$, respectively.

4 Effect of Misspecification

4.1 Data From Birnbaum-Saunders Distribution

In this section we study the effect of model misspefication when the data are actually from a Birnbaum-Saunders distribution, but it is wrongly assumed that it is coming from a log-normal distribution. Since both the distributions have similar probability density functions and the hazard functions, it is not very surprising that one may fit a wrong model. In reliability or survival analysis it is quite important to estimate percentile points. Hence, it is quite relevant to study the effect of misspecification in estimating the percentile points. Moreover, both these distributions have unimodal hazard functions. In an unimodal hazard function the time of the peak failure rate is of prime interest, see for example Greenwich [24]. In this section we study the effect of misspecification in estimating the mode of the hazard function. To study the effect of misspefication, we need to study the behavior of T.

First we will provide the asymptotic distribution of T under the assumption that the data are from BS(α, β), and then we will provide the effect of misspecification. We need the following lemma to provide the asymptotic distribution of T.

Lemma 1: Under the assumption that the data originated from $BS(\alpha, \beta)$, as $n \to \infty$,

(i) $\widehat{\alpha} \to \alpha$ a.s., $\widehat{\beta} \to \beta$ a.s., where

$$E_{BS}\left[\ln\left(f_{BS}(X;\alpha,\beta)\right)\right] = \max_{\bar{\alpha},\bar{\beta}} E_{BS}\left[\ln\left(f_{BS}(X;\bar{\alpha},\bar{\beta})\right)\right]$$

(ii) $\widehat{\sigma} \to \widetilde{\sigma}$ a.s. $\widehat{\theta} \to \widetilde{\theta}$ where

$$E_{BS}\left[\ln\left(f_{LN}(X;\widetilde{\sigma},\widetilde{\theta})\right)\right] = \max_{\sigma,\theta} E_{BS}\left[\ln\left(f_{LN}(X;\sigma,\theta)\right)\right]$$

Note that $\widetilde{\sigma}$ and $\widetilde{\theta}$ depend on α and β. The explicit expressions of $\widetilde{\theta}$ and $\widetilde{\sigma}$ are provided in (11). Let us denote,

$$T^* = \ln\left(\frac{L_{BS}(\alpha,\beta)}{L_{LN}(\widetilde{\sigma},\widetilde{\theta})}\right)$$

(iii) $n^{-1/2}\left[T - E_{BS}(T)\right]$ is asymptotically equivalent to $n^{-1/2}\left[T^* - E_{BS}(T^*)\right]$

Proof: The proof of this theorem can be obtained following the similar arguments as in White [14]. ∎

The explicit expressions of $\widetilde{\theta}$ and $\widetilde{\sigma}$ are as follows:

$$\widetilde{\theta} = \beta e^{E[\ln Y]} = \beta, \quad \text{and} \quad \widetilde{\sigma}^2 = V[\ln Y] = \frac{2\sqrt{2}}{\alpha\sqrt{2\pi}}\int_0^{\infty} y^2 \cosh(y) e^{-\frac{2}{\alpha^2}\sinh^2(y)}dy, \tag{11}$$

where $Y \sim$ BS$(\alpha,1)$. The proof of (11) is provided in the Appendix A.

The following theorem, stated with the help of the above lemma, will help us to derive the asymptotic distribution of T statistic.

Table 2: $\widetilde{\sigma}$, $AM_{BS}(\alpha)$ and $AV_{BS}(\alpha)$ for different α.

α	0.50	0.60	0.70	0.80	0.90	1.00	1.25	1.50	1.75	2.00
$AM_{BS}(\alpha)$	0.0009	0.0017	0.0027	0.0041	0.0058	0.0077	0.0138	0.0211	0.0293	0.0382
$AV_{BS}(\alpha)$	0.0016	0.0028	0.0044	0.0064	0.0088	0.3390	0.4685	0.6014	0.7363	0.8722
$\widetilde{\sigma}$	0.4863	0.5773	0.6657	0.7514	0.8344	0.9147	1.1044	1.2793	1.4411	1.5914

Theorem 1: When the data are from $BS(\alpha, \beta)$, for large n, the distribution of T is approximately normally distributed with mean $E_{BS}(T)$ and variance $V_{BS}(T^*) = V_{BS}(T)$.

Proof: Using central limit theorem and part (ii) of Lemma 1, it can be shown that $n^{-1/2}\left[T^* - E_{BS}(T^*)\right]$ is asymptotically normally distributed with mean 0 and variance $V_{BS}(T^*)$. Furthermore, applying part (iii) of Lemma 1, the result follows. ∎

Note that $\lim_{n\to\infty} \frac{E_{BS}(T)}{n}$ and $\lim_{n\to\infty} \frac{V_{BS}(T)}{n}$ exist. Let us denote

$$AM_{BS}(\alpha) = \lim_{n\to\infty} \frac{E_{BS}(T)}{n} \quad \text{and} \quad AV_{BS}(\alpha) = \lim_{n\to\infty} \frac{V_{BS}(T)}{n}.$$

The expressions of $AM_{BS}(\alpha)$ and $AV_{BS}(\alpha)$ are provided in the Appendix B. In the following Table 2 we provide $AM_{BS}(\alpha)$, $AV_{BS}(\alpha)$ and $\widetilde{\sigma}$, for different values of α. It will be used later.

The following result which was originally developed by Englehardt, Bain and Wright [25] will be useful for further development.

Theorem 2: If the data are from BS(α, β), then the asymptotic joint distribution of $\widehat{\alpha}$ and $\widehat{\beta}$ is bivariate normal and it is given by

$$\begin{pmatrix} \sqrt{n}(\widehat{\alpha} - \alpha) \\ \sqrt{n}(\widehat{\beta} - \beta) \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\alpha^2}{2} & 0 \\ 0 & \beta^2 v(\alpha) \end{pmatrix}\right],$$

where $v(\alpha) = (0.25 + \alpha^{-2} + I(\alpha))^{-1}$,

$$I(\alpha) = 2\int_0^\infty \left\{[1 + g(\alpha x)]^{-1} - 0.5\right\}^2 d\Phi(x), \quad \text{and}$$

$$g(y) = 1 + \frac{y^2}{2} + y\left(1 + \frac{y^2}{4}\right)^{1/2}.$$

Based on Theorem 1, we have the following corollaries:

Corollary 1: If the data are from BS(α, β), then the maximum likelihood estimator of the p-th percentile point of the Birnbaum-Saunders distribution, $\gamma_p(\alpha, \beta)$, is $\gamma_p(\widehat{\alpha}, \widehat{\beta})$, and the asymptotic distribution of $\gamma_p(\widehat{\alpha}, \widehat{\beta})$ is

$$\sqrt{n}(\gamma_p(\widehat{\alpha}, \widehat{\beta}) - \gamma_p(\alpha, \beta)) \sim N\left(0, \frac{\alpha^2}{2}(u_p'(\alpha)\beta)^2 + \beta^2 v(\alpha)(u_p(\alpha))^2\right).$$

Corollary 2: If the data are from BS(α, β), then the maximum likelihood estimator of the mode of the hazard function of the Birnbaum-Saunders distribution, $m_{BS}(\alpha, \beta)$, is $m_{BS}(\widehat{\alpha}, \widehat{\beta})$, and the asymptotic distribution of $m_{BS}(\widehat{\alpha}, \widehat{\beta})$ is

$$\sqrt{n}(m_{BS}(\widehat{\alpha}, \widehat{\beta}) - m_{BS}(\alpha, \beta)) \sim N\left(0, \frac{\alpha^2}{2}\left(\frac{\ddot{h}_{t\alpha}(m_{BS}(\alpha); \alpha, \beta)}{\ddot{h}_{tt}(m_{BS}(\alpha); \alpha, \beta)}\right)^2 + \beta^2 v(\alpha)\left(\frac{\ddot{h}_{t\beta}(m_{BS}(\alpha); \alpha, \beta)}{\ddot{h}_{tt}(m_{BS}(\alpha); \alpha, \beta)}\right)^2\right).$$

Here

$$\ddot{h}_{t\alpha}(m_{BS}(\alpha); \alpha, \beta) = \frac{\partial^2}{\partial t \partial \alpha} h_{BS}(t; \alpha, \beta)\Big|_{t=m_{BS}(\alpha)}, \quad \ddot{h}_{t\beta}(m_{BS}(\alpha); \alpha, \beta) = \frac{\partial^2}{\partial t \partial \beta} h_{BS}(t; \alpha, \beta)\Big|_{t=m_{BS}(\alpha)}$$

and

$$\ddot{h}_{tt}(m_{BS}(\alpha); \alpha, \beta) = \frac{\partial^2}{\partial t^2} h_{BS}(t; \alpha, \beta)\Big|_{t=m_{BS}(\alpha)}.$$

The following theorem will be useful to study the effect of misspefication, if the data are from BS(α, β), and the LN(σ, θ) is fitted to the data set.

Theorem 3: If the data are from BS(α, β), then the asymptotic joint distribution of $\widehat{\sigma}$ and $\widehat{\theta}$ is bivariate normal and it is given by

$$\begin{pmatrix} \sqrt{n}(\widehat{\sigma} - \widetilde{\sigma}) \\ \sqrt{n}(\widehat{\theta} - \widetilde{\theta}) \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\alpha^2}{2}(h'(\alpha))^2 & 0 \\ 0 & \beta^2 v(\alpha) \end{pmatrix}\right],$$

where $h(\alpha) = \widetilde{\sigma}$, see (11).

Proof: The proof can be easily obtained using the multivariate delta method. ∎

Corollary 3: If the data are from a BS(α, β), and wrongly a LN(σ, θ) has been fitted to that data set, then the maximum likelihood estimator of the p-th percentile of the fitted distribution is $\delta_p(\widehat{\sigma}, \widehat{\theta})$, and its asymptotic distribution is the following:

$$\sqrt{n}(\delta_p(\widehat{\sigma}, \widehat{\theta}) - \delta_p(\widetilde{\sigma}, \widetilde{\theta})) \sim N\left(0, \frac{\alpha^2}{2}(h'(\alpha))^2(\widetilde{\theta} v_P'(\widetilde{\sigma}))^2 + \beta^2 v(\alpha)(v_p(\widetilde{\sigma}))^2\right).$$

Corollary 4: If the data are from a BS(α, β), and wrongly a LN(σ, θ) has been fitted to that data set, then the maximum likelihood estimator of the mode of the hazard function of the fitted distribution is $m_{LN}(\widehat{\sigma}, \widehat{\theta})$, and its asymptotic distribution is the following:

$$\sqrt{n}(m_{LN}(\widehat{\sigma}, \widehat{\theta}) - m_{LN}(\widetilde{\sigma}, \widetilde{\theta})) \sim N\left(0, \frac{\alpha^2}{2}(h'(\alpha))^2\left(\frac{\ddot{h}_{t\alpha}(m_{LN}(\sigma); \sigma, \theta)}{\ddot{h}_{tt}(m_{LN}(\sigma); \sigma, \theta)}\right)^2 + \beta^2 v(\alpha)\left(\frac{\ddot{h}_{t\beta}(m_{LN}(\sigma); \sigma, \theta)}{\ddot{h}_{tt}(m_{LN}(\sigma); \sigma, \theta)}\right)^2\right).$$

Here

$$\ddot{h}_{t\alpha}(m_{LN}(\sigma); \sigma, \theta) = \frac{\partial^2}{\partial t \partial \sigma} h_{LN}(t; \sigma, \theta)\bigg|_{t=m_{LN}(\sigma)}, \quad \ddot{h}_{t\beta}(m_{LN}(\sigma); \sigma, \theta) = \frac{\partial^2}{\partial t \partial \theta} h_{LN}(t; \sigma, \theta)\bigg|_{t=m_{LNS}(\sigma)}$$

and

$$\ddot{h}_{tt}(m_{LN}(\sigma); \sigma, \theta) = \frac{\partial^2}{\partial t^2} h_{LN}(t; \sigma, \theta)\bigg|_{t=m_{LN}(\sigma)}.$$

We study the effect of model misspecification for different sample sizes and for different parameter values, when the data are coming from a Birnbaum-Saunders distribution, and wrongly a log-normal distribution has been fitted. All the results are based on 5000 replications. The average Birnbaum-Saunders percentile estimates, the associated mean squared errors, the average misspecified log-normal percentile estimates and the associated mean squared errors when the data are coming from a Birnbaum-Saunders are presented in Table 3. The average estimates of the mode of the Birnbaum-Saunders hazard function, the associated mean squared errors, the average estimates of the misspecified mode of the log-normal hazard function, the associated mean squared errors when the data are coming from a Birnbaum-Saunders are presented in Table 4. It is very clear from Tables 3 and 4 that the model misspecification has a more severe effect in estimating the mode of the hazard function than estimating the percentile points when the data are coming from a Birnbaum-Saunders distribution, particularly when the shape parameter is small.

Table 3: The average Birnbaum-Saunders percentile estimates ($\widehat{u}_p$), the associated mean squared errors ($MS(\widehat{u}_p)$), the average misspecified log-normal percentile estimates ($\widehat{r}_p$) and the associated mean squared errors ($MS(\widehat{r}_p)$)when the data are coming from a $BS(\alpha, 1)$

α	p	n	$\widehat{u}_p$	$MS(\widehat{u}_p)$	$\widehat{r}_p$	$MS(\widehat{r}_p)$	α	p	n	$\widehat{u}_p$	$MS(\widehat{u}_p)$	$\widehat{r}_p$	$MS(\widehat{r}_p)$
0.25	0.900	25	1.3433	0.0084	1.3620	0.0084	0.50	0.900	25	1.8261	0.0585	1.8349	0.0568
		50	1.3539	0.0045	1.3681	0.0044			50	1.8676	0.0291	1.8545	0.0281
		75	1.3102	0.0028	1.3696	0.0028			75	1.8724	0.0198	1.8596	0.0192
		100	1.3281	0.0021	1.3714	0.0021			100	1.8732	0.0151	1.8604	0.0145
	0.990	25	1.7433	0.0261	1.7552	0.0271		0.990	25	2.9421	0.2724	3.0127	0.3174
		50	1.7205	0.0143	1.7681	0.0148			50	2.9812	0.1365	3.0549	0.1601
		75	1.7476	0.0092	1.7708	0.0095			75	2.9962	0.0913	3.0712	0.1058
		100	1.7272	0.0067	1.7733	0.0070			100	3.0034	0.0680	3.0791	0.0790
	0.999	25	2.0532	0.0575	2.1069	0.0638		0.999	25	4.0029	0.6687	4.3131	1.0257
		50	2.1013	0.0290	2.1301	0.0321			50	4.0796	0.3355	4.4100	0.5094
		75	2.1019	0.0196	2.1375	0.0216			75	4.0999	0.2194	4.4350	0.3320
		100	2.1033	0.0193	2.1385	0.0213			100	4.1113	0.1659	4.4488	0.2506
0.75	0.900	25	2.4787	0.2158	2.4361	0.2051	1	0.900	25	3.2783	0.5932	3.1846	0.5639
		50	2.5097	0.1088	2.4645	0.1036			50	3.3086	0.2938	3.2024	0.2754
		75	2.5097	0.0736	2.4638	0.0696			75	3.3223	0.2015	3.2149	0.1903
		100	2.5124	0.0552	2.4659	0.0521			100	3.3243	0.1550	3.2158	0.1454
	0.990	25	4.6698	1.2132	5.0000	1.7442		0.990	25	7.0556	4.0476	7.9177	7.5879
		50	4.7705	0.6552	5.1217	0.9287			50	7.1795	2.0175	8.1261	3.6682
		75	4.7906	0.4435	5.1436	0.6220			75	7.1783	0.9883	8.0676	1.7523
		100	4.7987	0.3097	5.1554	0.4361			100	7.2165	1.0300	8.1222	1.8384
	0.999	25	6.9292	2.1325	8.5875	4.7748		0.999	25	11.0342	7.3688	16.3977	27.2773
		50	7.0343	1.0506	8.6827	2.3425			50	11.1772	3.6286	16.4891	13.1596
		75	7.1097	0.7238	8.7946	1.6459			75	11.2102	2.5154	16.5309	9.2168
		100	7.1289	0.5470	8.8114	1.2498			100	11.2730	1.8325	16.5753	6.7907

4.2 Data From Log-Normal Distribution

In this section along the same line as the previous one, we provide the necessary results when the data come from a log-normal distribution. First we provide Lemma 2, along the same line as Lemma 1.

Lemma 2: If the data originally come from a $LN(\sigma, \theta)$ distribution, then as $n \to \infty$,

Table 4: The average estimates of the mode of the Birnbaum-Saunders hazard function, the associated mean squared errors, the average estimates of the misspecified mode of the log-normal hazard function, the associated mean squared errors when the data are coming from a $BS(\alpha, 1)$.

α	n	$\widehat{m}_{BS}$	MSE_{BS}	$\widehat{m}_{LN}$	MSE_{LN}
0.5	25	4.4357	1.0971	1.8553	7.3585
	50	4.5260	0.7791	1.8331	7.4474
	75	4.5784	0.6008	1.8224	7.4946
	100	4.5738	0.5301	1.8183	7.5125
0.75	25	1.6019	0.6594	1.2874	0.8914
	50	1.4127	0.3365	1.2540	0.9012
	75	1.3362	0.2097	1.2390	0.9021
	100	1.2964	0.1116	1.2229	0.9217
1	25	0.6630	0.1559	0.8445	0.4578
	50	0.5790	0.0415	0.8021	0.4689
	75	0.5594	0.0248	0.7921	0.4778
	100	0.5482	0.0168	0.7869	0.4882

(i) $\widehat{\sigma} \to \sigma$ a.s., $\widehat{\theta} \to \theta$ a.s., where

$$E_{LN}\left[\ln\left(f_{LN}(X;\sigma,\theta)\right)\right] = \max_{\bar{\sigma},\bar{\theta}} E_{LN}\left[\ln\left(f_{LN}(X;\bar{\sigma},\bar{\theta})\right)\right]$$

(ii) $\widehat{\alpha} \to \tilde{\alpha}$ a.s. $\widehat{\beta} \to \tilde{\beta}$ a.s. where

$$E_{LN}\left[\ln\left(f_{BS}(X;\widetilde{\alpha},\widetilde{\beta})\right)\right] = \max_{\alpha,\beta} E_{LN}\left[\ln\left(f_{BS}(X;\alpha,\beta)\right)\right]$$

Note that $\widetilde{\alpha}$ and $\widetilde{\beta}$ may depend on σ and β, but their explicit forms are not discussed here.

Let us denote,

$$T^* = \ln\left(\frac{L_{BS}(\widetilde{\alpha},\widetilde{\beta})}{L_{LN}(\sigma,\theta)}\right)$$

(iii) $n^{-1/2}\left[T - E_{LN}(T)\right]$ is asymptotically equivalent to $n^{-1/2}\left[T^* - E_{LN}(T^*)\right]$

Proof: The proof can be obtained similar to the proof of Lemma 1. ∎

Table 5: $\widetilde{\alpha}$, $AM_{LN}(\sigma)$ and $AV_{LN}(\sigma)$ for different σ.

σ	0.50	0.60	0.70	0.80	1.00	1.25	1.50	1.75	2.00
$AM_{LN}(\sigma)$	-0.0012	-0.0025	-0.0045	-0.0075	-0.0173	-0.0392	-0.0754	-0.1299	-0.2062
$AV_{LN}(\sigma)$	0.0031	0.0069	0.0139	0.0260	0.0782	0.2605	0.7838	2.2660	6.6019
$\widetilde{\alpha}$	0.5160	0.6280	0.7451	0.8685	1.1391	1.5390	2.0397	2.6922	3.5746

The explicit expressions of $\widetilde{\alpha}$ and $\widetilde{\beta}$ are as follows:

$$\widetilde{\beta} = \theta \quad \text{and} \quad \widetilde{\alpha}^2 = 2(e^{\frac{\sigma^2}{2}} - 1). \tag{12}$$

Theorem 4: When the underlying data are from a $LN(\sigma, \theta)$, then the asymptotic distribution of T is normally distributed with mean $E_{LN}(T)$ and variance $V_{LN}(T^*) = V_{LN}(T)$.

Proof: Using Lemma 2, the proof can be established along the same line as the proof of Theorem 1. ∎

In this case also $\lim_{n\to\infty} \frac{E_{LN}(T)}{n}$ and $\lim_{n\to\infty} \frac{V_{LN}(T)}{n}$ exist, and we will denote them as $AM_{LN}(\alpha)$ and $AV_{LN}(\alpha)$, respectively. The explicit expressions of $AM_{LN}(\alpha)$ and $AV_{LN}(\alpha)$ are provided in the Appendix C. In the following Table 5 we provide $AM_{LN}(\sigma)$, $AV_{LN}(\sigma)$ and $\widetilde{\alpha}$, for different values of σ. It will be used later.

We have the following well known results. The proof is available in many places, see for example Alshunnar et al. [26].

Theorem 5: If the data are from a LN(σ, θ), then the asymptotic joint distribution of $\widehat{\sigma}$ and $\widehat{\theta}$ is given by

$$\begin{pmatrix} \sqrt{n}(\widehat{\sigma} - \sigma) \\ \sqrt{n}(\widehat{\theta} - \theta) \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sigma^2}{2} & 0 \\ 0 & \sigma^2\theta^2 \end{pmatrix} \right],$$

Corollary 5: If the data are from LN(σ, θ), then the maximum likelihood estimator of the p-th percentile point of the log-normal distribution, $\delta_p(\sigma, \theta)$, is $\delta_p(\widehat{\sigma}, \widehat{\theta})$, and the asymptotic distribution of $\delta_p(\widehat{\sigma}, \widehat{\theta})$ is

$$\sqrt{n}(\delta_p(\widehat{\sigma}, \widehat{\theta}) - \delta_p(\sigma, \theta)) \sim N\left(0, \frac{\sigma^2}{2}(r_p'(\sigma)\theta)^2 + \sigma^2\theta^2(r_p(\sigma))^2\right).$$

Corollary 6: If the data are from a LN(σ, θ), then the maximum likelihood estimator of the mode of the hazard function of the log-normal distribution, $m_{LN}(\sigma, \theta)$, is $m_{LN}(\widehat{\sigma}, \widehat{\theta})$, and the asymptotic distribution of $m_{LN}(\widehat{\sigma}, \widehat{\theta})$ is

$$\sqrt{n}(m_{LN}(\widehat{\sigma}, \widehat{\theta}) - m_{LN}(\sigma, \theta)) \sim N\left(0, \frac{\sigma^2}{2}\left(\frac{\ddot{h}_{t\sigma}(m_{LN}(\sigma); \sigma, \theta)}{\ddot{h}_{tt}(m_{LN}(\sigma); \sigma, \theta)}\right)^2 + \sigma^2\theta^2\left(\frac{\ddot{h}_{t\theta}(m_{LN}(\sigma); \sigma, \theta)}{\ddot{h}_{tt}(m_{LN}(\sigma); \sigma, \theta)}\right)^2\right).$$

Here

$$\ddot{h}_{t\sigma}(m_{LN}(\sigma); \sigma, \theta) = \frac{\partial^2}{\partial t \partial \sigma} h_{LN}(t; \sigma, \theta)\Big|_{t=m_{LN}(\sigma)}, \quad \ddot{h}_{t\theta}(m_{LN}(\sigma); \sigma, \theta) = \frac{\partial^2}{\partial t \partial \theta} h_{LN}(t; \sigma, \theta)\Big|_{t=m_{LN}(\sigma)}$$

and

$$\ddot{h}_{tt}(m_{LN}(\sigma); \sigma, \theta) = \frac{\partial^2}{\partial t^2} h_{LN}(t; \sigma, \theta)\Big|_{t=m_{LN}(\sigma)}.$$

Theorem 6: If the data are from a LN(σ, θ), then the asymptotic joint distribution of $\widehat{\alpha}$ and $\widehat{\beta}$ is given by

$$\begin{pmatrix} \sqrt{n}(\widehat{\alpha} - \widetilde{\alpha}) \\ \sqrt{n}(\widehat{\beta} - \widetilde{\beta}) \end{pmatrix} \sim N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{e^{\sigma^2}\sigma^4}{(e^{\frac{\sigma^2}{2}}-1)} & 0 \\ 0 & \sigma^2\theta^2 \end{pmatrix}\right],$$

Corollary 7: If the data are from a LN(σ, θ), and wrongly a BS(α, β) has been fitted to that data set, then the maximum likelihood estimator of the p-th percentile of the fitted distribution is $\gamma_p(\widehat{\alpha}, \widehat{\beta})$, and its asymptotic distribution is the following:

$$\sqrt{n}(\gamma_p(\widehat{\alpha}, \widehat{\beta}) - \gamma_p(\widetilde{\alpha}, \widetilde{\beta})) \sim N\left(\frac{e^{\sigma^2}\sigma^4}{(e^{\frac{\sigma^2}{2}} - 1)}(\theta v_p'(\sigma))^2 + \sigma^2\theta^2(v_p(\sigma))^2\right).$$

Corollary 8: If the data are from a LN(σ, θ), and wrongly a BS(α, β) has been fitted to that data set, then the maximum likelihood estimator of the mode of the hazard function of the fitted distribution is $m_{BS}(\widehat{\alpha}, \widehat{\beta})$, and its asymptotic distribution is the following:

$$\sqrt{n}(m_{BS}(\widehat{\alpha}, \widehat{\beta}) - m_{BS}(\widetilde{\alpha}, \widetilde{\beta})) \sim N\left(\frac{e^{\sigma^2}\sigma^4}{(e^{\frac{\sigma^2}{2}} - 1)}\left(\frac{\ddot{h}_{t\sigma}(m_{BS}(\alpha); \alpha, \beta)}{\ddot{h}_{tt}(m_{BS}(\alpha); \alpha, \beta)}\right)^2 + \sigma^2\theta^2\left(\frac{\ddot{h}_{t\theta}(m_{BS}(\alpha); \alpha, \beta)}{\ddot{h}_{tt}(m_{BS}(\alpha); \alpha, \beta)}\right)^2\right).$$

Here

$$\ddot{h}_{t\sigma}(m_{BS}(\alpha);\alpha,\beta) = \frac{\partial^2}{\partial t \partial \sigma} h_{BS}(t;\alpha,\beta)\Big|_{t=m_{BS}(\alpha)}, \quad \ddot{h}_{t\theta}(m_{BS}(\alpha);\alpha,\beta) = \frac{\partial^2}{\partial t \partial \theta} h_{BS}(t;\alpha,\beta)\Big|_{t=m_{BS}(\alpha)}$$

and

$$\ddot{h}_{tt}(m_{BS}(\alpha);\alpha,\beta) = \frac{\partial^2}{\partial t^2} h_{BS}(t;\alpha,\beta)\Big|_{t=m_{BS}(\alpha)}.$$

The average log-normal percentile estimates, the associated mean squared errors, the average misspecified Birnbaum-Saunders percentile estimates and the associated mean squared errors when the data are coming from a log-normal distribution are presented in Table 6. The average estimates of the mode of the log-normal hazard function, the associated mean squared errors, the average estimates of the misspecified mode of the Birnbaum-Saunders hazard function, the associated mean squared errors when the data are coming from a log-normal distribution are presented in Table 7. In this case also it is very clear that the model misspecification has a more severe effect in estimating the mode of the hazard function than estimating the percentile points when the data are coming from a log-normal distribution, particularly when the shape parameter is small.

5 Simulation Results

It is quite clear that the effect of misspefication will be less if the probability of correct selection of the model is more. It has already been observed in the previous section that asymptotically the probability of correct selection tends to one. It is expected that if the sample size is large, say for example 100 or 150, the probability of correct selection should be close to one, as it has been observed in case Weibull and log-normal distribution, see for example Kundu and Manglick [5]. Moreover, the probability of correct selection increases as the sample size increases. To verify the effect of sample size and the parameter values on the probability of correct selection, we have performed some simulations experiments

Table 6: The average log-normal percentile estimates ($\widehat{r}_p$), the associated mean squared errors ($MS(\widehat{r}_p)$), the average misspecified Birnbaum-Saunders percentile estimates ($\widehat{u}_p$) and the associated mean squared errors ($MS(\widehat{u}_p)$) when the data are coming from a $LN(\sigma, 1)$

σ	p	n	$\widehat{u}_p$	$MS(\widehat{u}_p)$	$\widehat{r}_p$	$MS(\widehat{r}_p)$	σ	p	n	$\widehat{u}_p$	$MS(\widehat{u}_p)$	$\widehat{r}_p$	$MS(\widehat{r}_p)$
0.25	0.900	25	1.3682	0.0086	1.3670	0.0086	0.50	0.900	25	1.8913	0.0685	1.8778	0.0650
		50	1.3738	0.0043	1.3724	0.0042			50	1.9096	0.0347	1.8939	0.0327
		75	1.3761	0.0030	1.3747	0.0030			75	1.9072	0.0234	1.8911	0.0220
		100	1.3757	0.0021	1.3742	0.0021			100	1.9107	0.0175	1.8944	0.0165
	0.990	25	1.7622	0.0286	1.7689	0.0296		0.990	25	3.0635	0.3316	3.1530	0.3875
		50	1.7694	0.0143	1.7759	0.0148			50	3.0955	0.1642	3.1794	0.1882
		75	1.7741	0.0099	1.7804	0.0102			75	3.0995	0.1082	3.1821	0.1241
		100	1.7788	0.0074	1.7853	0.0076			100	3.1038	0.0810	3.1850	0.0926
	0.999	25	2.1077	0.0604	2.1351	0.0672		0.999	25	4.1973	0.8407	4.5911	1.3147
		50	2.1233	0.0303	2.1505	0.0336			50	4.2733	0.4370	4.6633	0.6726
		75	2.1276	0.0204	2.1544	0.0226			75	4.2868	0.2840	4.6742	0.4333
		100	2.1341	0.0151	2.1612	0.0168			100	4.2809	0.2142	4.6634	0.3248
0.75	0.900	25	2.6727	0.3497	2.6047	0.2911	1	0.900	25	3.7946	1.5315	3.5688	0.9555
		50	2.6680	0.1618	2.5965	0.1357			50	3.8709	0.8147	3.6138	0.4945
		75	2.6865	0.1151	2.6107	0.0950			75	3.8626	0.4806	3.6055	0.3093
		100	2.6861	0.0848	2.6095	0.0703			100	3.8630	0.3775	3.6051	0.2405
	0.990	25	5.2470	2.1111	5.7011	2.9281		0.990	25	8.7570	12.0650	10.3099	17.3127
		50	5.2456	1.0491	5.6744	1.4008			50	8.8073	5.5262	10.2549	8.0153
		75	5.2819	0.7184	5.6987	0.9401			75	8.8250	3.7803	10.2199	5.2913
		100	5.3017	0.5340	5.7145	0.6963			100	8.8584	2.8349	10.2506	4.0806
	0.999	25	7.6262	3.6600	9.6760	8.2640		0.999	25	13.4479	18.1942	21.1481	65.3189
		50	7.8420	1.7384	9.8856	3.9483			50	13.7905	9.9792	21.4447	34.7519
		75	7.9121	1.1767	9.9735	2.6378			75	13.9883	6.8829	21.6539	23.2649
		100	7.9598	0.9415	10.0223	2.1042			100	13.9932	4.9697	21.5552	16.4918

Table 7: The average estimates of the mode of the log-normal hazard function, the associated mean squared errors, the average estimates of the misspecified mode of the Birnbaum-Saunders hazard function, the associated mean squared errors when the data are coming from Mean square error of estimated modes of hazards under misspecification when $LN(\sigma, 1)$ is the underlying model

σ	n	$\widehat{m}_{BS}$	MSE_{BS}	$\widehat{m}_{LN}$	MSE_{LN}
0.50	25	5.3369	9.4446	1.8147	0.0787
	50	4.6780	2.9649	1.7892	0.0360
	75	4.4538	1.7423	1.7734	0.0229
	100	4.3727	1.1907	1.7709	0.0176
0.75	25	1.3265	0.8203	1.1805	0.1001
	50	1.1431	0.2739	1.1458	0.0504
	75	1.0799	0.1372	1.1358	0.0315
	100	1.0415	0.0891	1.1237	0.0231
1	25	0.5153	0.1354	0.7185	0.0934
	50	0.4263	0.0313	0.6637	0.0374
	75	0.4058	0.0189	0.6481	0.0244
	100	0.3955	0.0126	0.6424	0.0179

and we have presented the results here. From the expression of T it is immediate that the distribution of T is independent of β if the data come from a Birnbaum-Saunders distribution and it is independent of θ if it is from a log-normal distribution. If the data come from a Birnbaum-Saunders distribution, the probability of correct selection can be defined as $PCS_{BS}(\alpha) = P(T > 0|\alpha)$, similarly, if it is from a log-normal distribution, it can be defined as $PCS_{LN}(\sigma) = P(T < 0|\sigma)$. Hence, an estimate of probability of correct selection can be obtained easily obtained as follows. We generate a random sample of size n from a BS$(\alpha, 1)$, and we calculate T. We replicate the process say N times, compute the proportion of time $T > 0$ out of N replications and this becomes as estimate of $PCS_{BS}(\alpha)$. Similarly, as estimate of $PCS_{LN}(\sigma)$ also can be obtained. Based on the asymptotic distributions of T, for large n an approximate value of probability of correct selection can be obtained as

$$PCS_{BS}(\alpha) = P(T > 0|\alpha) \approx 1 - \Phi\left(\frac{-E_{BS}(T)}{\sqrt{V_{BS}(T)}}\right) = 1 - \Phi\left(\frac{-n \times AM_{BS}(\alpha)}{\sqrt{nAV_{BS}(\alpha)}}\right) \tag{13}$$

and

$$PCS_{LN}(\sigma) = P(T < 0|\sigma) \approx \Phi\left(\frac{-E_{LN}(T)}{\sqrt{V_{LN}(T)}}\right) = \Phi\left(\frac{-n \times AM_{LN}(\sigma)}{\sqrt{nAV_{LN}(\sigma)}}\right). \quad (14)$$

We obtain the estimated $PCS_{BS}(\alpha)$ and $PCS_{LN}(\sigma)$, for different values of α and σ and for different sample sizes. We have taken, small, moderate, large and very large sample sizes in both the cases. We have considered different parameter values also. All the results are based on 10,000 replications. Along with the estimated probability of correct selections, we have also reported the asymptotic probability of correct selections as it has been described above, in brackets below. In Table 8 we have reported the results when the parent distribution is Birnbaum-Saunders, and the corresponding results when the parent distribution is log-normal has been reported in Table 9. For any α and n the upper value indicates the proportion of correct selection based on 10,000 replications and the value within brackets indicates the theoretical probability of correct selection.

Table 8: Estimated probability of correct selections (given in the brackets) and the asymptotic probability of correct selections when the data are from a BS$(\alpha, 1)$.

$\alpha \downarrow n \rightarrow$	25	50	75	100	200	500	1000	10000
0.25	0.7194	0.6694	0.6537	0.6405	0.6170	0.6078	0.6154	0.7333
	(0.5124)	(0.5175)	(0.5215)	(0.5248)	(0.5350)	(0.5553)	(0.5779)	(0.7329)
0.50	0.7377	0.7015	0.6497	0.6914	0.6955	0.7336	0.7806	0.9783
	(0.5460)	(0.5649)	(0.5793)	(0.5913)	(0.6281)	(0.6973)	(0.7676)	(0.9896)
0.75	0.7597	0.7461	0.7502	0.7555	0.7835	0.8628	0.9234	1.0000
	(0.5916)	(0.6284)	(0.6559)	(0.6784)	(0.7438)	(0.8499)	(0.9286)	(1.0000)
1.00	0.7887	0.7859	0.8017	0.8164	0.8638	0.9470	0.9849	1.0000
	(0.6405)	(0.6946)	(0.7334)	(0.7641)	(0.8456)	(0.9462)	(0.9886)	(1.0000)
1.50	0.8383	0.8607	0.8889	0.9092	0.9577	0.9956	0.9998	1.0000
	(0.7287)	(0.8055)	(0.8542)	(0.8883)	(0.9575)	(0.9968)	(0.9999)	(1.0000)
2.00	0.8795	0.9128	0.9399	0.9586	0.9880	1.0000	1.0000	1.0000
	(0.7959)	(0.8789)	(0.9240)	(0.9509)	(0.9903)	(0.9999)	(1.0000)	(1.0000)
3.00	0.9255	0.9650	0.9811	0.9905	0.9993	1.0000	1.0000	1.0000
	(0.8795)	(0.9514)	(0.9789)	(0.9905)	(0.9995)	(1.0000)	(1.0000)	(1.0000)

Some of the points are quite clear from these simulation results. It is apparent that al-

Table 9: Estimated probability of correct selections (given in the brackets) and the asymptotic probability of correct selections when the data are from a LN(σ, 1).

$\sigma \downarrow n \rightarrow$	25	50	75	100	200	500	1000	10000
0.25	0.2962 (0.5123)	0.3594 (0.5173)	0.3829 (0.5213)	0.4039 (0.5246)	0.4487 (0.5348)	0.4956 (0.5548)	0.5397 (0.5773)	0.7330 (0.7313)
0.50	0.3209 (0.5439)	0.4052 (0.5620)	0.4345 (0.5758)	0.4682 (0.5873)	0.5456 (0.6224)	0.6537 (0.6890)	0.7399 (0.7572)	0.9947 (0.9863)
0.75	0.3606 (0.5839)	0.4661 (0.6178)	0.5158 (0.6432)	0.5675 (0.6641)	0.6806 (0.7255)	0.8285 (0.8283)	0.9224 (0.9099)	1.0000 (1.0000)
1.00	0.4083 (0.6214)	0.5428 (0.6690)	0.6159 (0.7038)	0.6813 (0.7318)	0.8098 (0.8091)	0.9494 (0.9166)	0.9924 (0.9747)	1.0000 (1.0000)
1.50	0.5083 (0.6650)	0.7049 (0.7266)	0.8017 (0.7698)	0.8716 (0.8030)	0.9647 (0.8860)	0.9996 (0.9717)	1.0000 (0.9965)	1.0000 (1.0000)
2.00	0.6156 (0.7353)	0.8333 (0.8241)	0.9213 (0.8851)	0.9604 (1.0000)	0.9981 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)
3.00	0.8131 (0.8604)	0.9639 (0.9953)	0.9941 (1.0000)	0.9980 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)	1.0000 (1.0000)

though the traditional likelihood ratio approach works quite well when the shape parameters are greater than one, but it does not work well when the shape parameters are less than one. In fact when the parent distribution is log-normal it is observed that the probability of correct selection is less than 0.5 in many cases. Clearly one should not use this method in this case as one can get a better result simply by tossing a fair coin. In fact when the shape parameters are less than one some counter intuitive results are also observed, i.e. even if the sample size increases, the probability of correct selection does not increase. It is quite apparent that in certain range of shape parameters, choosing the correct model is quite difficult in this case. It seems special attention is needed when the shape parameters are less than one, and it has been addressed in the next section.

6 A Modified Approach

6.1 A Modified likelihood Ratio Approach

In this section we propose a modified likelihood ratio approach mainly for small and moderate sample sizes, and when the shape parameters are less than one. The modified approach is based on the following two points: (i) it should converge to the standard likelihood ratio approach as the sample size increases, and (ii) the probability of correct selection should be at least 0.50 in both the cases. Suppose we have a random sample of size n, say $\{x_1, \ldots, x_n\}$ either from BS(α, β) or LN(σ, θ), and we want to detect the parent distribution. The following algorithm is proposed for this purpose. We will be using the following notations for brevity

$$\mu_1 = \lim_{n\to\infty} \frac{E_{BS}(T)}{n}, \quad \sigma_1^2 = \lim_{n\to\infty} \frac{V_{BS}(T)}{n}, \quad \mu_2 = \lim_{n\to\infty} \frac{E_{LN}(T)}{n}, \quad \sigma_2^2 = \lim_{n\to\infty} \frac{V_{LN}(T)}{n}.$$

Algorithm:

Step 1: Fit BS(α, β) to the data set and obtain $\widehat{\alpha}$ and $\widehat{\beta}$, the maximum likelihood estimators of α and β, respectively.

Step 2: Fit LN(σ, θ) to the data set and obtain $\widehat{\sigma}$ and $\widehat{\theta}$, the maximum likelihood estimators of σ and θ, respectively.

Step 3: Calculate the statistic T, say T_0, as defined in (4).

Step 4: Generate a sample of size n from BS$(\widehat{\alpha}, \widehat{\beta})$, calculate the statistic T, say T_{11}, as defined in (4).

Step 5: Repeat Step 4, B times, and obtain $T_{11}, \ldots, T_{B1}$.

Step 6: Based on $T_{11}, \ldots, T_{B1}$, obtain robust estimators of μ_1 and σ_1. Let us denote them as $\widehat{\mu}_1$ and $\widehat{\sigma}_1$, respectively.

Step 7: Repeat Step 4 - Step 6, by replacing BS$(\widehat{\alpha}, \widehat{\beta})$, with LN$(\widehat{\sigma}, \widehat{\theta})$, and obtain $\widehat{\mu}_2$ and $\widehat{\sigma}_2$.

Step 8: Calculate the critical point (CP) as

$$\text{CP} = \frac{\widehat{\mu}_1\widehat{\sigma}_2 + \widehat{\mu}_2\widehat{\sigma}_1}{\widehat{\sigma}_2 + \widehat{\sigma}_1}$$

- **Step 9:** Choose the Birnbaum-Saunders distribution as the parent distribution if $T_0 >$ CP, otherwise, choose it as the log-normal distribution.

It may be noted that Step 8 is obtained based on the discriminant analysis of two normal populations with means μ_1, μ_2 and variances σ_1^2, σ_2^2, respectively. It can be easily shown that the critical point $\dfrac{\mu_1\sigma_2 + \mu_2\sigma_1}{\sigma_2 + \sigma_1}$ is the optimal critical point, in the sense it provides the minimum probability of missclassfication based on the assumption $\ln \sigma_1 \approx \ln \sigma_2$.

6.2 Simulation Results

In this section we have performed some small scale simulations to show how the above proposed algorithm improves the probability of correct selection. We have already seen in Section 2 that discriminating between the BS(0.1704,131.8188) and LN(0.1695,131.8629) or BS(0.7654,77.3728) and (LN(0.7104,77.0394) is very difficult. We have presented the probability of correct selection based on the maximized likelihood ratio method as suggested in Section 3, in Tables 10 and 12 and based on the modified likelihood ration in Tables 11 and 13. In all the calculations $B = 1000$. From the Tables 10 to 13, it is clear that the proposed modified method is working quite well even for small sample sizes. It can maintain the probability of correct selection to be at least 0.5 in both the cases.

Table 10: Probability of correct selection based on the maximized likelihood ratio method when BS(0.1704,131.8188) and LN(0.1695,131.8629) are the two competing models. The first (second) row indicates the probability of correct selection when the data are generated from BS(0.1704,131.8188) (LN(0.1695,131.8629)).

Distribution	25	50	75	100	200	500	1000	5000	10000
BS	0.7159	0.6637	0.6450	0.6289	0.6013	0.5811	0.5689	0.6031	0.6810
LN	0.2913	0.3508	0.3724	0.3911	0.4298	0.4685	0.5055	0.5562	0.4970

Table 11: Probability of correct selection based on the modified maximized likelihood ratio method when BS(0.1704,131.8188) and LN(0.1695,131.8629) are the two competing models. The first (second) row indicates the probability of correct selection when the data are generated from BS(0.1704,131.8188) (LN(0.1695,131.8629)).

Distribution	25	50	75	100	200	500	1000	5000	10000
BS	0.5048	0.5060	0.4825	0.5093	0.5168	0.5264	0.5369	0.5554	0.5887
LN	0.5052	0.5066	0.5361	0.5111	0.5143	0.5258	0.5364	0.6035	0.5967

7 Conclusions

In this paper we have considered a classical problem of model discrimination between two non-nested distributions. We have considered two-parameter Birnbaum-Saunders and two-parameter log-normal distributions in this paper. We have used the traditional likelihood ratio approach to discriminate between these two distribution functions. It is observed that for certain range of shape parameters it is very difficult to discriminate between these two distribution functions. Some of the other methods like the ratio of maximized product of spacings and minimized Kolmogorov-Smirnov distance, have also been explored although not reported here, since they are not very satisfactory. We have provided a modified likelihood ratio method and it performs reasonably well. The Bayesian approach may be used in

Table 12: Probability of correct selection based on the maximized likelihood ratio method when BS(0.7654,77.3728) and LN(0.7104,77.0394) are the two competing models. The first (second) row indicates the probability of correct selection when the data are generated from BS(0.7654,77.3728) (LN(0.7104,77.0394)).

Distribution	25	50	75	100	200	500	1000	5000	10000
BS	0.7613	0.7487	0.7533	0.7589	0.7881	0.8689	0.9294	0.9991	1.0000
LN	0.3525	0.4567	0.5015	0.5523	0.6598	0.8028	0.9010	0.9990	1.0000

Table 13: Probability of correct selection based on the modified maximized likelihood ratio method when BS(0.7654,77.3728) and LN(0.7104,77.0394) are the two competing models. The first (second) row indicates the probability of correct selection when the data are generated from BS(0.7654,77.3728) (LN(0.7104,77.0394)).

Distribution	25	50	75	100	200	500	1000	5000	10000
BS	0.5595	0.5889	0.6174	0.6416	0.7027	0.8066	0.8924	0.9973	1.0000
LN	0.5728	0.6115	0.6473	0.6742	0.7460	0.8612	0.9323	0.9997	1.0000

this case to discriminate between these distribution functions if some prior information is available. Further, the work may be extended for censored observations or for one-shot device data, see for example Ling and Balakrishnan [27] or Prajapati et al. [28] in this respect. More work is needed in that direction.

8 Acknowledgements

The authors would like to thank the Associate Editor and two unknown reviewers for their constructive suggestions which have helped to improve the paper significantly.

Conflict of Interest

None of the authors has any conflict of interest in preperation of this manuscript.

Availability of Data and Materials

All the data which have been used are available in the manuscript itself.

Appendix A:

Proof of (11):

Suppose

$$\begin{aligned} g(\sigma,\theta) &= E_{BS}\left[\ln\left(f_{LN}(X;\sigma,\theta)\right)\right] \\ &= -\frac{1}{2}\ln(2\pi) - E_{BS}[\ln X] - \ln\sigma - \frac{1}{2\sigma^2}E[(\ln X)^2] + \frac{\ln\theta E[\ln X]}{\sigma^2} - \frac{(\ln\theta)^2}{2\sigma^2} \end{aligned}$$

Now, $\widetilde{\sigma}, \widetilde{\theta}$ are obtained by solving the partial derivatives of $g(\sigma,\theta)$.

$$\begin{aligned} 0 &= \frac{\partial}{\partial\theta}g(\sigma,\theta) = \frac{E_{BS}[\ln X]}{\theta\sigma^2} - \frac{\ln\theta}{\theta\sigma^2} \\ 0 &= \frac{\partial}{\partial\sigma}g(\sigma,\theta) = -\frac{1}{\sigma} + \frac{E_{BS}[\ln X]^2}{\sigma^3} + \frac{(\ln\theta)^2}{\sigma^3} - \frac{2\ln\theta E_{BS}[\ln X]}{\sigma^3}. \end{aligned}$$

Hence (11) can be obtained from the above two equations and using the expressions of $E(\ln Y)$ and $V(\ln Y)$ from Rieck and Nedelman [29].

Appendix B:

$$
\begin{aligned}
AM_{BS}(\alpha) &= E_{BS}\left[\ln(f_{BS}(\alpha,\beta)) - \ln(f_{LN}(\tilde{\sigma},\tilde{\theta}))\right] \\
&= \ln\left(\frac{\tilde{\sigma}}{2\alpha}\right) - \frac{1}{2}E_{BS}\left[\ln Y\right] + E_{BS}\left[\ln(1+Y)\right] + \frac{1}{2\tilde{\sigma}^2}E_{BS}\left[\ln\left(\frac{\beta Y}{\tilde{\theta}}\right)\right]^2 - \frac{1}{2} \\
&= \ln\left(\frac{\tilde{\sigma}}{2\alpha}\right) - \frac{1}{2}E_{BS}\left[\ln Y\right] + E_{BS}\left[\ln(1+Y)\right] + \frac{1}{2\tilde{\sigma}^2}\left\{E_{BS}\left[\ln Y\right]^2 - \left[\ln\left(\frac{\tilde{\theta}}{\beta}\right)\right]^2\right\} - \frac{1}{2} \\
&= \ln\left(\frac{\tilde{\sigma}}{2\alpha}\right) + E_{BS}\left[\ln(1+Y)\right] - \frac{1}{2}E_{BS}[\ln Y]
\end{aligned}
$$

and

$$
\begin{aligned}
AV_{BS}(\alpha) &= V_{BS}\left[\ln(f_{BS}(\alpha,\beta)) - \ln(f_{LN}(\tilde{\sigma},\tilde{\theta}))\right] \\
&= V_{BS}\left[\ln\left\{\left(\frac{\beta}{X}\right)^{1/2} + \left(\frac{\beta}{X}\right)^{3/2}\right\} - \frac{1}{2\alpha^2}\left\{\frac{X}{\beta} + \frac{\beta}{X}\right\} + \ln X + \frac{1}{2\tilde{\sigma}^2}\left\{\ln\left(\frac{X}{\tilde{\theta}}\right)\right\}^2\right] \\
&= V_{BS}\left[\ln(1+Y) - \frac{1}{2}\ln Y - \frac{1}{2\alpha^2}\left(Y + Y^{-1}\right) + \frac{1}{2\tilde{\sigma}^2}\left(\ln Y\right)^2 + \frac{1}{\tilde{\sigma}^2}\ln\left(\frac{\beta}{\tilde{\theta}}\right)\ln Y\right] \\
&= V_{BS}\left[\ln(1+Y) + \left(\frac{1}{\tilde{\sigma}^2}\ln\left(\frac{\beta}{\tilde{\theta}}\right) - \frac{1}{2}\right)\ln Y - \frac{1}{2\alpha^2}\left(Y + Y^{-1}\right) + \frac{1}{2\tilde{\sigma}^2}\left(\ln Y\right)^2\right] \\
&= V_{BS}\left[\ln(1+Y)\right] + \left(\frac{1}{\tilde{\sigma}^2}\ln\left(\frac{\beta}{\tilde{\theta}}\right) - \frac{1}{2}\right)^2 V_{BS}[\ln Y] + \frac{1}{4\alpha^4}V_{BS}\left[\left(Y + Y^{-1}\right)\right] \\
&+ \frac{1}{4\tilde{\sigma}^4}V_{BS}\left[(\ln Y)^2\right] + 2\left(\frac{1}{\tilde{\sigma}^2}\ln\left(\frac{\beta}{\tilde{\theta}}\right) - \frac{1}{2}\right)cov_{BS}\left[\ln(1+Y), \ln Y\right] \\
&- \frac{1}{\alpha^2}cov_{BS}\left[\ln(1+Y), \left(Y + Y^{-1}\right)\right] + \frac{1}{\tilde{\sigma}^2}cov_{BS}\left[\ln(1+Y), (\ln Y)^2\right] \\
&- \left(\frac{1}{\alpha^2\tilde{\sigma}^2}\ln\left(\frac{\beta}{\tilde{\theta}}\right) - \frac{1}{2\alpha^2}\right)cov_{BS}\left[\ln Y, (Y + Y^{-1})\right] \\
&+ \left(\frac{1}{\tilde{\sigma}^4}\ln\left(\frac{\beta}{\tilde{\theta}}\right) - \frac{1}{2\tilde{\sigma}^2}\right)cov_{BS}\left[\ln Y, (\ln Y)^2\right] - \frac{1}{2\alpha^2\tilde{\sigma}^2}cov_{BS}\left[(Y + Y^{-1}), (\ln Y)^2\right]
\end{aligned}
$$

Appendix C

$$
\begin{aligned}
AM_{LN}(\sigma) &= E_{LN}\left[\ln(f_{BS}(\widetilde{\alpha}, \widetilde{\beta})) - \ln(f_{LN}(\sigma, \theta))\right] \\
&= \ln\left(\frac{\sigma}{2\widetilde{\alpha}}\right) + \frac{1}{\sqrt{2\pi}}\int_0^\infty \frac{1}{\sigma z}\ln(z+1)e^{-\frac{(\ln z)^2}{2\sigma^2}}dz \\
AV_{LN}(\sigma) &= V_{LN}\left[\ln\left(Z+1\right)\right] + \frac{\sigma^2}{4} + \frac{e^{\sigma^2}(e^{\sigma^2}-1)}{2\widetilde{\alpha}^4} - cov_{LN}\left[\ln(Z+1), \ln Z\right] \\
&\quad -\frac{1}{\widetilde{\alpha}^2}cov_{LN}\left[\ln(Z+1), Z\right] - \frac{1}{\widetilde{\alpha}^2}cov_{LN}\left[\ln(Z+1), \frac{1}{Z}\right] + \frac{1-e^{\sigma^2}}{2\widetilde{\alpha}^4} + \frac{1}{2} \\
&\quad +\frac{1}{\sigma^2}cov_{LN}\left[\ln(Z+1), (\ln Z)^2\right] - \frac{\sigma^2 e^{\frac{\sigma^2}{2}}}{\widetilde{\alpha}^2}
\end{aligned}
$$

References

[1] B.B. Pereira and C.A.B. Pereira. *Model choice in nonnested families.* Springer, Berlin, Germany, 2016.

[2] J.K. Ghosh and K. Subramanyam. Inference about separated families in large samples. *Sankhya, Ser. A.*, 37:502–513, 1975.

[3] M.H. Pesaran. Global and partial nonnested hypothesis and asymptotic local power. *Economic Theory*, 3:69–75, 1987.

[4] R. Dumonceaux and C.E. Antle. Discrimination between the log-normal and the weibull distributions. *Technometrics*, 15(4):923–926, 1973.

[5] D. Kundu and A. Manglick. Discriminating between the weibull and log-normal distributions. *Naval Research Logistics*, 51:893–905, 2004.

[6] B. Basu, D. Tiwari, D. Kundu, and R. Prasad. Is weibull distribution the most appropriate statistical strength distribution for brittle materials? *Ceramics International*, 35:237–246, 2009.

[7] A. Basavalingappa, J.M. Passage, M.Y. Shen, and Lloyd J.H.H. Stassen. Electromigration: Log-normal versus weibull distribution. *Integrated Reliability Workshop (IIRW), IEEE International, Fallen Leaf Lake, CA, IEEE*, pages 1–4, 2017.

[8] S. Pasari. Stochastic modeling of earthquake interoccurence times in northwest himalaya and adjoining regions. *Geometics, Natural Hazards and Risk*, 9:568–588, 2018.

[9] A.K. Dey and D. Kundu. Discriminating between the weibull and log-normal distributions for type-ii censored data. *Statistics*, 46(2):197–214, 2012.

[10] J.S. Kim and B-J Yum. Selection between weibull and log-normal dis- tributions: A comparative simulation study. *Computational Statistics and Data Analysis*, 43:477–485, 2008.

[11] Z.W. Birnbaum and S.C. Saunders. A new family of life distribution. *Journal of Applied Probability*, 6:319–327, 1969.

[12] Z.W. Birnbaum and S.C. Saunders. Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, 6:328–347, 1969.

[13] A.J. Lemonte. A note on the fisher information matrix of the birnbaum-saunders distribution. *Journal of Statistical Theory and Applications*, 15:196–205, 2016.

[14] H. White. Maximum likelihood estimation of mis-specified models. *Econometrica*, 50:1–25, 1982.

[15] T. Bjerkedal. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *American Journal of Hygenes*, 72:130–148, 1960.

[16] R.C. Gupta, N. Kannan, and A. Raychaudhury. Analysis of lognormal survival data. *Mathematical Biosciences*, 139:103–115, 1997.

[17] D. Kundu, N. Kannan, and N. Balakrishnan. On the hazard function of birnbaum-saunders distribution and associated inference. *Computational Statistics and Data Analysis*, 52:2692–2702, 2008.

[18] H.K.T. Ng, D. Kundu, and Balakrishnan. Modified moment estimators for the two-parameter birnbaum-saunders distributions. *Computational Statistics and Data Analysis*, 43:283–298, 2003.

[19] L.J. Bain and M. Englehardt. Probability of correct selection of weibull versus gamma based on likelihood ratio. *Communications in Statistics - Theory and Methods*, 9:375–381, 1980.

[20] D.H. Fearn and E. Nebenzahl. On the maximum likelihood ratio method of deciding between the weibull and gamma distributions. *Communications in Statistics-Theory and Methods*, 20:579–593, 1991.

[21] A.W. Marshall, J.C. Meza, and I. Olkin. Can data recognize its parent distribution? *Journal of Computational and Graphical Statistics*, 10:555–580, 2001.

[22] A.K. Dey and D. Kundu. Discriminating among the log-normal, weibull and generalized exponential distributions. *IEEE Transactions on Reliability*, 58:416–424, 2009.

[23] N. Balakrishnan and D. Kundu. Birnbaum-saunders distribution: A review of models, analyses and applications (with discussions). *Applied Stochastic Models in Business and Industry*, 35:4–132, 2019.

[24] K. Greenwich. A unimodal hazard rate function and its failure distribution. *Statistical Papers*, 33:187–202, 1992.

[25] M. Engelhardt, L.J. Bain, and F.T. Wright. Inferences on the parameters of the birnbaum-saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics*, 23:251–255, 1981.

[26] F.S. Alshunnar, M.Z. Raqab, and D. Kundu. On the comparison of the fisher information of the log-normal and generalized rayleigh distributions. *Journal of Applied Statistics*, 37:391–404, 2010.

[27] M.H. Ling and N. Balakrishnan. Model mis-specification analysis of weibull and gamma models based on one-shot device test data. *IEEE Transactions on Reliability*, 66:641–650, 2017.

[28] D. Prajapati, M.H. Ling, P.S. Chan, and D. Kundu. Misspecification of copula for one-shot devices under constant stress accelerated life-tests. *Part O: Journal of Risks and Reliability*, 237:725–740, 2023.

[29] J.R. Rieck and J.R. Nedelman. A log-linear model for the birnbaum-saunders distribution. *Technometrics*, 33:51–60, 1991.