

Article

On Modeling Cancer and Tuberculosis Data Using the Birnbaum–Saunders Lifetime Model Established on a Logistic Kernel

Farouq Mohammad A. Alam ^{1,*}  and Abeer Mansour Almalki ^{1,2} 

¹ Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia; aalmalki2@uj.edu.sa

² Department of Statistics, Faculty of Science, University of Jeddah, Jeddah 21959, Saudi Arabia

* Correspondence: fmalam@kau.edu.sa

Abstract: Due to their importance in representing, explaining, and analyzing phenomena, statistical lifetime distributions are widely used in science. As a result, this paper discusses a modern lifetime model called Birnbaum–Saunders logistic distribution. This distribution extends the Birnbaum–Saunders distribution, as it has proven to be characterized by great flexibility in data modeling in practice. Different features of this distribution have been discussed. The parameters of the model are estimated using the maximum likelihood and modified moment estimation methods. To evaluate the performance of the methods, a simulation study with data contamination scenarios is presented. Finally, the new model’s flexibility is tested using real datasets.

Keywords: logistic Birnbaum–Saunders; hazard rate; critical value; maximum likelihood estimation; moment estimation; data contamination



Citation: Alam, F.M.A.; Almalki, A.M. On Modeling Cancer and Tuberculosis Data Using the Birnbaum–Saunders Lifetime Model Established on a Logistic Kernel. *Appl. Sci.* **2022**, *12*, 5000. <https://doi.org/10.3390/app12105000>

Academic Editors: Federico Divina and Giancarlo Mauri

Received: 24 March 2022

Accepted: 12 May 2022

Published: 16 May 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Lifetime models are a probability distribution with either a non-negative or positive support which are considered in survival analysis. In the literature, there exist many well-known life models, including, but not limited to, the exponential distribution and its generalization; the Weibull distribution. The main limitation of these two models is that they are not suitable for data with non-monotonic hazard rates. Examples of non-monotonic hazard rates are the bathtub and the upside-down hazard rates. From a medical perspective, the bathtub hazard rate represents three life phases. The first phase is called the infant mortality period which is a duration of time with decreasing hazard rate. The second part is the normal life period in which a constant hazard rate is maintained. The final phase is a period of time in which the hazard rate increases due to aging. Additionally, the upside-down hazard rate can be considered to describe the behavior of malicious diseases, such as cancer. In fact, ref. [1] observed such a hazard rate in the case of a certain type of breast cancer. They concluded that the mortality (i.e., hazard) rate increased to its highest peak which was approximately three years after the cancer was diagnosed. Afterwards, they observed that over a specified amount of time, the mortality rate decreases slowly.

In the literature, skewed life distributions such as the inverse Gaussian, the log-Gaussian, and the Birnbaum–Saunders distribution have been used to model phenomena with an upside-down hazard rate [2]. The latter model, however, received considerable attention from many researchers due to its desirable properties and physical interpretation. In fact, at least 200 articles and a single study monograph have previously been published detailing many aspects and advancements connected to this lifetime model. For more details, see [2]. The Birnbaum–Saunders (BS) distribution [3,4] belongs to a generalized BS (GBS) distribution, which was presented by [5] and later discussed by [6,7].

Researchers today deal with a variety of types and huge amounts of data from different fields, due to recent developments in applied sciences. In practice, there’s no guarantee

that these data are free of contamination, such as outliers, since contamination might originate from a variety of sources. Accordingly, estimation efficiency may be affected by this contamination. Therefore, the pollution aspect in data has motivated many researchers to suggest alternative estimators to those derived by the maximum likelihood theory for many distributions; see, for example, [8–12], among other papers.

This paper's aim is to study a specific instance of the GBS distribution; namely, the BS distribution established on a logistic kernel [13]. The considered model will henceforth be called the Logistic BS (LBS) distribution. The LBS distribution is considered for several reasons: the first is to investigate the behavior of LBS distribution compared to BS distribution in terms of distribution properties such as hazard (failure rate) function (HF). The second is to compare the application of LBS distribution to BS and other well-known distributions by analyzing two real-life medical datasets. Alongside these reasons, we are going to study the properties of the distribution and derive estimators for the model parameters using two estimation methods. Although the LBS distribution was considered previously by [14], this study differs from the latter contribution in a few points. First and foremost, the latter contribution pointed out that the hazard rate has a single nonmonotonic shape (i.e., inverse bathtub shape), while our study indicates that the hazard rate has two forms as shown in a later section. Second, while both [14] and this study derived similar statistical properties (e.g., mean, median, etc.), we have discussed some distributional aspects of the LBS distribution such as the corresponding mean residual life of the LBS distribution as well as the distribution of order statistics. Third, in terms of estimation, Ref. [14] considered six methods of estimation alongside interval estimation. In our study, however, we adopted the maximum likelihood method which is a vital method that was not considered by the previously mentioned study. Finally, ref. [14] analyzed simulated data, while our study analyzed real datasets to examine the distribution's application. The remaining parts of this article are presented in the following manner: in Section 2, the LBS distribution's fundamental and statistical properties are described. We also discuss the inference for the distribution by using two methods: Maximum likelihood estimation (MLE) and modified moment estimation (MME) in Section 3. In Section 4, a Monte Carlo simulation study is conducted, followed by a presentation of the real data analysis in Section 5.

2. Properties of the LBS Distribution

This section discusses the LBS distribution's fundamental and statistical properties.

2.1. Fundamental Properties

If a non-negative random variable X follows the GBS distribution established on an elliptically contoured kernel; say, $G(\cdot)$ with shape parameter $\alpha > 0$, scale parameter $\beta > 0$, and a location parameter γ , then the associated cumulative distribution function (CDF) is:

$$F(x; \alpha, \beta, \gamma) = G\left(\frac{1}{\alpha} \left[\sqrt{\frac{x - \gamma}{\beta}} - \sqrt{\frac{\beta}{x - \gamma}} \right]\right), \quad x > \gamma \geq 0, \alpha, \beta > 0. \quad (1)$$

Now, assuming $\gamma = 0$, we can substitute the CDF of the logistic distribution $(1 + e^{[-z]})^{-1}$ for the kernel $G(\cdot)$ in (1), then the CDF of a non-negative random variable X represented as:

$$F(x; \alpha, \beta) = \frac{1}{1 + e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)}}, \quad (2)$$

is said to follow the two-parameter LBS distribution with $\alpha > 0$ and $\beta > 0$ as shape and scale parameters, i.e., $X \sim \text{LBS}(\alpha, \beta)$, and the probability density function (PDF) that corresponds to it is:

$$f(x; \alpha, \beta) = \frac{1}{2\alpha x} \left(\sqrt{\frac{x}{\beta}} + \sqrt{\frac{\beta}{x}} \right) \frac{e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)}}{\left(1 + e^{-\frac{1}{\alpha} \left[\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right]} \right)^2}. \quad (3)$$

For different values of α and β , Figures 1–4 show the various shapes of the PDF and HF of the LBS distribution.

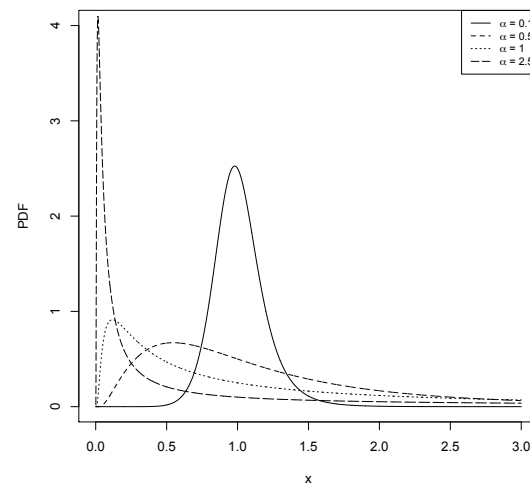


Figure 1. The PDF for the LBS distribution with different values for α and $\beta = 1$.

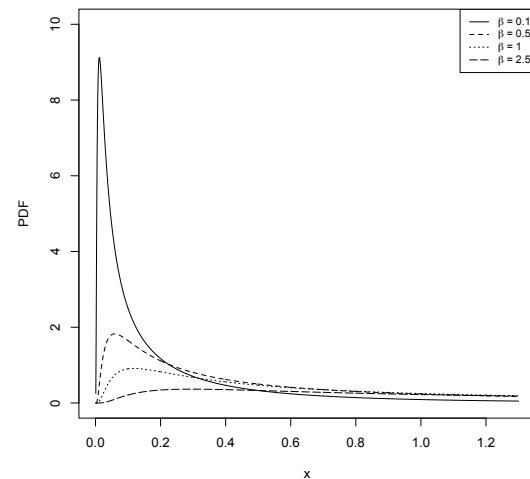


Figure 2. The PDF for the LBS distribution with different values for β and $\alpha = 1$.

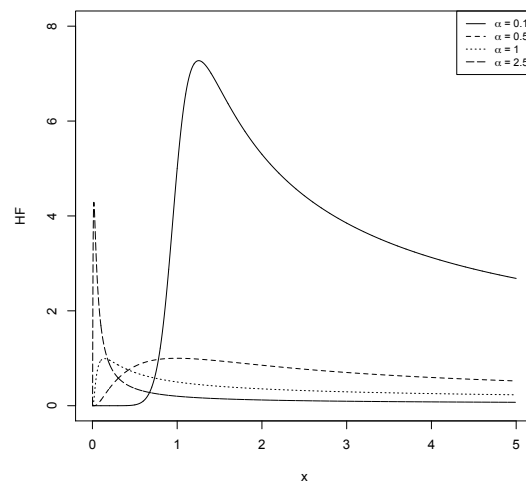


Figure 3. The HF for the LBS distribution with different values for α and $\beta = 1$.

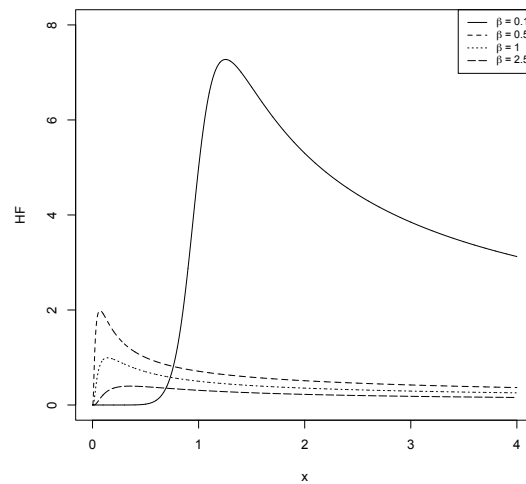


Figure 4. The HF for the LBS distribution with different values for β and $\alpha = 1$.

Evidently, if $X \sim \text{LBS}(\alpha, \beta)$ the quantile function is then as follows:

$$F^{-1}(u, \alpha, \beta) = \frac{\beta}{4} \left(\alpha q^{-1}(u) + \sqrt{[\alpha q^{-1}(u)]^2 + 4} \right)^2, \quad 0 < u < 1. \quad (4)$$

Clearly, using the inverse probability integral transform on the CDF (2), Equation (4) can be used to construct random variables having $\text{LBS}(\alpha, \beta)$ as its distribution, where $q^{-1}(u)$ is the inverse of the CDF of the standard logistic distribution. Accordingly, the median will be equal to β .

Upon using Equations (2) and (3), the HF can be obtained, which is defined as:

$$h(x; \alpha, \beta) = \frac{1}{2\alpha x} \left(\sqrt{\frac{x}{\beta}} + \sqrt{\frac{\beta}{x}} \right) \frac{1}{1 + e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)}}. \quad (5)$$

The HF in (5) is unimodal when $\alpha \rightarrow 0$ and decreases when $\alpha \rightarrow \infty$.

Lemma 1. To study the shape of HF, let us consider the following:

- a. The HF is decreasing for all values of x as $\alpha \rightarrow \infty$.
- b. If $\alpha \rightarrow \infty$, then the HF is an upside down function with critical point; say $c_{\alpha,\beta}$, such that $c_{\alpha,\beta} \geq (<) \beta$ when $\alpha \leq (>) 0.5$. For the proof of Lemma 1, see [15].

In the reliability analysis, the reliability (survival) function and the mean residual life (MRL) are significant characteristics of a lifetime model. On this basis and using the CDF we can obtain the reliability (survival) function as shown below:

$$R(x; \alpha, \beta) = P(X \geq x) = \frac{e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)}}{1 + e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)}}, \quad (6)$$

and

$$m(t; \alpha, \beta) = E(X - t | X > t) = \int_{\mathbb{R}} \exp \left(- \int_t^{t+x} h(\tau; \alpha, \beta) d\tau \right) dx, \quad (7)$$

respectively, where $h(\tau; \mu, \sigma)$ is the considered distribution's HF; see [16,17] in this regard. Hence, it can readily conclude from the MRL expression that it has an opposite attitude to the HF.

2.2. Order Statistics

Order statistics are another significant concept in the field of reliability analysis. Therefore, the r th order statistic's PDF can be computed as follows: Let $X_1, X_2, \dots, X_n$ denote an n -sized random sample having the LBS distribution and $X_{r:n}$ indicate the r th order statistic. Ref. [18] define the density of the order statistic, $X_{r:n}$, as:

$$f_{r:n}(x) = \frac{f(x)}{B(r, n-r+1)} \sum_{i=1}^{n-r} (-1)^i \binom{n-r}{i} [F(x)]^{r+i-1}, \quad (8)$$

where $B(a, b)$ refers to the beta function. The $f_{r:n}(x)$ of the LBS distribution is produced by substituting Equations (2) and (3) into Equation (8).

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)(2\alpha)} \sum_{i=1}^{n-r} \frac{(-1)^i}{x} \binom{n-r}{i} \left(\sqrt{\frac{x}{\beta}} + \sqrt{\frac{\beta}{x}} \right) \frac{e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)}}{\left(1 + e^{-\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right)} \right)^{r+i+1}}. \quad (9)$$

2.3. Statistical Properties

Property 1. If T follows standard logistic distribution, i.e., $T \sim \text{Logistic}(0, 1)$, then:

$$X = \frac{\beta}{4} \left(\alpha T + \sqrt{[\alpha T]^2 + 4} \right)^2 \sim \text{LBS}(\alpha, \beta).$$

Property 1 was used to calculate the LBS distribution's expected value, variance, skewness, and kurtosis.

Property 2. If $T \sim \text{LBS}(\alpha, \beta)$, then

1. $\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}} \right) \sim \text{Logistic}(0, 1)$;
2. $cX \sim \text{LBS}(\alpha, c\beta)$;
3. $X^{-1} \sim \text{LBS}(\alpha, \beta^{-1})$.

Property 3. If $T \sim LBS(\alpha, \beta)$, then, we can show that $E(X) = \beta(\frac{1}{6}\alpha^2\pi^2 + 1)$, $V(X) = \frac{(\alpha\pi\beta)^2}{3}(\frac{37}{60}\pi^2\alpha^2 + 1)$, $\gamma_1(X) = \frac{30(155\alpha^6\pi^6 + 294\alpha^4\pi^4 + 315\alpha^2\pi^2 + 210)^2}{49(7\alpha^4\pi^4 + 20\alpha^2\pi^2 + 30)^3}$ and $\gamma_2(X) = \frac{30(889\alpha^8\pi^8 + 1240\alpha^6\pi^6 + 980\alpha^4\pi^4 + 560\alpha^2\pi^2 + 210)}{7(7\alpha^4\pi^4 + 20\alpha^2\pi^2 + 30)^2}$. We also can easily obtain $E(X^{-1}) = \beta^{-1}(\frac{1}{6}\alpha^2\pi^2 + 1)$ and $V(X^{-1}) = \frac{(\alpha\pi\beta^{-1})^2}{3}(\frac{37}{60}\pi^2\alpha^2 + 1)$ by using Property 2.

3. Inference for the LBS Distribution

This section examines the inference for the LBS distribution using two methods: MLE and MME.

3.1. Maximum Likelihood Estimation

The MLEs for LBS distribution parameters are discussed in this section. Let $(x_1, x_2, \dots, x_n)$ denote an n -sized random sample having the LBS distribution with unknown parameter vector $\zeta = (\alpha, \beta)^T$. Then, for ζ , the log-likelihood function was constructed as follows:

$$\begin{aligned} \ell(\zeta) &= -n\ln(2\alpha) - \sum_{i=1}^n \ln(x_i) + \sum_{i=1}^n \ln\left(\frac{x_i + \beta}{\sqrt{\beta x_i}}\right) - \frac{1}{\alpha} \sum_{i=1}^n \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right) - 2 \sum_{i=1}^n \ln\left(1 + e^{-\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right]}\right) \\ &= -n\ln(2\alpha) - \sum_{i=1}^n \ln(x_i) + \sum_{i=1}^n \ln(x_i + \beta) - \frac{n}{2} \ln(\beta) - \frac{1}{2} \sum_{i=1}^n \ln(x_i) - \frac{1}{\alpha} \sum_{i=1}^n \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right) \\ &\quad - 2 \sum_{i=1}^n \ln\left(1 + e^{-\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right]}\right). \end{aligned} \quad (10)$$

The MLEs $\hat{\zeta} = (\hat{\alpha}, \hat{\beta})^T$ of $\zeta = (\alpha, \beta)^T$ are found by solving the nonlinear systems of equations shown below:

$$\frac{\partial \ell(\zeta)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^2} \sum_{i=1}^n \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right) - \frac{2}{\alpha^2} \sum_{i=1}^n \frac{\left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right)}{\left(1 + e^{\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right]}\right)} = 0 \quad (11)$$

$$\frac{\partial \ell(\zeta)}{\partial \beta} = \sum_{i=1}^n \left(\frac{1}{x_i + \beta}\right) - \frac{n}{2\beta} + \frac{1}{2\alpha\beta} \sum_{i=1}^n \left(\sqrt{\frac{\beta}{x_i}} + \sqrt{\frac{x_i}{\beta}}\right) - \frac{1}{\alpha\beta} \sum_{i=1}^n \frac{\left(\sqrt{\frac{\beta}{x_i}} + \sqrt{\frac{x_i}{\beta}}\right)}{\left(1 + e^{\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}}\right]}\right)} = 0. \quad (12)$$

The second derivatives of ℓ can be calculated in the following way:

$$\begin{aligned} \frac{\partial^2 \ell(\zeta)}{\partial \alpha^2} &= \frac{n}{\alpha^2} - \frac{2}{\alpha^3} \sum_{i=1}^n \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right) \\ &+ \sum_{i=1}^n \frac{2 \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right) \left[\left(\sqrt{\frac{\beta}{x_i}} - \sqrt{\frac{x_i}{\beta}} + 2\alpha \right) e^{\frac{1}{\alpha} \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right)} + 2\alpha \right]}{\alpha^4 \left(1 + e^{\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right]} \right)^2} = A_1, \end{aligned} \quad (13)$$

$$\begin{aligned} \frac{\partial^2 \ell(\zeta)}{\partial \alpha \partial \beta} &= -\frac{1}{2\alpha^2 \beta} \sum_{i=1}^n \left(\sqrt{\frac{\beta}{x_i}} + \sqrt{\frac{x_i}{\beta}} \right) \\ &+ \sum_{i=1}^n \frac{\left(x_i^2 \sqrt{\frac{\beta}{x_i}} + \beta^2 \sqrt{\frac{x_i}{\beta}} \right) \left[\left(\sqrt{\frac{\beta}{x_i}} - \sqrt{\frac{x_i}{\beta}} + \alpha \right) e^{\frac{1}{\alpha} \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right)} + \alpha \right]}{\alpha^3 \beta^2 x_i \sqrt{\frac{x_i}{\beta}} \sqrt{\frac{\beta}{x_i}} \left(1 + e^{\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right]} \right)^2} = A_2, \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial^2 \ell(\zeta)}{\partial \beta^2} &= -\sum_{i=1}^n \left(\frac{1}{(x_i + \beta)^2} \right) + \frac{n}{2\beta^2} - \frac{1}{4\alpha\beta^2} \sum_{i=1}^n \left(3\sqrt{\frac{\beta}{x_i}} + \sqrt{\frac{x_i}{\beta}} \right) \\ &- \sum_{i=1}^n \frac{\left[\left(\sqrt{\frac{x_i}{\beta}} - 3\alpha \right) x_i^4 \left(\frac{\beta}{x_i} \right)^{\frac{3}{2}} + \left(\sqrt{\frac{\beta}{x_i}} - \alpha \right) \beta^4 \left(\frac{x_i}{\beta} \right)^{\frac{3}{2}} + 2x_i^2 \beta^2 \right] e^{\frac{1}{\alpha} \left(\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right)} - 3\alpha x_i^4 \left(\frac{\beta}{x_i} \right)^{\frac{3}{2}} - \alpha \beta^4 \left(\frac{x_i}{\beta} \right)^{\frac{3}{2}}}{2\alpha^2 \beta^4 x_i \left(\frac{x_i}{\beta} \right)^{\frac{3}{2}} \left(\frac{\beta}{x_i} \right)^{\frac{3}{2}} \left(1 + e^{\frac{1}{\alpha} \left[\sqrt{\frac{x_i}{\beta}} - \sqrt{\frac{\beta}{x_i}} \right]} \right)^2} \\ &= A_3. \end{aligned} \quad (15)$$

The observed information matrix is then calculated as follows:

$$\begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = - \begin{pmatrix} A_1 & A_2 \\ A_2 & A_3 \end{pmatrix}.$$

Then, we can approximate the observed variance-covariance matrix as:

$$\begin{aligned} V &= \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}^{-1} \\ &= \frac{1}{\det I} \begin{pmatrix} I_{22} & -I_{21} \\ -I_{21} & I_{11} \end{pmatrix}. \end{aligned}$$

Accordingly, $\hat{\alpha}$ and $\hat{\beta}$ asymptotic joint distribution is approximately bivariate normal, as shown by:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \approx N \left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \right]. \quad (16)$$

3.2. Modified Moment Estimation

The MMEs were proposed by [19] for BS distribution. Inspired by their research, the MMEs for LBS are proposed here by equalizing $E(T)$ and $E(\frac{1}{T})$ with the relevant

sample moments. In this process, the MMEs for α and β , represented by $\tilde{\alpha}$ and $\tilde{\beta}$, are procured as follows:

$$\tilde{\alpha} = \left[\frac{6}{\pi^2} \left[\left(\frac{s}{r} \right)^{\frac{1}{2}} - 1 \right] \right]^{\frac{1}{2}} \quad \text{and} \quad \tilde{\beta} = (sr)^{\frac{1}{2}}, \quad (17)$$

where $s = \frac{1}{n} \sum_{i=1}^n x_i$ and $r = \left[\frac{1}{n} \sum_{i=1}^n x_i^{-1} \right]^{-1}$.

The strong law of large numbers states that s and r , respectively, converge to $E(T)$ and $E(T^{-1})$. The asymptotic joint distribution of $\tilde{\alpha}$ and $\tilde{\beta}$ can be determined using CLT as shown below:

$$\sqrt{n} \begin{pmatrix} s - E(T) \\ r^{-1} - E(T^{-1}) \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right],$$

where

$$\Sigma = \begin{bmatrix} \frac{(\alpha\pi\beta)^2}{3} (1 + \frac{37}{60}\pi^2\alpha^2) & 1 - (1 + \frac{\pi^2}{6}\alpha^2)^2 \\ 1 - (1 + \frac{\pi^2}{6}\alpha^2)^2 & \frac{(\alpha\pi\beta^{-1})^2}{3} (1 + \frac{37}{60}\pi^2\alpha^2) \end{bmatrix}. \quad (18)$$

By using Taylor's expansion, we get:

$$\tilde{\alpha} \approx \alpha + \frac{1}{4\alpha\beta}s + \frac{\beta}{4\alpha}r^{-1} \quad \text{and} \quad \tilde{\beta} \approx \beta + \frac{1}{2(1 + \frac{\pi^2}{6}\alpha^2)}s - \frac{\beta^2}{2(1 + \frac{\pi^2}{6}\alpha^2)}r^{-1}.$$

We can now calculate the asymptotic joint distribution of $(\tilde{\alpha}, \tilde{\beta})$ by utilizing the asymptotic joint distribution of (s, r^{-1}) .

$$\begin{pmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{pmatrix} \sim N \left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \frac{\pi^4\alpha^2}{45n} & 0 \\ 0 & (\pi\alpha\beta)^2 \left(\frac{3(20+7\pi^2\alpha^2)}{5n(6+\pi^2\alpha^2)^2} \right) \end{pmatrix} \right]$$

4. Simulation Study

As previously mentioned, it is necessary to assess the estimators' performance when the data are contaminated. In a similar way to [11,20], we take the following scenarios into consideration:

Model 1: A model without any contamination.

Model 2: A model with 10% of upper contamination.

Model 3: A model with 10% of lower contamination.

Model 4: A model with 20% of two-tailed contamination.

Therefore, this section presents the findings of the MC simulations study based on $M = 10,000$ simulation runs with various combinations of shape parameter and sample size values. By multiplying the upper or lower order statistics by 5 or 1/5, data contamination was achieved. In addition, the simulation study uses $n = 30, 50, 100, 150, 200, 500$, $\alpha = 0.5, 0.75, 1, 1.5, 2.0$, and $\beta = 1$ for each scenario, with no loss of generality. It should be mentioned that all calculations were carried out using an **R** program, which is accessible from the authors upon request.

The simulated bias and root mean squared error (RMSE) are generated to assess estimation efficiency.

$$\text{Bias}(\alpha) = \frac{1}{M} \sum_{i=1}^M (\hat{\alpha}_i - \alpha), \quad \text{Bias}(\beta) = \frac{1}{M} \sum_{i=1}^M (\hat{\beta}_i - \beta),$$

and

$$\text{RMSE}(\alpha) = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{\alpha}_i - \alpha)^2}, \quad \text{RMSE}(\beta) = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{\beta}_i - \beta)^2}.$$

The simulation study's results are visually presented in Figures 5–8, and from these figures, we can note the following:

- When there is no contamination in the data, all figures show that when the sample size is increased, all methods perform well;
- The bias values of the α parameter increase as the α value increases in models 2 and 3;
- In the case of model 4, the bias values show that the MLE method outperforms the MME method, with MLE performing significantly better as the α value decreases;
- The RMSE values of the α parameter in models 2 and 3 are slightly increased when the α value increases in the two methods;
- However, the RMSE values of model 4 indicate that MLE outperforms MME and that this performance improves as α decreases;
- In the case of models 1 and 4, the bias values of the β parameter appear to be constant around zero for both methods;
- As the α value decreases in the MLE method, the bias values of models 2 and 3 approach zero;
- The RMSE values of the β parameter perform similarly in the MLE and MME methods, with the values of all models getting closer to zero as the α value decreases.

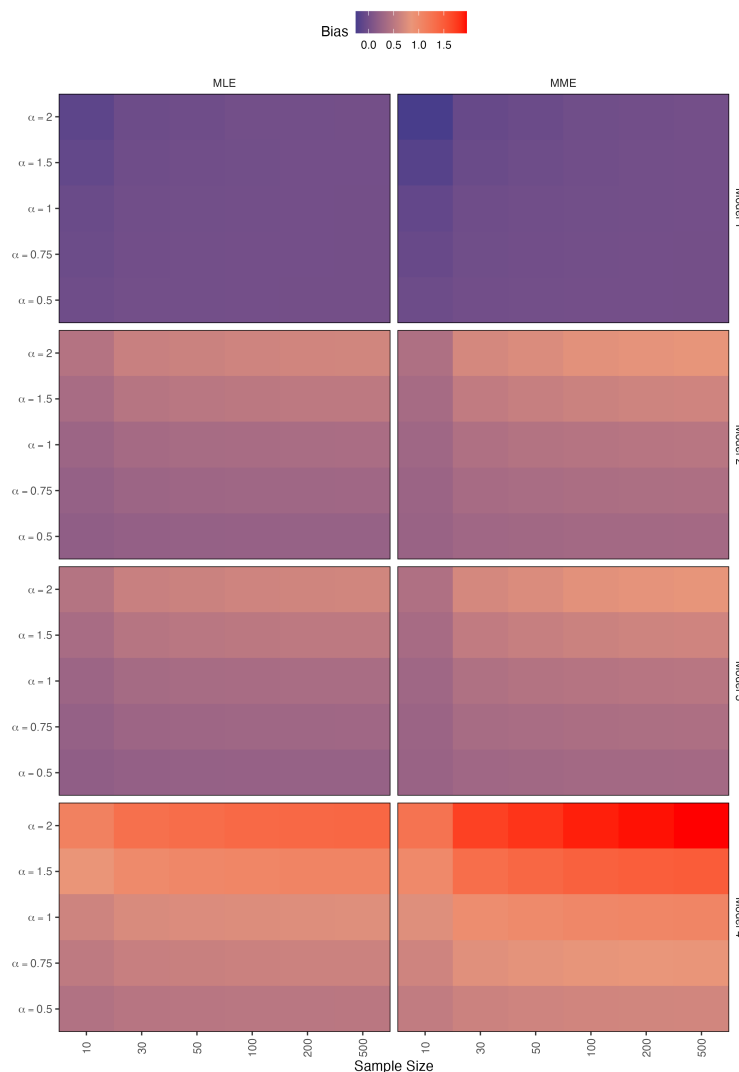


Figure 5. Simulated biases for the estimators of α .

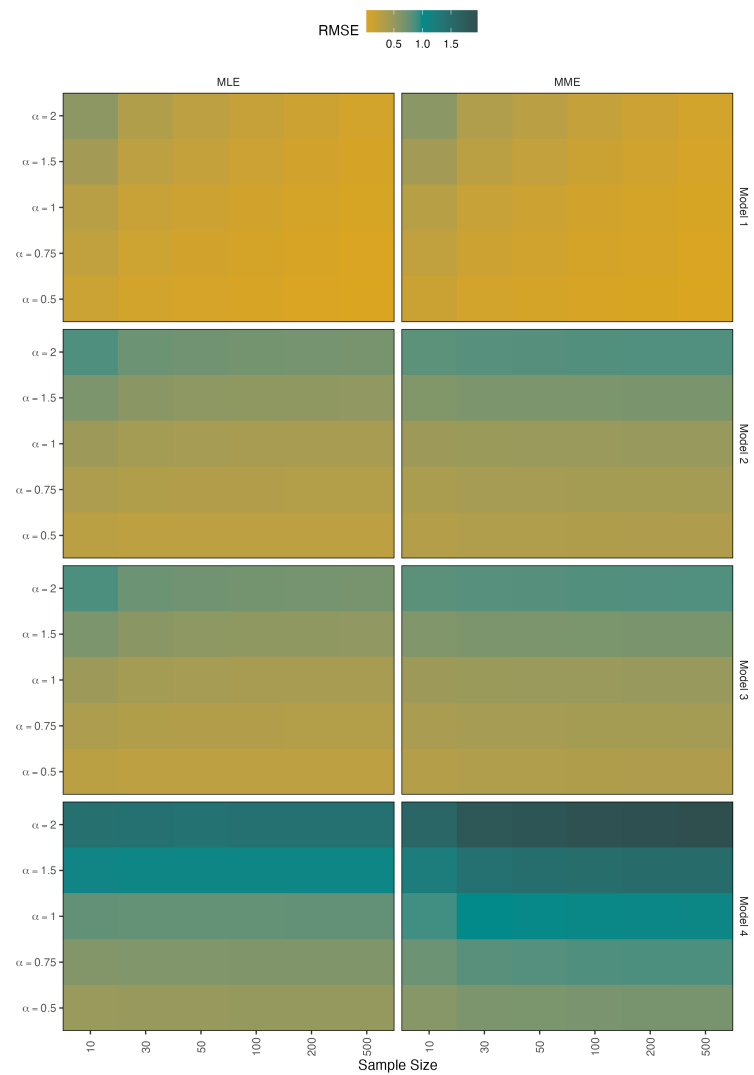


Figure 6. Simulated RMSEs for the estimators of α .

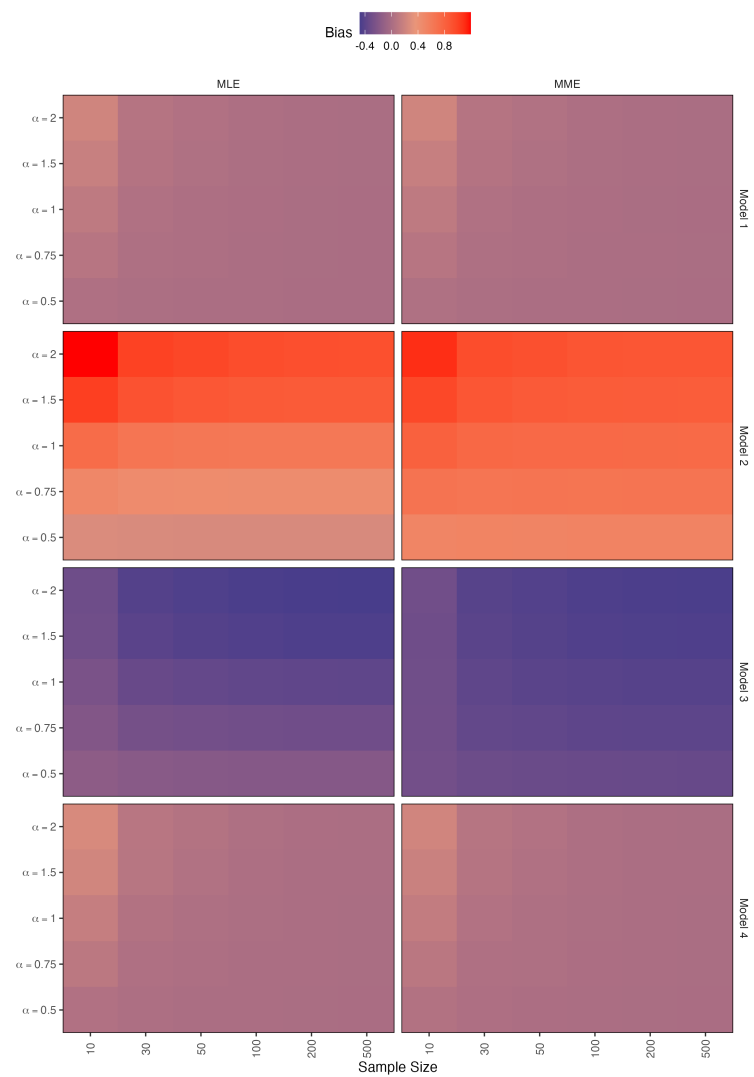


Figure 7. Simulated biases for the estimators of β .

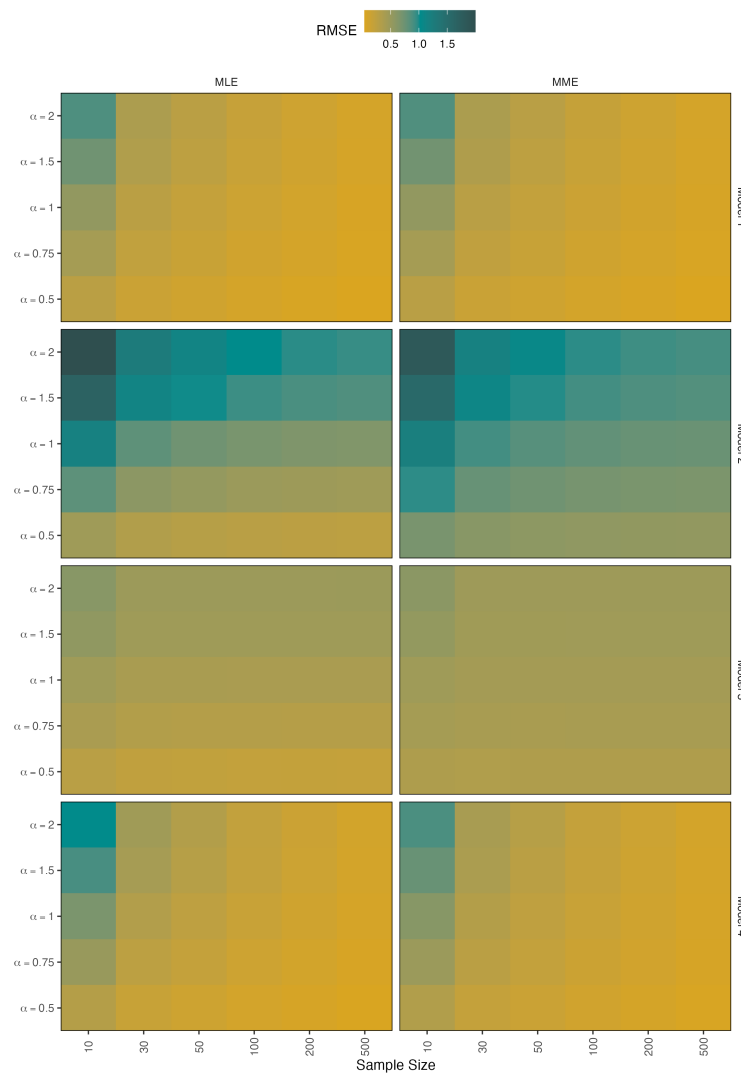


Figure 8. Simulated RMSEs for the estimators of β .

5. Data Analyses

We perform two data analyses in this section to compare the performance of LBS distribution with some other well-known distributions. The first set of data under examination consists of the lifetimes of 42 individuals who were diagnosed with head and neck cancer, which had previously been used by [21]. Cancer that evolves in the mouth, throat, nose, salivary glands or other parts of the head and neck is referred to as head and neck cancer. Table 1 presents the dataset:

Table 1. Lifetimes of individuals who were diagnosed with Head and Neck Cancer.

7	91	140	160	248	440
34	108	140	165	273	523
42	112	146	173	277	583
63	129	149	176	297	594
64	133	154	218	405	1101
83	133	157	225	417	1146
84	139	160	241	420	1417

It is important to mention that we eliminated data based on participants being lost to follow-up, as indicated by [21]; that is, we have omitted the censored data and considered

the remaining ones as a complete sample. Modeling the actual data by [21] (i.e., the case of censored data) is considered for future research.

The second set of data under examination represents the lifetimes of 72 *Cavia Porcellus* (guinea pigs) injected with various doses of *Mycobacterium tuberculosis*, which was previously studied by [22] among others researchers. *Mycobacterium tuberculosis* is a bacterium that causes tuberculosis, a contagious disease that primarily affects the lungs. Table 2 displays the dataset:

Table 2. Lifetimes of *Cavia Porcellus* injected with various doses of *Mycobacterium Tuberculosis*.

12	44	60	70	95	146
15	48	60	72	96	175
22	52	60	73	98	175
24	53	60	75	99	221
24	54	61	76	109	233
32	54	62	76	110	258
32	55	63	81	121	258
33	56	65	83	127	263
34	57	65	84	129	297
38	58	67	85	131	341
38	58	68	87	143	341
43	59	70	91	146	376

First, we estimate the MLEs, Bias, and RMSE of parameters, AIC and BIC of Exponential, Rayleigh, Weibull, BS, and LBS distributions, then depending on the considered dataset, we compute many explanatory data analysis (EDA) measurements, as well as their approximations based on the estimated model parameters.

Tables 3 and 4 list the estimators of the models' parameters as well as the AIC and BIC, while Tables 5 and 6 present the outcomes of EDA measurements. Overall, when comparing these tables, one can observe that the estimated EDA measurements assuming the LBS distribution were very near to their observed counterparts which are calculated from the sample directly. Furthermore, the AIC and BIC indicate that the LBS distribution is a better fit. In comparison to the fitted PDFs, Figures 9 and 10 show the histogram of the datasets. As can be seen in the two figures, the LBS distribution fits the data better than other distributions.

Table 3. Values of MLEs, AIC and BIC of Exponential, Rayleigh, Weibull, BS, LBS and Laplace BS distributions for Head and Neck Cancer data.

Model	α			β			AIC	BIC
	Value	Bias	RMSE	Value	Bias	RMSE		
Exponential	-	-	-	280.1667	0	46.7732	559.3724	561.1101
Rayleigh	-	-	-	410.1102	-7.6502	79.2048	601.3328	603.0705
Weibull	1.0918	0.0491	0.13	290.9757	6.558	43.4886	560.7841	564.2595
BS	1.1724	-0.0684	0.2855	162.4521	9.8421	45.9678	565.1072	568.5826
LBS	0.5736	0.0485	0.1076	179.6688	37.6861	27.0372	557.3078	560.7831

Table 4. Values of MLEs, AIC and BIC of Exponential, Rayleigh, Weibull, BS, LBS and Laplace BS distributions for Cavia Porcellus data.

Model	α			β			AIC	BIC
	Value	Bias	RMSE	Value	Bias	RMSE		
Exponential	-	-	-	99.81944	0	9.5598	808.8843	811.1610
Rayleigh	-	-	-	128.2679	-0.7152	13.5447	818.5921	820.8688
Weibull	1.39295	0.0168	0.0995	110.5393	-2.8794	10.0684	798.2955	802.8488
BS	0.75999	-0.0099	0.0740	77.52778	0.2859	6.7819	785.8348	790.3881
LBS	0.41460	-0.0033	0.0457	76.01035	1.1905	6.1102	783.7526	788.3060

Table 5. EDA outcomes based on Head and Neck Cancer data.

	Q1	Q2	Q3	Mean	Kurtosis	Skewness	SD
Sample	130	160	292	280.1667	8.0993	2.3223	303.1249
Exponential	80.5989	194.1967	388.3935	280.1667	9	2	280.1667
Rayleigh	4311.0799	482.8679	682.8783	363.4507	3.2451	0.6311	268.6781
Weibull	92.9533	208.0012	392.4504	281.4662	7.4714	1.7535	258.0723
BS	75.1050	162.4521	351.3837	274.0991	14.7086	2.7626	314.0061
LBS	96.6347	179.6688	334.0506	276.9076	23.1251	11.8393	323.9002

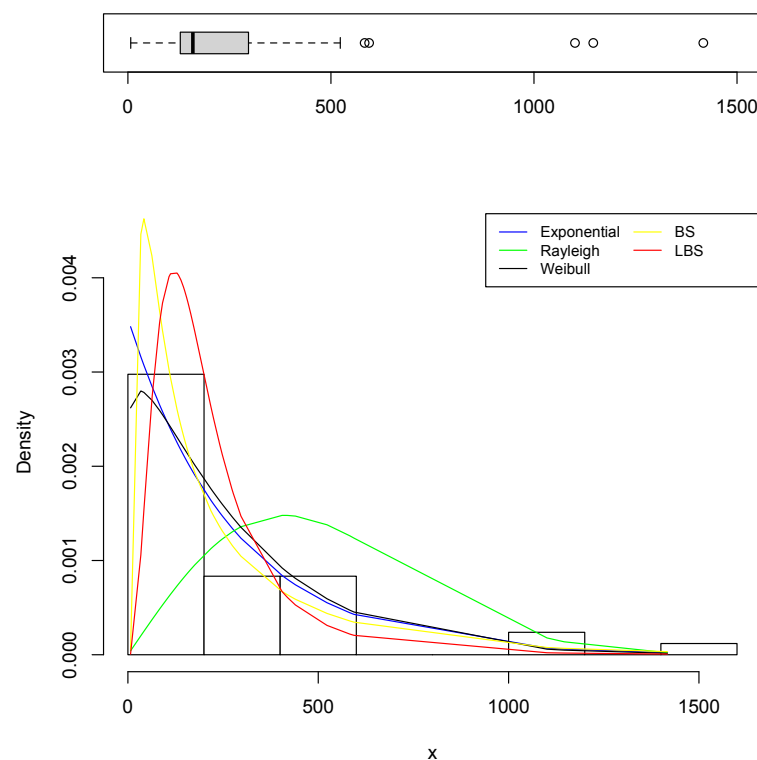
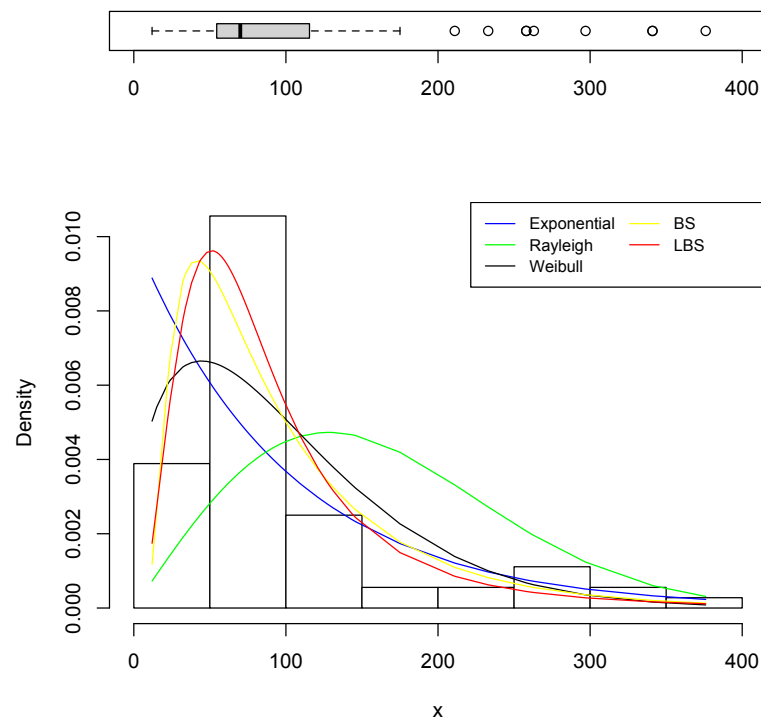
**Figure 9.** The fitted PDFs based on Head and Neck Cancer data.

Table 6. EDA outcomes based on Cavia Porcellus data.

	Q1	Q2	Q3	Mean	Kurtosis	Skewness	SD
Sample	54.7500	69.1896	112.7500	99.81944	5.6144	1.7962	81.1180
Exponential	28.7163	70	138.3791	99.81944	9	2	99.8194
Rayleigh	97.2947	151.0239	213.5801	113.6745	3.2451	0.6311	84.0330
Weibull	45.1928	84.9660	139.7509	100.8287	4.8757	1.2081	73.3186
BS	46.6880	77.5278	128.7388	99.9169	9.8454	2.0774	77.3171
LBS	48.3869	76.0104	119.4038	97.5030	10.5492	5.6765	81.7657

**Figure 10.** The fitted PDFs based on Cavia Porcellus data.

6. Conclusions

The features of the LBS distribution and certain associated inferential methods have been discussed in this paper. The PDF of the LBS distribution might take on many shapes due to the inclusion of the shape parameter. It has either a decreasing or a unimodal HF function. When there is no contamination in the data, all approaches perform well, according to the MC simulations study. The contamination in the data may have an impact on the shape and scale parameters, with the exception of two-tailed contamination, which has no impact on the scale parameter. Finally, based on the results of the real data analysis, it is concluded that the LBS distribution, as compared to the well-known distributions, may provide a better fit to real-life data. In practice, data contamination is not the only issue researchers encounter. Another significant research challenge that needs to be addressed in future studies is data censoring, since it negatively impacts estimation efficiency and robustness. Comparing the examined estimators to Bayesian estimation in terms of performance is the last study area that one may consider.

Author Contributions: Under the idea and supervision of the corresponding author F.M.A.A., A.M.A. carried out the numerical calculations. Both F.M.A.A. and A.M.A. contributed to the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors express their sincere gratitude to the anonymous reviewers and the editors for providing useful suggestions on earlier versions of this manuscript which resulted in this improved version.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Langlands, A.O.; Pocock, S.J.; Kerr, G.R.; Gore, S.M. Long-term survival of patients with breast cancer: A study of the curability of the disease. *Br. Med. J.* **1979**, *2*, 1247–1251. [[CrossRef](#)] [[PubMed](#)]
2. Balakrishnan, N.; Kundu, D. Birnbaum–Saunders distribution: A review of models, analysis, and applications. *Appl. Stoch. Model. Bus. Ind.* **2019**, *35*, 4–49. [[CrossRef](#)]
3. Birnbaum, Z.W.; Saunders, S.C. A new family of life distributions. *J. Appl. Probab.* **1969**, *6*, 319–327. [[CrossRef](#)]
4. Birnbaum, Z.W.; Saunders, S.C. Estimation for a family of life distributions with applications to fatigue. *J. Appl. Probab.* **1969**, *6*, 328–347. [[CrossRef](#)]
5. Díaz-García, J.A.; Leiva-Sánchez, V. A new family of life distributions based on the elliptically contoured distributions. *J. Stat. Plan. Infer.* **2005**, *128*, 445–457; Erratum in *J. Stat. Plan. Infer.* **2007**, *137*, 1512–1513. [[CrossRef](#)]
6. Leiva, V.; Riquelme, M.; Balakrishnan, N.; Sanhueza, A. Lifetime analysis based on the generalized Birnbaum–Saunders distribution. *Comput. Stat. Data Anal.* **2008**, *52*, 2079–2097. [[CrossRef](#)]
7. Sanhueza, A.; Leiva, V.; Balakrishnan, N. The generalized Birnbaum–Saunders distribution and its theory, methodology, and application. *Commun. Stat. Methods* **2008**, *37*, 645–670. [[CrossRef](#)]
8. Lawson, C.; Keats, J.B.; Montgomery, D.C. Comparison of robust and least-squares regression in computer-generated probability plots. *IEEE Trans. Reliab.* **1997**, *46*, 108–115. [[CrossRef](#)]
9. Boudt, K.; Caliskan, D.; Croux, C. Robust explicit estimators of Weibull parameters. *Metrika* **2011**, *73*, 187–209. [[CrossRef](#)]
10. Agostinelli, C.; Marazzi, A.; Yohai, V.J. Robust estimators of the generalized log-gamma distribution. *Technometrics* **2014**, *56*, 92–101. [[CrossRef](#)]
11. Wang, M.; Park, C.; Sun, X. Simple robust parameter estimation for the Birnbaum–Saunders distribution. *J. Stat. Distrib. Appl.* **2015**, *2*, 1–11. [[CrossRef](#)]
12. Alam, F.M.A.; Nassar, M. On Modeling Concrete Compressive Strength Data Using Laplace Birnbaum–Saunders Distribution Assuming Contaminated Information. *Crystals* **2021**, *11*, 830. [[CrossRef](#)]
13. Balakrishnan, N. *Handbook of the Logistic Distribution*; CRC Press: Boca Raton, FL, USA, 1991.
14. Xu, X.; Wang, R.; Gu, B. Statistical Analysis of Two-Parameter Generalized BS-Logistic Fatigue Life Distribution. *J. Syst. Sci. Complex.* **2019**, *32*, 1231–1250. [[CrossRef](#)]
15. Alam, F.M.A.; Almalki, A.M. The hazard rate function of the logistic Birnbaum–Saunders distribution: Behavior, associated inference, and application. *J. King Saud Univ.-Sci.* **2021**, *33*, 101580. [[CrossRef](#)]
16. Watson, G.; Wells, W. On the possibility of improving the mean useful life of items by eliminating those with short lives. *Technometrics* **1961**, *3*, 281–298. [[CrossRef](#)]
17. Gupta, R.C.; Bradley, D.M. Representing the mean residual life in terms of the failure rate. *Math. Comput. Model.* **2003**, *37*, 1271–1280. [[CrossRef](#)]
18. David, H.A.; Nagaraja, H.N. *Order Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
19. Ng, H.; Kundu, D.; Balakrishnan, N. Modified moment estimation for the two-parameter Birnbaum–Saunders distribution. *Comput. Stat. Data Anal.* **2003**, *43*, 283–298. [[CrossRef](#)]
20. Dupuis, D.J.; Mills, J.E. Robust estimation of the Birnbaum–Saunders distribution. *IEEE Trans. Reliab.* **1998**, *47*, 88–95. [[CrossRef](#)]
21. Efron, B. Logistic regression, survival analysis, and the Kaplan–Meier curve. *J. Am. Stat. Assoc.* **1988**, *83*, 414–425. [[CrossRef](#)]
22. Gupta, C.R.; Kannan, N.; Raychaudhuri, A. Analysis of lognormal survival data. *Math. Biosci.* **1997**, *139*, 103–115. [[CrossRef](#)]