Failure rate of Birnbaum-Saunders distributions: shape, change-point, estimation and robustness

Emilia Athayde¹,², Assis Azevedo¹,², Michelli Barros³, Víctor Leiva⁴,⁵
¹Centre of Mathematics, Universidade do Minho, Portugal
²Department of Mathematics, Universidade do Minho, Portugal
³Department of Statistics, Universidade Federal de Campina Grande, Brazil
⁴School of Industrial Engineering, Pontificia Universidad Católica de Valparaíso, Chile
⁵Faculty of Administration, Accounting and Economics, Universidade Federal de Goiás, Brazil

Abstract

The Birnbaum-Saunders (BS) distribution has been largely studied and applied. A random variable with BS distribution is a transformation of another random variable with standard normal distribution. Generalized BS distributions are obtained when the normally distributed random variable is replaced by another symmetrically distributed random variable. This allows us to obtain a wide class of positively skewed models with lighter and heavier tails than the BS model. Its failure rate admits several shapes, including the unimodal case, with its change-point being able to be used for different purposes. For example, to establish the reduction in a dose, and then in the cost of the medical treatment. We analyze the failure rates of generalized BS distributions obtained by the logistic, normal and Student-t distributions, considering their shape and change-point, estimating them, evaluating their robustness, assessing their performance by simulations, and applying the results to real data from different areas.

1 Introduction

The failure rate (FR) is a useful indicator for reliability and survival analyses. It can be increasing (IFR class), decreasing (DFR class) or constant (exponential distribution), but the non-monotone cases are also detected in practice. In these latter cases, the change-point of the FR, conducting to bathtub (BT) or inverse bathtub (IBT) shapes, is of interest, as well as its limit behavior approaching to zero, to a positive constant (not at zero), or to infinity. Then, the change-point of the FR is an important value, because the hazard begins to decrease at this point.

Atypical data are present in several fields, including reliability and survival, which have a harmful effect on the estimation of parameters. Distributions of lifetime data have often positive skewness (asymmetry to the right), which are called life distributions; see Marshall and Olkin (2007). A life distribution known as Birnbaum-Saunders (BS) has received considerable attention due to its good properties. Because of its genesis from fatigue of materials, the BS distribution has been largely applied to engineering, but recent applications range across diverse fields as biology, business, environment and industry, which have been conducted by an international, transdisciplinary group of researchers; see, for example, Bhatti (2010), Leiva et al. (2011, 2014ab, 2015ab, 2016ab, 2017), Villegas et al. (2011), Santana et al. (2011), Ferreira et al. (2012), Paula et al. (2012), Lemonte (2013), Marchant et al. (2013), Saulo et al. (2013, 2018), Kundu (2015), Sanchez et al. (2015), Wanke and Leiva (2015), Santos-Neto et al. (2016), Vanegas and Paula (2016), Bourguignon et al. (2017), Desousa et al. (2017), Garcia-Papani et al. (2017), Leão et al. (2017ab, 2018) and Lillo et al. (2018).

A random variable (RV) with BS distribution can be considered as a transformation of another basis RV distributed as standard normal, that we call normal generator. Generalized BS (GBS) distributions
are obtained by considering different symmetric distributions for the generator; see Sanhueza et al. (2008). It allows us to reach a high flexibility essentially in the kurtosis of GBS distributions. The BS (normal generator), BS-L (logistic generator) and BS-t (Student-t generator) distributions, among others, are particular cases of GBS distributions, which are implemented in the gbs package of the R statistical software; see Barros et al. (2009) and http://www.r-project.org. GBS distributions have several shapes for their FRs, including the unimodal (non-monotone) case; see Bebbington et al. (2008), Kundu et al. (2008), Azevedo et al. (2012) and Athayde (2017), for detailed analyses of the FR of the BS and BS-t distributions. Studies of robustness in the parameter estimation for the BS distribution have been considered by Dupuis and Mills (1998) and Wang et al. (2013, 2015), whereas Paula et al. (2012) considered the robustness in the estimation of parameters for a logarithmic version of the BS-t distribution, but no studies have been considered for the BS-L and BS-t distributions.

The objectives of this paper are (i) to analyze the shape and change-point of the BS-L FR, comparing it to the BS and BS-t FRs; (ii) to propose estimation procedures for these FRs and their change-points, detecting their statistical robustness; and (iii) to evaluate the obtained results numerically with simulated and real-world data.

The remaining of this paper is organized as follows. Section 2 presents preliminary aspects useful to develop our work. In Section 3, shape and change-point analyses of the BS-L FR are conducted, comparing it to the BS and BS-t FRs. In Section 4, we derive estimates of the BS, BS-L and BS-t FRs with the maximum likelihood (ML) and moment methods, including a robustness study. Section 5 evaluates the performance of the proposed estimators with simulated data, analyzing the robustness of the ML estimation procedure. Section 6 illustrates the obtained results with real-world data. In Section 7, we provide the conclusions of this work and some comments on future research related to the topic under study.

2 Preliminary aspects

If an RV $T$ has a BS distribution with shape ($\alpha > 0$) and scale ($\beta > 0$) parameters, we denote it by $T \sim \text{BS}(\alpha, \beta)$. Here, $\beta$ is also the median of the distribution. RVs with BS and N(0, 1) distributions are related by

$$T = \beta(\alpha Z/2 + \sqrt{(\alpha Z/2)^2 + 1})^2 \sim \text{BS}(\alpha, \beta) \quad \text{and} \quad Z = (1/\alpha) \xi(T/\beta) \sim \text{N}(0, 1), \quad (2.1)$$

where $\xi(x) = \sqrt{x} - 1/\sqrt{x}$. If the assumption given in (2.1) is relaxed supposing $Z$ to follow any standard symmetric distribution in $\mathbb{R}$, with probability density function (PDF) $f$ and cumulative distribution function (CDF) $F$, then we obtain the class of GBS distributions. In this case, the notation $T \sim \text{GBS}(\alpha, \beta; f)$ is used and its PDF and CDF are, respectively, $f_T(t) = f(\xi(t/\beta)/\alpha)\psi(t/\beta)/(2\alpha)$ and $F_T(t) = F(\xi(t/\beta)/\alpha)$, for $t > 0$, where $\psi(x) = \sqrt{x} + 1/\sqrt{x}$. Some properties of GBS distributions are: (A1) $cT \sim \text{GBS}(\alpha, c\beta; f)$, for $c > 0$; (A2) $1/T \sim \text{GBS}(\alpha, 1/\beta; f)$; and (A3) $U = (\xi(T/\beta)/\alpha)^2 \sim G\chi^2(1, f)$, that is, $U$ follows the generalized $\chi^2$ distribution with one degree of freedom, which has the $\chi^2(1)$ distribution as a special case if $f$ is the standard normal PDF given by $f(z) = \phi(z) = \exp(-z^2/2) / \sqrt{2\pi}$, for $z \in \mathbb{R}$.

If the RVs $S$ and $V$ follow standard logistic and Student-t (simply t thereafter) distributions, which are symmetrical in $\mathbb{R}$, we use the notations $S \sim \text{L}(0, 1)$ and $V \sim t(\nu)$, where $\nu$ is a shape parameter often corresponding to the degrees of freedom of the t distribution, for $\nu = 1, 2, \ldots$. The PDFs of $S$ and $V$ are, respectively,

$$\phi_L(s) = \frac{\exp(-s)}{1 + \exp(-s))^2} \quad \text{and} \quad \phi_t(v; \nu) = \frac{\Gamma((\nu + 1)/2)}{\sqrt{\nu \pi} \Gamma(\nu/2)} \left(1 + v^2/\nu\right)^{-\nu + 1/2}, \quad s, v \in \mathbb{R}, \nu > 0.$$

Special cases of the t distribution are the Cauchy and normal distributions, when $\nu = 1$ and $\nu \rightarrow +\infty$, respectively.
BS-L and BS-t distributions are obtained replacing $Z \sim N(0,1)$ in (2.1) by $S \sim L(0,1)$ and $V \sim t(\nu)$, respectively, leading to $T = \beta(\alpha S/2 + \sqrt{(\alpha S/2)^2 + 1})^2 \sim \text{BS-L}(\alpha, \beta)$ and $T = \beta(\alpha V/2 + \sqrt{(\alpha V/2)^2 + 1})^2 \sim \text{BS-t}(\alpha, \beta, \nu)$. The logistic distribution has heavier tails than the normal distribution in the sense that $\lim_{t \to \pm\infty} \phi_L(t)/\phi(t) = +\infty$. In addition, its coefficient of kurtosis (CK) is $\beta_2 = 4.2$. As a consequence, the BS-L distribution has heavier tails than the BS distribution, thus having the ability to accommodate atypical cases well, but not in a flexible way as the t distribution does. Note that the t distribution has CK given by $\beta_2 = 3(\nu - 2)/(\nu - 4)$, for $\nu > 4$.

### 3 Shape and change-point analyses of the BS-L failure rate

A good property of the FR is that it allows the behavior of statistical distributions to be characterized. A misspecification of the FR can produce inconveniences in the estimation procedure; see, for example, Bhatti (2010). The FR of $T$ is given in general by $h_T(t) = f_T(t)/(1 - F_T(t))$, for $t > 0$ and $0 < F_T(t) < 1$, where $f_T$ is the PDF of the RV $T$ and $F_T$ is its CDF. The FR of an RV $T$ can be characterized by its total time on test (TTT) function; see details in Appendix A and its theoretical TTT plot in Figure 1 and Athayde (2017). Next, we present the FR of the BS and BS-t distributions and study the FR of the BS-L distribution, including its change-point, denoted by $t_c$, and its limiting behavior.

![Figure 1: Scaled TTT functions for distributions with the indicated FR shape.](image)

Let $T \sim \text{BS}(\alpha, \beta)$. Then, the FR of $T$ is given by

$$h_T(t) = \frac{\phi(\xi(t/\beta)/\alpha)}{2t \alpha \Phi(\xi(t/\beta)/\alpha)}, \quad t > 0,$$

where $\phi$ and $\Phi$ are the N(0, 1) PDF and CDF, respectively. Thus, the following holds: (B1) $h_T(t)$ is unimodal for any $\alpha$, increasing for $t < t_c$, and decreasing for $t > t_c$, where $t_c$ is the change-point of $h_T(t)$; (B2) $\lim_{t \to +\infty} h_T(t) = 1/(2\alpha^2 \beta)$; and (B3) $h_T(t)$ tends to be increasing as $\alpha \to 0$. For more details about (B1)-(B3), see Chang and Tang (1993) and Kundu et al. (2008).

Let $T \sim \text{BS-t}(\alpha, \beta, \nu)$. Then, the FR of $T$ is given by

$$h_T(t; \nu) = \frac{\phi_t(\xi(t/\beta)/\alpha; \nu)}{2t \alpha \Phi_t(-\xi(t/\beta)/\alpha; \nu)}, \quad t > 0, \nu > 0,$$

where $\phi_t$ and $\Phi_t$ are the t PDF and CDF, respectively. Thus, the following holds: (C1) $\lim_{t \to 0} h_T(t; \nu) = \lim_{t \to 0} \phi_t(\xi(t/\beta)/\alpha; \nu)$ and equals $+\infty$ if $\nu = 1$, $\alpha^2/(2\beta)$ if $\nu = 2$, and zero if $\nu > 2$; also
Remark 1. Let see Figure 2. An approximation for the change-point $t$ can be obtained as in Kundu et al. (2008), for $\alpha$ different from the BS case, where the change-point is a decreasing function of shape of $h_\alpha$. Ng et al. (2003) obtained modified moment (MM) estimates of $\beta$, and $\alpha > 1.2649$; and (C4) for $\nu > 2$, $h_T(t; \nu)$ has an IBT shape for $\alpha^2 > 1/\nu$. For more details about (C1)-(C4), see Azevedo et al. (2012).

Let $T \sim BS-L(\alpha, \beta)$. Then, the FR of $T$ is given by

$$h_T(t) = \frac{\phi_L((t/\beta)/\alpha)\psi(t/\beta)}{2t\alpha \Phi_L(-((t/\beta)/\alpha))} = \frac{\psi(t/\beta)}{2t\alpha (1 + \exp(-((t/\beta)/\alpha)))}, \quad t > 0,$$

where $\phi_L$ and $\Phi_L$ are the standard logistic PDF and CDF, respectively. As $\beta$ is a scale parameter, the shape of $h_T$ does not depend on $\beta$ in the sense that, for the same value of $\alpha$, two FRs for different values of $\beta$ differ only by a scale factor, since $h_{aT}(t) = h_T(t/\alpha)/a$, for any $a > 0$. Thus, from now on we suppose $\beta = 1$, without loss of generality. First, it is easy to prove that: (D1) $\lim_{t \to 0} h_T(t) = \lim_{t \to +\infty} h_T(0)$; and (S2) $h_T'(t) = 0$ if and only if $f_\alpha(t) = 0$, where

$$f_\alpha(t) = 1 + \exp\left(\frac{t - 1}{\alpha \sqrt{t}}\right) - \frac{(t + 1)^2}{\alpha \sqrt{t}(t + 3)}.$$

Therefore, the following theorem holds, with its proof being given in Appendix B.

**Theorem 1.** Let $T \sim BS-L(\alpha, \beta)$. Then, $h_T(t)$, with $t > 0$, is unimodal for all $\alpha > 0$.

**Remark 1.** Let $t_c = t_c(\alpha)$ denote the change-point of the BS-L FR, that is, $t_c(\alpha)$ is the solution in $t$ of $f_\alpha(t) = 0$. Then, it can be proved that $t_c(\alpha)$ increases if $\alpha \in [0, \alpha_0]$, and it decreases if $\alpha \in [\alpha_0, +\infty]$. Here, $\alpha_0 = ((4x - 4x^2 + 8 - 8\sqrt{2 - x}/(\alpha^2(x - 2)(x + 2)))^{1/2} \approx 0.2383$, where $x \approx 1.2785$ is the unique solution of $\exp(x) = 1/(x - 1)$. Moreover, $t_c(\alpha_0) \approx 1.3545$; see Figure 3 (left). This is quite different from the BS case, where the change-point is a decreasing function of $\alpha$.

Similarly to the BS and BS-t cases, the change-point of the BS-L($\alpha, \beta$) FR has no closed-form. For several values of $\alpha$, the turning points have been calculated by a standard root solving technique; see Figure 2. An approximation for the change-point $t_c(\alpha)$ (with $\beta = 1$, without loss of generality) can be obtained as in Kundu et al. (2008), for $\alpha > 1.4$, fitting a first-order linear function to $y^{-1/2}$ depending on $\alpha$. This approximation is given by $t_c(\alpha) \approx 1/(3.1683\alpha - 0.3235)^2$. Moreover, for $\beta \neq 1$, the change-point is $t_c(\alpha, \beta) = \beta t_c(\alpha)$, due to the fact that $\beta$ is a scale parameter, and thus a similar approximation is possible. As an example, for $\beta = 10$, it is easy to check that the approximation is $t_c(\alpha, 10) \approx 1/(1.0019\alpha - 0.1023)^2 \approx 10/(3.1683\alpha - 0.3235)^2$.

## 4 Estimation and robustness of the FR and its change-point

Let $T \sim BS-L(\alpha, \beta)$. Then, the ML estimates of $\alpha$ and $\beta$ are given from

$$\hat{\alpha} = \left(\frac{1}{n} \sum_{i=1}^n w_i \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right)^{1/2} \quad \text{and} \quad \hat{\beta} = \left(\frac{1}{2\alpha^2} \sum_{i=1}^n w_i t_i + \frac{n}{2\beta} + \frac{\sum_{i=1}^n w_i t_i}{\sum_{i=1}^n \frac{1}{t_i + \beta}}\right)^{1/2}, \quad (4.1)$$

where $w_i = \tanh((\xi(t_i)/\beta)/(2\alpha))((\xi(t_i)/\beta)/\alpha)$. An iterative procedure to solve the equations in (4.1) has been implemented in R code, which is available under request from the authors and not in an appendix due to restrictions of space. The iterative procedure considers the following as starting values. Ng et al. (2003) obtained modified moment (MM) estimates of $\alpha$ and $\beta$ for the BS distribution, taking
the mean of $T$ and the mean of $1/T$ (instead of $T^2$) by using property (A2). For the BS-t and BS-L distributions, the corresponding MM estimates are

$$\tilde{\alpha} = \left( \frac{2}{u_1} \left( \left( \frac{1}{n} \sum_{i=1}^{n} t_i \right)^{1/2} \right) - 1 \right)^{1/2} \quad \text{and} \quad \tilde{\beta} = \left( \frac{1}{n} \sum_{i=1}^{n} t_i \left( \frac{1}{n} \sum_{i=1}^{n} \frac{1}{t_i} \right)^{-1} \right)^{1/2},$$

where for the BS-t case, $u_1 = E(U) = \nu/(\nu - 2)$, with $U = V^2$, $V \sim t(\nu)$ and $\nu > 2$; whereas for the BS-L case, $u_1 = E(U) = \pi^2/3 \approx 3.2899$, with $U = S^2$ and $S \sim L(0, 1)$. Once the parameters $\alpha$ and $\beta$ are estimated by the ML method, considering as mentioned the MM estimates as starting values, we can use the invariance property of the ML estimators to obtain the estimators of the FR and its change-point. Note that, for both ML and MM methods, the asymptotic joint distribution of the estimators of $\alpha$ and $\beta$ of GBS distributions is bivariate normal; see Engelhardt et al. (1981) for the ML estimators in the BS case, Ng et al. (2003) for the MM estimators in the BS case and Sanhueza et al. (2008) for ML and MM estimators in the general case of GBS distributions, under regularity conditions (for example, that $\nu > 4$ in the BS-t case). The delta method was applied by Kundu et al. (2008) to obtain a normal asymptotic distribution for the change-point of the BS FR, based on MM estimators. Actually, it can be proved that the estimated change-point, based on MM and ML estimators, are also asymptotically normal, following the same procedures used by Ng et al. (2003) and Kundu et al. (2008) for the BS case. We illustrate this procedure for the BS-L case in Appendix D.

An estimation procedure is qualitatively robust if the corresponding influence function (IF), defined by

$$IF(t, \theta; f) = -\left( E\left( \frac{\partial}{\partial \theta} \Psi(t, \theta; f) \right) \right)^{-1} \Psi(t, \theta; f), \quad \text{with} \quad \theta = (\theta_1, \ldots, \theta_k)^T,$$ (4.2)

is bounded, where $\Psi$ is such that $\sum_{i=1}^{n} \Psi(t_i, \theta, f) = 0$ produces the ML estimate of $\theta$ and $f$ is the corresponding PDF. For the three distributions under study, we prove that the indicated limits in Table 1 hold.

To obtain the IF given in (4.2), for $\alpha$ and $\beta$ of the BS and BS-L distributions, and for $\alpha$, $\beta$ and $\nu$ of the BS-t distribution, we first compute the derivatives of $\Psi$ given in (4.2), then its expected value, the
corresponding inverse matrix and finally we multiply by $\Psi$; see derivatives of $\Psi$, as well as the IFs of BS, BS-L and BS-t distributions in Appendix C. Thus, based on these results, we have the following.

Let $T \sim \text{BS}(\alpha, \beta)$ and $\theta = (\alpha, \beta)$. Then, the corresponding IF is proportional to its score, namely $\Psi$, since $E(\partial \Psi(t, \theta; \phi)/\partial \theta)$ is a diagonal matrix. Now, as $t \to 0$, we have $\Psi(t, \alpha; \phi) \to +\infty$ and $\Psi(t, \beta; \phi) \to -\infty$; and, as $t \to +\infty$, we have $\Psi(t, \alpha; \phi) \to +\infty$ and $\Psi(t, \beta; \phi) \to +\infty$. Hence, $\Psi(t, \alpha; \phi)$ and $\Psi(t, \beta; \phi)$ are unbounded, and therefore, their IFs are unbounded as well. This confirms that the estimation procedure based on the BS distribution is non-robust, providing estimates that are sensitive to atypical cases, which is consequent with what occurs in the case of the normal generator; see Leiva et al. (2010) and Paula et al. (2012).

Let $T \sim \text{BS-t}(\alpha, \beta, \nu)$ and now $\theta = (\alpha, \beta, \nu)$. Then, the functions $\Psi(t, \alpha; \phi_1)$ and $\Psi(t, \beta; \phi_1)$ are bounded, whereas $\Psi(t, \nu; \phi_1)$ is unbounded. In fact, we have the following limits as $t \to 0$: $\Psi(t, \alpha; \phi_1) \to \nu/\alpha$, $\Psi(t, \beta; \phi_1) \to -\nu/(2\beta)$ and $\Psi(t, \nu; \phi_1) \to -\infty$; and as $t \to +\infty$: $\Psi(t, \alpha; \phi_1) \to \nu/\alpha$, $\Psi(t, \beta; \phi_1) \to \nu/(2\beta)$ and $\Psi(t, \nu; \phi_1) \to +\infty$. Therefore, the IF for $\beta$ is bounded, since it is proportional to $\Psi(t, \beta; \phi_1)$. However, the IFs for $\alpha$ and $\nu$ are unbounded, because they are linear combinations of $\Psi(t, \alpha; \phi_1)$ and $\Psi(t, \nu; \phi_1)$ (both unbounded). Indeed, both IF functions tend to $+\infty$ as $t$ goes to zero or $+\infty$. Notice that if $\nu$ is known, then the IFs for $\alpha$ and $\beta$ are bounded, because they are proportional to $\Psi(t, \alpha; \phi_1)$ and $\Psi(t, \beta; \phi_1)$, respectively. In this way, the procedure of estimation for $\beta$ based on the BS-t distribution is qualitatively robust. When $\nu$ is estimated, the procedure of estimation for $\alpha$ based on this distribution is sensitive to atypical cases and, as a consequence, it is non-robust. Hence, the unboundedness of the IF of the ML estimator of $\alpha$ crucially depends on the fact that $\nu$ is estimated rather than fixed, due to the unboundedness of the score associated with $\nu$. This is because if $\nu$ is assumed to be known instead of estimated, the IF for $\alpha$ is proportional to its score and functionally independent of $\Psi(t, \nu, \phi_1)$. Therefore, we have demonstrated the non-robustness of the ML estimation procedure using the BS distribution and highlighted that this procedure based on the BS-t distribution is robust only when $\nu$ is fixed, which is consequent with what occurs in the case of the t generator; see Lucas (1997) and Paula et al. (2012).

Let $T \sim \text{BS-L}(\alpha, \beta)$ and again $\theta = (\alpha, \beta)$. Then, the functions $\Psi(t, \alpha; \phi_L)$ and $\Psi(t, \beta; \phi_L)$ are unbounded and $E(\partial \Psi(t, \theta; \phi)/\partial \theta)$ is a diagonal matrix, such as in the BS case, which implies that the ML estimation procedure is robust only in the BS-t case, when $\nu$ is fixed. Thus, although the BS-L distribution has a heavier tail than the BS distribution, the estimation procedure based on the ML method is not robust. This is consequent with what occurs in the case of the logistic generator; see Balakrishnan (2013).

<table>
<thead>
<tr>
<th>Function</th>
<th>Limit as $t \to 0$</th>
<th>Limit as $t \to +\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi(t, \alpha; \phi)$</td>
<td>$+\infty$</td>
<td>$+\infty$</td>
</tr>
<tr>
<td>$\Psi(t, \beta; \phi)$</td>
<td>$-\infty$</td>
<td>$+\infty$</td>
</tr>
<tr>
<td>$\Psi(t, \alpha; \phi_1)$</td>
<td>$\nu/\alpha$</td>
<td>$\nu/\alpha$</td>
</tr>
<tr>
<td>$\Psi(t, \beta; \phi_1)$</td>
<td>$-\nu/(2\beta)$</td>
<td>$2\beta$</td>
</tr>
<tr>
<td>$\Psi(t, \nu; \phi_1)$</td>
<td>$-\infty$</td>
<td>$+\infty$</td>
</tr>
<tr>
<td>$\Psi(t, \alpha; \phi_L)$</td>
<td>$+\infty$</td>
<td>$+\infty$</td>
</tr>
<tr>
<td>$\Psi(t, \beta; \phi_L)$</td>
<td>$-\infty$</td>
<td>$+\infty$</td>
</tr>
</tbody>
</table>
5 Simulation study

We use the R software and $R = 5000$ Monte Carlo simulations to evaluate the behavior of the ML and MM estimators of the change-point $t_c$ of the BS-L FR. A similar study that compares BS and BS-t FRs is presented in Azevedo et al. (2012). Our simulation considers a sample size $n \in \{30, 50, 75, 100, 500\}$, a BS-L parameter $\alpha \in \{0.5, 1.0, 1.5, 2.0\}$ and $\beta = 1.0$, without loss of generality, because as mentioned $\beta$ is a scale parameter. To analyze the point estimation results, we compute, for each $n$ and $\alpha$, the empirical mean, bias and mean squared error (MSE) of the change-point estimators of the BS-L FR. For interval estimation, we compute the empirical coverage probabilities (CPs) from the relative frequencies, at which the true parameter value belongs to the confidence interval (CI). Bootstrap CIs are computed with Algorithm 1. We consider $B = 5000$ bootstrap replications. Tables 2 and 3 show that, in general, as $n$ and $\alpha$ increase, the bias significantly decreases (in absolute value), for both ML and MM estimators, as expected. The same occurs with the MSE. The ML method performs better than the MM method. In addition, the CPs decrease with $\alpha$ and increase with $n$, which are closer to the nominal level of 95% for the ML method.

Algorithm 1 Bootstrap CIs for the BS-L distribution

1: Fix $\alpha$ and $\beta$ and generate $n$ data $t_1, \ldots, t_n$ from the BS-L distribution.
2: Estimate $\alpha$ and $\beta$ and then $t_c$ as $\hat{t}_c$ by the ML method with a quasi-Newton technique implemented by the optim command of the R software with $t_1, \ldots, t_n$.
3: Collect a bootstrap sample from $t_1, \ldots, t_n$.
4: Compute the bootstrap estimate of the change-point $\hat{t}_c^*$. $\hat{t}_c^*$.
5: Repeat steps 2 and 3 $B$ times (bootstrap replications) and obtain $\hat{t}_c^*, \ldots, \hat{t}_c^{**}$.
6: Calculate a CI of $\rho \times 100\%$ for $t_c$ with the percentile bootstrap method as $CI(t_c) = [\hat{t}_c^*(\frac{\rho}{2}); \hat{t}_c^*(1-\frac{\rho}{2})]$.
7: Repeat steps 1 to 6 $R$ times (Monte Carlo replications) and obtain the empirical CPs from the relative frequencies, at which the true parameter value belongs to the CI.

Next, we demonstrate the robustness of the ML estimation with a simulated data set based on survival times of $n = 72$ pigs; see Kundu et al. (2008). We generated 72 cases from a BS-t($0.61, 75.6, 3$) distribution, which parameters were chosen suitably; see data sets in Appendix E. Table 4 (where SD, CV and CS stand for standard deviation and coefficients of variation and skewness, respectively) provides a descriptive summary of these data. Figure 3 presents (left) the frequency plot by kernel estimation (see Saulo et al., 2013), (center) boxplot and (right) TTT plot for these data. From the results in Table 4 and Figure 3 (left), a positively skewed distribution (CS = 4.30) seems to be appropriate. In addition, a high level of kurtosis (CK = 26.98) indicates a heavy-tailed positively skewed distribution generating the data. From Figure 3 (center), some atypical data are detected by the boxplot on the upper tail. In Figure 3 (right), note that the empirical scaled TTT plot indicates that a $\cap$-shaped FR seems to be appropriate for describing the data; see theoretical TTT plot in Figure 1. Therefore, based on the mentioned points, we realize BS-t or BS-L distributions can be good candidates for modeling these data, since they take into account their degrees of skewness and kurtosis, as well as the shape of their FR.

We compute the MM and ML estimates of the change-point of the corresponding FR, as well as the fitted log-likelihood (log-L), such as Kundu et al. (2008) and Azevedo et al. (2012) did for the BS and BS-t models, respectively. The results for the ML case are summarized in Table 5. Figure 2 shows the empirical and fitted survival functions (SF) and the estimated FR with the ML method for the BS, BS-t and BS-L distributions, as well as the true SF and FR. Similar calculations were obtained with the MM method, whose change-points are estimated as 51.40, 92.47 and 79.64 for the BS, BS-t and BS-L distributions, respectively. The change-point for the true distribution is 84.26. In Figure 6 we present the graphical plots of diagnostics based on local influence, which show the robustness of the ML estimation procedure when the BS-t distribution with $\nu$ known is used; see Barros et al. (2009) for...
Table 2: Empirical mean, bias and MSE of the ML and MM estimators of $t_c$ for the indicated values with simulated data.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha$</th>
<th>$t_c$</th>
<th>MM method</th>
<th>ML method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td>Bias</td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>1.000</td>
<td>1.0222</td>
<td>0.0222</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1417</td>
<td>0.1924</td>
<td>0.0508</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0649</td>
<td>0.0132</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0336</td>
<td>0.0062</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>1.000</td>
<td>1.0167</td>
<td>0.0167</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1417</td>
<td>0.1708</td>
<td>0.0292</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0594</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0311</td>
<td>0.0036</td>
</tr>
<tr>
<td>75</td>
<td>0.5</td>
<td>1.000</td>
<td>1.0084</td>
<td>0.0084</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1417</td>
<td>0.1594</td>
<td>0.0178</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0564</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0297</td>
<td>0.0022</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>1.000</td>
<td>1.0076</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1417</td>
<td>0.1580</td>
<td>0.0131</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517744</td>
<td>0.0554</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274056</td>
<td>0.0292</td>
<td>0.0017</td>
</tr>
<tr>
<td>500</td>
<td>0.5</td>
<td>1.000</td>
<td>1.0003</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1417</td>
<td>0.1438</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517744</td>
<td>0.0523</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274056</td>
<td>0.0277</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

details on the local influence method.

Figure 3: PDF plot with kernel estimation (left), boxplot (center) and TTT plot (right) for simulated data.

Algorithm 2 provides goodness-of-fit (GOF) graphical tools based on the Kolmogorov-Smirnov (KS) test to evaluate whether a GBS distribution fits the data set well; see Barros et al. (2014). The KS test is associated with the probability versus probability (PP) plot. We consider p-values of the KS
Table 3: Empirical CP and lower (LL) and upper limits (UL) of 95% CIs of $t_c$ for the indicated value with simulated data.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha$</th>
<th>$t_c$</th>
<th>MM method</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>LL</td>
<td>UL</td>
<td>CP(%)</td>
<td>LL</td>
<td>UL</td>
<td>CP(%)</td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>1.0000000</td>
<td>0.7329</td>
<td>1.2886</td>
<td>88.0</td>
<td>0.6527</td>
<td>1.2808</td>
<td>89.9</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1416738</td>
<td>0.1090</td>
<td>0.5531</td>
<td>80.3</td>
<td>0.0908</td>
<td>0.5256</td>
<td>89.3</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0408</td>
<td>0.1710</td>
<td>79.6</td>
<td>0.0347</td>
<td>0.1597</td>
<td>89.1</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0218</td>
<td>0.0794</td>
<td>79.1</td>
<td>0.0186</td>
<td>0.0744</td>
<td>89.1</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.0000000</td>
<td>0.7545</td>
<td>1.2500</td>
<td>86.2</td>
<td>0.7029</td>
<td>1.2391</td>
<td>92.2</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1416738</td>
<td>0.1051</td>
<td>0.3952</td>
<td>85.2</td>
<td>0.0940</td>
<td>0.3724</td>
<td>91.9</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0399</td>
<td>0.1169</td>
<td>84.9</td>
<td>0.0361</td>
<td>0.1109</td>
<td>91.8</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0214</td>
<td>0.0573</td>
<td>84.8</td>
<td>0.0193</td>
<td>0.0544</td>
<td>91.8</td>
</tr>
<tr>
<td>75</td>
<td>0.5</td>
<td>1.0000000</td>
<td>0.7735</td>
<td>1.2159</td>
<td>88.4</td>
<td>0.7459</td>
<td>1.2063</td>
<td>92.7</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1416738</td>
<td>0.1047</td>
<td>0.3108</td>
<td>87.7</td>
<td>0.0978</td>
<td>0.2976</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0399</td>
<td>0.0955</td>
<td>87.5</td>
<td>0.0375</td>
<td>0.0919</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0214</td>
<td>0.0480</td>
<td>87.4</td>
<td>0.0201</td>
<td>0.0462</td>
<td>92.8</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>1.0000000</td>
<td>0.7941</td>
<td>1.1949</td>
<td>89.8</td>
<td>0.7800</td>
<td>1.1862</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1416738</td>
<td>0.1060</td>
<td>0.2728</td>
<td>89.2</td>
<td>0.1017</td>
<td>0.2632</td>
<td>92.8</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0404</td>
<td>0.0864</td>
<td>88.6</td>
<td>0.0389</td>
<td>0.0837</td>
<td>92.8</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0217</td>
<td>0.0440</td>
<td>88.5</td>
<td>0.0208</td>
<td>0.0426</td>
<td>92.9</td>
</tr>
<tr>
<td>500</td>
<td>0.5</td>
<td>1.0000000</td>
<td>0.8942</td>
<td>1.096</td>
<td>94.1</td>
<td>0.9003</td>
<td>1.0928</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.1416738</td>
<td>0.1182</td>
<td>0.1811</td>
<td>94.0</td>
<td>0.1189</td>
<td>0.1791</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0517645</td>
<td>0.0444</td>
<td>0.0631</td>
<td>93.9</td>
<td>0.0446</td>
<td>0.0625</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>0.0274099</td>
<td>0.0237</td>
<td>0.0330</td>
<td>93.9</td>
<td>0.0238</td>
<td>0.0327</td>
<td>94.4</td>
</tr>
</tbody>
</table>

Table 4: Descriptive statistics for simulated data.

<table>
<thead>
<tr>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>CV</th>
<th>CS</th>
<th>CK</th>
<th>Range</th>
<th>Min.</th>
<th>Max.</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>120.70</td>
<td>80.74</td>
<td>147.28</td>
<td>122%</td>
<td>4.30</td>
<td>26.98</td>
<td>1075.20</td>
<td>10.23</td>
<td>1085.43</td>
<td>72</td>
</tr>
</tbody>
</table>

To examine the goodness of fit of these distributions to the data; see Table 5. Figure 5 shows the good fitting of the mentioned distributions to the data by PP plots with 95% acceptance bands constructed using the KS test. Here, the BS-t distribution is the best (see KS p-value in Table 5) among the BS, BS-t and BS-L distributions, as expected due to that the data were simulated from such a distribution. This good fitting is corroborated by CDF plots of GBS distributions shown in Figure 4.

Table 5: Indicated model parameter and change-point estimate, KS p-value and log-L for simulated data.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\nu$ (fixed)</th>
<th>$\hat{t}_c$</th>
<th>KS p-value</th>
<th>log-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS</td>
<td>0.9396</td>
<td>84.1955</td>
<td>–</td>
<td>51.66</td>
<td>0.0725</td>
<td>-408.8250</td>
</tr>
<tr>
<td>BS-t (ν fixed)</td>
<td>0.6474</td>
<td>79.6503</td>
<td>3</td>
<td>85.59</td>
<td>0.5331</td>
<td>-406.4600</td>
</tr>
<tr>
<td>BS-L</td>
<td>0.5024</td>
<td>81.3337</td>
<td>–</td>
<td>80.81</td>
<td>0.1927</td>
<td>-406.6249</td>
</tr>
<tr>
<td>True model</td>
<td>0.6100</td>
<td>75.6000</td>
<td>3</td>
<td>84.26</td>
<td>–</td>
<td>-406.7489</td>
</tr>
</tbody>
</table>

Observations with a potential influence detected by the total local influence method are usually analyzed by the relative change (RC) of each parameter estimate. This is obtained by dropping the influential cases and re-estimating the distribution parameters according to the expression $RC(\hat{\theta}(i)) =$
Algorithm 2 GOF test for GBS distributions

1: Collect data $t_1, \ldots, t_n$ and order them as $t_{1:n}, \ldots, t_{n:n}$.
2: Estimate $\alpha$ and $\beta$ of a GBS distribution by $\hat{\alpha}$ and $\hat{\beta}$, respectively, with $t_1, \ldots, t_n$.
3: Compute $\hat{u}_{j:n} = F_T(t_{j:n}; \hat{\alpha}, \hat{\beta})$, for $j = 1, \ldots, n$, with $F_T$ being the GBS CDF.
4: Calculate $\hat{\gamma}_j = \Phi^{-1}(\hat{u}_{j:n})$, where $\Phi^{-1}$ is the inverse function of the $N(0, 1)$ CDF.
5: Define $\hat{\gamma}_j = \Phi^{-1}(\hat{u}_{j:n})$, with $\hat{\gamma}_j = (\hat{\gamma}_j - \bar{\gamma})/s_Y$, where $\bar{\gamma} = \sum_{j=1}^n \hat{\gamma}_j/n$, $s_Y = (\sum_{j=1}^n (\hat{\gamma}_j - \bar{\gamma})^2/(n-1))^{1/2}$.
6: Draw the PP plot with points $w_{j:n} = (2j-1)/(2n)$ versus $\hat{u}_{j:n}$, for $j = 1, \ldots, n$.
7: Construct acceptance bands according to $(\max\{w - d\zeta + 1/(2n), 0\}, \min\{w + d\zeta - 1/(2n), 1\})$, where $d\zeta$ is the $100\times\zeta$th percentile of the KS distribution and $w$ is a continuous version of $w_{j:n}$.
8: Determine the p-value of the KS statistic and reject the null hypothesis of a GBS distribution for the specified significance level based on this p-value.
9: Corroborate coherence between steps 7 and 8.

Figure 4: Empirical and fitted SF (left) and estimated FR (right) for the indicated model with simulated data.

<table>
<thead>
<tr>
<th>DC</th>
<th>$\hat{\alpha}$</th>
<th>RC</th>
<th>$\hat{\beta}$</th>
<th>RC</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>0.9396</td>
<td>84.20</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>0.8967</td>
<td>4.56</td>
<td>87.84</td>
<td>4.33</td>
</tr>
<tr>
<td>72</td>
<td>0.8566</td>
<td>8.83</td>
<td>78.34</td>
<td>6.96</td>
</tr>
</tbody>
</table>

Table 6: ML estimates and RC (in %) for the indicated parameter, dropped case (DC) and model with simulated data.
Figure 5: PP plots with 95% acceptance bands for the BS (left), BS-t with $\nu = 3$ fixed (center) and BS-L (right) distributions using simulated data.

Figure 6: Local influence index plot for BS (left), BS-L (center) and BS-t with $\nu = 3$ fixed (right) distributions using simulated data.
6 Examples with real-world data

We illustrate the potential use of the proposed models for practical purposes by means of the following three data sets, coming from different areas of application and detailed in Appendix E:

(S1) Survival times (in days) of \( n = 72 \) pigs injected with the same dose of tubercle bacilli, corresponding to \( 4.0 \times 10^6 \) bacillary units per 0.5 ml (\( \log_{10}(4.0 \times 10^6) = 6.6 \)), that is, to a regimen number of the base-10 logarithm of bacillary units in 0.5 ml of challenge solution, denoted by \( \text{survpig66} \), see Kundu et al. (2008).

(S2) Claim amounts corresponding to \( n = 542 \) injuries paid by an insurance Australian company, denoted by \( \text{amount} \), see Paula et al. (2012).

(S3) Total phosphorus (mg/l) in Melides lagoon, Portugal, measured from 05-April-2004 to 26-Jan-2013 with \( n = 104 \), denoted by \( \text{phosphorus} \), see http://www.snirh.pt (Portuguese National System of Information for Water Resources).

Table 7 provides a descriptive summary of the three data sets. Figure 8 presents (left) PDF plots by kernel estimation, (center) boxplot and (right) TTT plot from these data sets. We first conduct exploratory data analyses (EDA) considering that distributions skewed to the right (see CS in Table 7) with positive support seem to be appropriate. In addition, a high level of kurtosis (see CK in Table 7), as mentioned, indicates a heavy right tail. Moreover, from Figure 8(right), observe that the empirical scaled TTT plots suggest distributions with \( \cap \)-shaped or nearly \( \cap \)-shaped FR seem to be appropriate for describing these data; see theoretical TTT plot in Figure 1. Therefore, we realize GBS distributions (such as BS, BS-t or BS-L) can be good candidates for modeling these data, since they take into account their degrees of skewness and kurtosis, as well as the shape of their FR. Note that BS-L FRs are IBT, such as BS and most BS-t (\( \nu \geq 3 \)) FRs. This includes the nearly-IBT case, which leads to TTT plots with a clearly convex graph on the left side, and then a concave or nearly linear graph at the right side, as in the data sets S1, S2 and S3. In Figure 7, we illustrate this fact with three TTT plots for simulated data (\( n = 1000 \)) from a BS-t with \( \alpha = 0.75 \) and \( \nu = 3, 5 \) and 7, respectively, showing a wide variety of IBT shapes.

Table 7: Descriptive statistics for the indicated data set.

<table>
<thead>
<tr>
<th>Set</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>CV</th>
<th>CS</th>
<th>CK</th>
<th>Range</th>
<th>Min.</th>
<th>Max.</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>99.82</td>
<td>70.00</td>
<td>81.12</td>
<td>81.27%</td>
<td>1.76</td>
<td>5.46</td>
<td>364.0</td>
<td>12</td>
<td>376.0</td>
<td>72</td>
</tr>
<tr>
<td>S2</td>
<td>8996.41</td>
<td>6765.25</td>
<td>8790.71</td>
<td>97.71%</td>
<td>5.58</td>
<td>51.11</td>
<td>116477.7</td>
<td>109</td>
<td>116586.7</td>
<td>542</td>
</tr>
<tr>
<td>S3 (( \times 100 ))</td>
<td>13.93</td>
<td>9.95</td>
<td>14.93</td>
<td>107.1%</td>
<td>3.53</td>
<td>16.68</td>
<td>98.1</td>
<td>1</td>
<td>99.1</td>
<td>104</td>
</tr>
</tbody>
</table>

For S1, S2 and S3, we compute the ML estimates of the BS, BS-t and BS-L distribution parameters, modified KS statistic (see Chen and Balakrishnan, 1995; D’Agostino and Stephens, 1986) and fitted log-L, such as Kundu et al. (2008), Azevedo et al. (2012) and Paula et al. (2012) did for the BS and/or BS-t models. The results for the ML method are summarized in Table 8. Figure 9 shows the empirical and fitted SFs and estimated FR based on the ML method for the BS, BS-t and BS-L models with the analyzed data sets. These results show a better performance of the BS-t distribution in all data sets; see KS p-values in Table 8 and PP plots with 95% acceptance bands in Figure 9. Once the best model is selected, we can use the estimated change-point of its FR for different purposes. For example, in the case of the data set S1, based on this estimate, we can establish the reduction in the dose, and therefore, in the cost of the treatment.

As an example, for the sample \( \text{survpig66} \), we have estimated the asymptotic standard errors of the corresponding ML and MM estimators, \( \hat{t}_c \) and \( \tilde{t}_c \), for the three models; see Table 9. In the BS-L case, these results agree with the simulation results obtained in Table 2 for \( n = 75, \alpha = 0.5 \) and \( \beta = 1.0 \), where the empirical MSEs are 0.0161 and 0.0177 for the ML and MM methods, respectively. Then,
Table 8: Parameter and change-point ML estimates, KS statistic and p-value, as well as log-L for the indicated model and data.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\nu$ (fixed)</th>
<th>$\hat{t}_c$</th>
<th>KS p-value</th>
<th>log-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S1) survpig66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS</td>
<td>0.760</td>
<td>77.535</td>
<td>–</td>
<td>90.4</td>
<td>0.039</td>
<td>-390.917</td>
</tr>
<tr>
<td>BS-t</td>
<td>0.608</td>
<td>75.588</td>
<td>5</td>
<td>90.2</td>
<td>0.177</td>
<td>-390.054</td>
</tr>
<tr>
<td>BS-L</td>
<td>0.415</td>
<td>75.998</td>
<td>–</td>
<td>90.7</td>
<td>0.164</td>
<td>-389.876</td>
</tr>
<tr>
<td>(S2) amount</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS</td>
<td>0.844</td>
<td>6592.406</td>
<td>–</td>
<td>3296.2</td>
<td>0.000</td>
<td>-5438.130</td>
</tr>
<tr>
<td>BS-t</td>
<td>0.509</td>
<td>6853.522</td>
<td>3</td>
<td>3426.7</td>
<td>0.125</td>
<td>-5353.927</td>
</tr>
<tr>
<td>BS-L</td>
<td>0.406</td>
<td>6882.437</td>
<td>–</td>
<td>3441.2</td>
<td>0.000</td>
<td>-5372.379</td>
</tr>
<tr>
<td>(S3) phosphorus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS</td>
<td>0.787</td>
<td>0.107</td>
<td>–</td>
<td>0.1130</td>
<td>0.000</td>
<td>118.995</td>
</tr>
<tr>
<td>BS-t</td>
<td>0.499</td>
<td>0.101</td>
<td>3</td>
<td>0.11867</td>
<td>0.533</td>
<td>128.959</td>
</tr>
<tr>
<td>BS-L</td>
<td>0.399</td>
<td>0.102</td>
<td>–</td>
<td>0.12515</td>
<td>0.024</td>
<td>126.060</td>
</tr>
</tbody>
</table>

Figure 7: TTT plots for the BS-t($\alpha = 0.75, \beta = 1, \nu$) with $\nu = 3$ (left), $\nu = 5$ (center) and $\nu = 7$ (right) using simulated data.

using the fact that $t_c(\alpha, \beta) = \beta t_c(\alpha, 1)$, the estimated standard errors of the change-point, based on the ML and MM estimates of $\beta$, employing survpig66 data, are 9.643 and 10.284, respectively. Note that these values are similar to the asymptotic estimated standard errors given in Table 9 for the BS-L case.

Table 9: ML and MM estimates of the change-point (with estimated asymptotic standard errors in parenthesis) for survpig66 data.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{t}_c$</th>
<th>$\tilde{t}_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS</td>
<td>90.3951 (24.37)</td>
<td>90.2996 (25.13)</td>
</tr>
<tr>
<td>BS-t ($\nu = 5$)</td>
<td>90.1533 (8.71)</td>
<td>93.7848 (10.32)</td>
</tr>
<tr>
<td>BS-L</td>
<td>90.6961 (9.40)</td>
<td>91.8113 (10.34)</td>
</tr>
</tbody>
</table>
Figure 8: PDF plot by using kernel estimation (left), boxplot (center) and TTT plot (right) for S1 (first panel), S2 (second panel) and S3 (third panel) data sets.

7 Concluding remarks and future research

We have studied the shape of the failure rate of the Birnbaum-Saunders-logistic distribution, which is unimodal (upside down-shaped). Its change-point is an important value, because the hazard begins to decrease at this point. Such a change-point can be obtained as a solution of a non-linear equation. We have provided an approximation to this change-point and have shown that it works very well whenever the shape parameter is not too small. We have discussed and compared some robustness issues related to the Birnbaum-Saunders, Birnbaum-Saunders-logistic and Birnbaum-Saunders-t distributions. In addition, we have carried out a simulation to assess the performance of two estimation procedures of the change-point in these three distributions. Finally, we have illustrated the obtained results, including estimation and robustness, by means of data analyses.

We hope to report new findings related to modeling (of fixed and mixed effects), diagnostics and censored data analysis based on generalized Birnbaum-Saunders distributions beyond the normal and Student-t generators. Also, semi-parametric and error-in-variables formulations, as well as non-parametric estimation of kernel can be conducted with this wide family of distributions. Multivariate versions, copula methods and spatio-temporal models may also be addressed based on these distributions, some of
Figure 9: PP plots with 95% acceptance bands for the BS (left), BS-t with $\nu$ fixed according to values in Table 8 (center) and BS-L (right) distribution using S1 (first panel), S2 (second panel) and S3 (third panel) data sets.

which have been already partially considered for Birnbaum-Saunders and Birnbaum-Saunders-t distributions, but not for other members of the generalized Birnbaum-Saunders family; see Vilca et al. (2014), Kundu (2015a,b), Vanegas and Paula (2015), Marchant et al. (2016a,b, 2018) and Garcia-Papani et al. (2017).

Appendix A: TTT plot

The TTT function and its scaled version, if $H^{-1}(1) = E(T)$ is finite, are respectively

$$H^{-1}(u) = \int_0^{F_T^{-1}(u)} (1 - F_T(y)) \, dy \quad \text{and} \quad W(y) = \frac{H^{-1}(y)}{H^{-1}(1)}, \quad 0 \leq y \leq 1,$$
Figure 10: Estimated SF (first panel) and FR (second panel) using S1 (left), S2 (center) and S3 (right) data sets, with \( \nu \) for the BS-t distribution fixed according to values in Table 8.

where \( F_T^{-1} \) is the inverse function of \( F_T \). Now, \( W \) can be empirically approximated allowing the empirical scaled TTT function to be constructed by plotting the consecutive points \((k/n, W_n(k/n))\), where

\[
W_n(k/n) = \frac{\sum_{i=1}^{k} t(i) + (n-k) t(k)}{\sum_{i=1}^{n} t(i)}, \quad k = 1, \ldots, n,
\]

with \( t(i) \) being the corresponding observed \( i \)th order statistic. By means of the empirical scaled TTT function, we can detect the type of FR that the lifetime data have and, as a consequence, the type of distribution that these data can have. From Figure 11, note the different theoretical shapes for the scaled TTT function. For more details about the TTT method, see Aarset (1987). Thus, a TTT function that is concave (or convex) corresponds to the IFR (or DFR) class. A TTT function that is concave (convex) for \( t < t_c \) and convex (concave) for \( t > t_c \) corresponds to an IBT (BT) FR. A TTT plot expressed by a straight line is an indication that the underlying distribution is exponential.

Appendix B: Proof of Theorem 1

For \( \beta = 1.0 \), we have

\[
f_T'(t) = \frac{\sqrt{t} t(1 + \exp(-a_t)) - (1 + t) \left( \frac{3}{2} (1 + \exp(-a_t)) - t \exp(-a_t) A_t \right)}{(1 + \exp(-a_t))^2 t^3},
\]

where

\[
a_t = \frac{\xi(t)}{\alpha} = \frac{t - 1}{\alpha \sqrt{t}} \quad \text{and} \quad A_t = a_t' = \frac{(1 + t)^{3/2} - t^{3/2}}{2\alpha}.
\]

Thus,

\[
f_T'(t) = 0 \iff t(1 + \exp(-a_t)) - (1 + t) \left( \frac{3}{2} (1 + \exp(-a_t)) - t \exp(-a_t) A_t \right) = 0 \iff f_\alpha(t) = 0,
\]
where
\[ f_\alpha(t) = 1 + \exp\left(\frac{t - 1}{\alpha \sqrt{t}}\right) - \frac{(t + 1)^2}{\alpha \sqrt{t} (t + 3)}. \]

In addition, we have
\[ f'_\alpha(t) = \frac{(t + 1)g_\alpha(t)}{2\alpha \sqrt{t} (t + 3)^2}, \]

with
\[ g_\alpha(t) = \exp\left(\frac{t - 1}{\alpha \sqrt{t}}\right)(t + 3) - (t + 3)^2 + 12. \]

Note that: (E1) \( f_\alpha(1) = 2 - 1/\alpha \) (then \( f_{1/2}(1) = 0 \)); (E2) \( \lim_{t \to 0} f_\alpha(t) = -\infty \) and \( \lim_{t \to +\infty} f_\alpha(t) = +\infty \) (then \( f_\alpha(t) = 0 \) for at least one \( t > 0 \)); (E3) if \( \alpha > \delta \), \( f_\alpha(t) > f_\delta(t) \), for \( t \leq 1 \), and \( g_\alpha(t) > g_\delta(t) \), for \( t < 1 \); (E4) \( f_{1/2}(t) > 0 \) if and only if \( g_{1/2}(t) > 0 \); and (E5) if \( t \geq 1 \), \( f_{1/2}'(t) > 0 \), because \( g_{1/2}(t) \geq 12 \).

Thus, if we prove that \( f_{1/2}(t) > 0 \), then: from (E1), it follows that \( f_{1/2}(t) < 0 \), for all \( t \in [0,1] \), and \( f_{1/2}(t) > 0 \), for all \( t \in [1, +\infty] \); if \( \alpha > 1/2 \), \( f_{1/2}'(t) > 0 \) by (E5), for \( t \geq 1 \), and by (E3) and (E4), for \( t < 1 \); consequently \( f_\alpha \) has a unique root, which is necessarily less than one, since \( f_\alpha(1) > 0 \) from (E1); if \( \alpha < 1/2 \), \( f_\alpha \) has exactly one root, because (F3) implies \( f_\alpha(t) < f_{1/2}(t) < 0 \), for \( t \leq 1 \), and \( f_\alpha(t) \) increases for \( t > 1 \), from (E5); also, this root is greater than one. This means that \( f_\alpha \) has a unique root and also \( h'_T \), which proves the theorem. Now, we have to prove that \( f_{1/2}'(t) > 0 \), that is, \( g_{1/2}(t) > 0 \). In fact, it suffices to prove that \( g_{1/2}(t) > 0 \) for \( t < 1 \), according to (E5).

Let \( r(t) = \exp(2(t - 1)/\sqrt{t}) \), for \( t > 0 \). Note that \( r \) is an increasing positive function over \( t \). Then,
\[ g_{1/2}(t) = r(t)(t + 3)^2 - (t + 3)^2 + 12 \geq - (t + 3)^2 + 12 > 0, \]

if \( t < c_1 = -3 + 2\sqrt{3} \approx 0.4641 \). However, if \( t \geq c_1 \), then
\[ g_{1/2}(t) = r(t)(t + 3)^2 - (t + 3)^2 + 12 \geq r(c_1)(t + 3)^2 - (t + 3)^2 + 12 = (r(c_1) - 1)(t + 3)^2 + 12. \]

Thus, \( g_{1/2}(t) \geq 0 \) for \( t < c_2 = -3 + 2\sqrt{3}/\sqrt{1 - r(c_1)} \approx 0.8909 \). If \( t \geq c_2 \), then
\[ g_{1/2}(t) \geq (r(c_2) - 1)(t + 3)^2 + 12 > 0, \]

for \( t < -3 + 2\sqrt{3}/\sqrt{1 - r(c_2)} \approx 4.6260 > 1 \). This completes the proof.

\[ \text{Appendix C: Score functions for proving robustness} \]

**BS distribution**

\[ \Psi(t, \alpha; \phi) = \frac{1}{\alpha^3} \left( \frac{t}{\beta} + \frac{\beta}{t} - 2 \right) - \frac{1}{\alpha}, \quad \Psi(t, \beta; \phi) = \frac{1}{t + \beta} - \left( \frac{1}{2\alpha^2} \right) \left( \frac{1}{t} - \frac{t}{\beta^2} \right) - \frac{1}{2\beta}. \]

**BS-t distribution**

\[ \Psi(t, \alpha; \phi_t) = \left( \frac{\nu + 1}{\nu + (\xi(t/\beta)/\alpha)^2} \right) \left( \frac{1}{\alpha^3} \right) \left( \frac{t}{\beta} + \frac{\beta}{t} - 2 \right) - \frac{1}{\alpha}, \]
\[ \Psi(t, \beta; \phi_t) = \frac{1}{t + \beta} - \frac{1}{2\alpha^2} \left( \frac{\nu + 1}{\nu + (\xi(t/\beta)/\alpha)^2} \right) \left( \frac{1}{t} - \frac{t}{\beta^2} \right) - \frac{1}{2\beta}, \]
\[ \Psi(t, \nu; \phi_t) = \frac{1}{2} \left( \gamma \left( \frac{\nu + 1}{2} \right) - \gamma \left( \frac{\nu}{2} \right) \right) - \log \left( \frac{1 + \frac{(\xi(t/\beta)/\alpha)^2}{2}}{1 + \frac{(\xi(t/\beta)/\alpha)^2}{\nu + (\xi(t/\beta)/\alpha)^2}} \right) - \frac{\nu - (\xi(t/\beta)/\alpha)^2}{\nu + (\xi(t/\beta)/\alpha)^2}. \]
where \( \gamma \) is the digamma function given by \( \gamma(x) = d\log(\Gamma(x))/dx = \Gamma'(x)/\Gamma(x) \). Now, with \( \theta = (\alpha, \beta, \nu) \), we have

\[
E\left( \frac{\partial}{\partial \theta} \Psi(t, \theta; \phi_1) \right) = \begin{pmatrix} a & 0 & d \\ 0 & b & 0 \\ d & 0 & c \end{pmatrix},
\]

(7.1)

where

\[
a = E\left( \frac{\partial}{\partial \alpha} \Psi(t, \alpha; \phi_1) \right) = -\frac{2\nu}{\alpha^2(3 + \nu)}, \quad b = E\left( \frac{\partial}{\partial \beta} \Psi(t, \beta; \phi_1) \right) = \frac{3\alpha^2 - 4(2 + \nu)}{4\alpha^2\beta^2(3 + \nu)} - B\nu,
\]

\[
c = E\left( \frac{\partial}{\partial \nu} \Psi(t, \nu; \phi_1) \right) = -\frac{1}{(1 + \nu)(3 + \nu)} \frac{1}{4\gamma(\nu)} - \frac{1}{4\gamma(\nu + 1)}.
\]

\[
d = E\left( \frac{\partial}{\partial \alpha} \Psi(t, \alpha; \phi_1) \right) = E\left( \frac{\partial}{\partial \nu} \Psi(t, \nu; \phi_1) \right) = \frac{2}{\alpha(3 + 4\nu + \nu^2)},
\]

with \( B\nu = E(1/(\beta + T)^2) \) and \( T \sim \text{BS-t}(\alpha, \beta, \nu) \). Then, the inverse matrix of (7.1) is given by

\[
E\left( \frac{\partial}{\partial \theta} \Psi(t, \theta; \phi_1) \right)^{-1} = \frac{1}{abc - bd^2} \begin{pmatrix} bc & 0 & -bd \\ 0 & ac - d^2 & 0 \\ -bd & 0 & ab \end{pmatrix}.
\]

Therefore, we have

\[
\text{IF}(t, \theta; \phi_1) = -\left[ \frac{1}{ac - d^2} \left[ c\Psi(t, \alpha; \phi_1) - d\Psi(t, \nu; \phi_1) \right] - \frac{1}{b} \Psi(t, \beta; \phi_1) \right].
\]

**BS-L distribution**

\[
\Psi(t, \alpha; \phi_L) = \left( \frac{\tanh(\frac{1}{2}(\xi(\nu)/\alpha))}{\xi(\nu)/\alpha} \right) \frac{1}{\alpha^3} \left( \frac{t}{\beta} + \frac{\beta}{t} - 2 \right) - \frac{1}{\alpha},
\]

\[
\Psi(t, \beta; \phi_L) = \frac{1}{t + \beta} - \frac{1}{2\alpha^2} \left( \frac{\tanh(\frac{1}{2}(\xi(\nu)/\beta))}{\xi(\nu)/\beta} \right) \left( \frac{1}{\beta} - \frac{t}{\beta^2} \right) - \frac{1}{2\beta}.
\]

Here, again \( E[\partial\Psi(t, \theta; \phi)/\partial \theta] \) is a diagonal matrix, but now with \( \theta = (\alpha, \beta) \) and

\[
E\left( \frac{\partial}{\partial \alpha} \Psi(t, \alpha; \phi_L) \right) = -\frac{3 + \pi^2}{9\alpha^2}, \quad E\left( \frac{\partial}{\partial \beta} \Psi(t, \beta; \phi_L) \right) = -\frac{12 + \alpha^2(21 + \pi^2) + \frac{126\zeta(3)}{\pi^2}}{72\alpha^2\beta^2},
\]

where \( \zeta \) denotes the Riemann zeta function, given by \( \zeta(s) = \sum_{k \geq 1} k^{-s} \). Note that \( \zeta(3) \approx 1.20206 \).

Therefore, we have

\[
\text{IF}(t, \theta; \phi_L) = \begin{pmatrix} \frac{9\alpha^2}{3 + \pi^2} \Psi(t, \alpha; \phi_L) \\ \frac{72\alpha^2\beta^2}{20206(21 + \pi^2) + \frac{126\zeta(3)}{\pi^2}} \Psi(t, \beta; \phi_L) \end{pmatrix}.
\]
Appendix D: Asymptotic normality of estimators

For the MM method, as $u_1 = E(U) = \pi^2/3$ and $u_2 = E(U^2) = 7\pi^4/15$ in the BS-L case, it follows that the asymptotic distribution of the MM estimators $\hat{\alpha}$ and $\hat{\beta}$ is bivariate normal with mean $(\alpha, \beta)$ and variance-covariance matrix

$$\Sigma_\theta = \begin{pmatrix} \frac{8}{10} & 0 \\ 0 & \frac{9}{15 (6 + \pi^2 \alpha^2)^2} \end{pmatrix},$$

whereas the estimator of the change-point $\tilde{t}_c$ follows a normal asymptotic distribution with mean $t_c$ and variance $C/n$, where

$$C = \frac{8\alpha^2}{10} \left( \frac{\partial^2 \ell(t, \alpha, \beta)}{\partial \alpha \partial \beta} \right)^2 + \frac{9}{15} \left( \frac{\partial^2 \ell(t, \alpha, \beta)}{\partial \alpha^2} \right)^2 \left( \frac{\partial^2 \ell(t, \alpha, \beta)}{\partial \beta^2} \right)^2.$$

For the ML method, the asymptotic distribution of $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ is bivariate normal with mean $(\alpha, \beta)$ and covariance matrix $\Sigma_\hat{\theta}$. This matrix can be approximated by $-\hat{\ell}^{-1}$, with $-\hat{\ell}$ being evaluated at $\hat{\theta}$ and obtained from the log-likelihood $\ell$; see Sanhueza et al. (2008). The elements of $-\hat{\ell}$ are given by

$$\hat{\ell}_{\alpha\alpha} = \frac{n}{\alpha^2} - \frac{2}{\alpha^6} \sum_{i=1}^{n} v_i \left( \frac{t_i}{\beta} + \beta - 2 \right)^2 - \frac{3}{\alpha^4} \sum_{i=1}^{n} \left( \frac{t_i}{\beta} + \beta - 2 \right),$$

$$\hat{\ell}_{\alpha\beta} = \frac{1}{\alpha^3} \left( \frac{1}{\alpha} - \frac{2}{\alpha^2} \sum_{i=1}^{n} v_i' \left( \frac{t_i}{\beta} - 2 \right) \right),$$

$$\hat{\ell}_{\beta\beta} = \frac{n}{\alpha^2 \beta^2} - \frac{n}{\alpha^2 \beta^2} \sum_{i=1}^{n} \left( \frac{1}{\alpha} \right)^2 - \frac{2}{\alpha^4} \sum_{i=1}^{n} v_i' \left( \frac{t_i}{\beta} - 2 \right)^2,$$

where $v_i$ and $v_i'$ depend on the underlying model, namely, letting $a_i = (\xi(t_i/\beta)/\alpha)^2$. Therefore, (i) for the BS case, $v_i = 1$ and $v_i' = 0$; (ii) for the BS-t case,

$$v_i = \frac{\nu + 1}{\nu + a_i} \quad \text{and} \quad v_i' = -\frac{\nu + 1}{(\nu + a_i)^2};$$

(iii) for the BS-L case,

$$v_i = \frac{\tanh(\sqrt{a_i})}{\sqrt{a_i}} \quad \text{and} \quad v_i' = \frac{\sqrt{a_i} - \sinh(\sqrt{a_i})}{2\sqrt{a_i} (1 + \cosh(\sqrt{a_i}))}.$$

Computing $-\hat{\ell}^{-1}$ and letting its diagonal elements be denoted by $b_{11}$ and $b_{22}$, we obtain the asymptotic normal distribution for the estimator of the change-point $\hat{t}_c$ using the standard delta method, as Kundu et al. (2008) did for the BS case. This asymptotic distribution is $N(t_c, \sigma^2)$ with

$$\sigma^2 = b_{11} \left( \frac{\partial \ell(t, \alpha, \beta)}{\partial \alpha} \right)^2 + b_{22} \left( \frac{\partial \ell(t, \alpha, \beta)}{\partial \beta} \right)^2.$$

Appendix E: Data sets

[S0] Simulated data set 10.23, 13.95, 17.75, 17.76, 22.81, 27.02, 29.69, 30.12, 30.92, 31.21, 34.88, 35.29, 35.90, 37.24, 43.01, 46.69, 46.75, 48.10, 50.28, 53.08, 57.11, 60.73, 63.04, 64.30, 64.58, 64.63, 65.65, 68.03, 68.53, 69.21, 71.34,
Acknowledgements

The authors thank the Editors and the referees for their constructive comments on an earlier version of this manuscript. Emilia Athayde and Assis Azevedo were partially supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the “Fundaçao para a Ciência e a Tecnologia”, through the Project UID/MAT/00013/2013. Michelli Barros was partially supported by the Brazilian Council for Scientific and Technological Development (CNPq in Portuguese). Victor Leiva was partially supported by the Chilean Council for Scientific and Technology Research (Conicyt in Spanish), grant FONDECYT 1160868.

References


