Computational Statistics and Data Analysis 56 (2012) 698–718

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A log-linear regression model for the β-Birnbaum–Saunders distribution with censored data

Edwin M.M. Ortega [a], Gauss M. Cordeiro [b], Artur J. Lemonte [c,*]

[a] Departamento de Ciências Exatas, Universidade de São Paulo, Piracicaba/SP, Brazil
[b] Departamento de Estatística, Universidade Federal de Pernambuco, Recife/PE, Brazil
[c] Departamento de Estatística, Universidade de São Paulo, São Paulo/SP, Brazil

ARTICLE INFO

Article history:
Received 18 April 2011
Received in revised form 25 July 2011
Accepted 16 September 2011
Available online 22 September 2011

Keywords:
Birnbaum–Saunders distribution
Censored data
Cure fraction
Generating function
Lifetime data
Maximum likelihood estimation
Regression model

ABSTRACT

The β-Birnbaum–Saunders (Cordeiro and Lemonte, 2011) and Birnbaum–Saunders (Birnbaum and Saunders, 1969a) distributions have been used quite effectively to model failure times for materials subject to fatigue and lifetime data. We define the log-β-Birnbaum–Saunders distribution by the logarithm of the β-Birnbaum–Saunders distribution. Explicit expressions for its generating function and moments are derived. We propose a new log-β-Birnbaum–Saunders regression model that can be applied to censored data and be used more effectively in survival analysis. We obtain the maximum likelihood estimates of the model parameters for censored data and investigate influence diagnostics. The new location-scale regression model is modified for the possibility that long-term survivors may be presented in the data. Its usefulness is illustrated by means of two real data sets.

1. Introduction

The fatigue is a structural damage that occurs when a material is exposed to stress and tension fluctuations. Statistical models allow to study the random variation of the failure time associated to materials exposed to fatigue as a result of different cyclical patterns and strengths. The most popular model for describing the lifetime process under fatigue is the Birnbaum–Saunders (BS) distribution (Birnbaum and Saunders, 1969a,b). The crack growth caused by vibrations in commercial aircrafts motivated these authors to develop this new family of two-parameter distributions for modeling the failure time due to fatigue under cyclic loading. Relaxing some assumptions made by Birnbaum and Saunders (1969a), Desmond (1985) presented a more general derivation of the BS distribution under a biological framework. The relationship between the BS and inverse Gaussian distributions was explored by Desmond (1986) who demonstrated that the BS distribution is an equal-weight mixture of an inverse Gaussian distribution and its complementary reciprocal. The two-parameter BS model is also known as the fatigue life distribution. It is an attractive alternative distribution to the Weibull, gamma and log-normal models, since its derivation considers the basic characteristics of the fatigue process. Furthermore, it has the appealing feature of providing satisfactory tail fitting due to the physical justification that originated it, whereas the Weibull, gamma and log-normal models typically provide a satisfactory fit in the middle portion of the data, but oftentimes fail to deliver a good fit at the tails, where only a few observations are generally available.

In many medical problems, for example, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric models to estimate univariate survival functions for censored

* Corresponding author.
E-mail address: arturlemonte@gmail.com (A.J. Lemonte).

0167-9473/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2011.09.018

data regression problems are widely used. Different forms of regression models have been proposed in survival analysis. Among them, the location-scale regression model (Lawless, 2003) is distinguished since it is frequently used in clinical trials. Recently, the location-scale regression model has been applied in several research areas such as engineering, hydrology and survival analysis. Lawless (2003) also discussed the generalized log-gamma regression model for censored data. Xie and Wei (2007) developed the censored generalized Poisson regression models, Barros et al. (2008) proposed a new class of lifetime regression models when the errors have the generalized BS distribution, Carrasco et al. (2008) introduced a modified Weibull regression model, Silva et al. (2008) studied a location-scale regression model using the Burr XII distribution and Silva et al. (2009) worked with a location-scale regression model suitable for fitting censored survival times with bathtub-shaped hazard rates. Ortega et al. (2009a,b) proposed a modified generalized log-gamma regression model to allow the possibility that long-term survivors may be presented in the data, Hashimoto et al. (2010) developed the log-exponentiated Weibull regression model for interval-censored data and Silva et al. (2010) discussed a regression model considering the Weibull extended distribution.

For the first time, we define a location-scale regression model for censored observations, based on the β-Birnbaum–Saunders (βBS for short) introduced by Cordeiro and Lemonte (2011), referred to as the log-βBS (LβBS) regression model. The proposed regression model is much more flexible than the log-BS regression model proposed by Rieck and Nedelman (1991). Further, some useful properties of the proposed model to study asymptotic inference are investigated. For some recent references about the log-BS linear regression model the reader is refereed to Lemonte et al. (2010), Lemonte and Ferrari (2011a,b,c), Lemonte (2011) and references therein. A log-BS nonlinear regression model was proposed by Lemonte and Cordeiro (2009); see also Lemonte and Cordeiro (2010) and Lemonte and Patriota (2011).

Another issue tackled is when in a sample of censored survival times, the presence of an immune proportion of individuals who are not subject to death, failure or relapse may be indicated by a relatively high number of individuals with large censored survival times. In this note, the log-βBS model is modified for the possible presence of long-term survivors in the data. The models attempt to estimate the effects of covariates on the acceleration/deceleration of the timing of a given event and the surviving fraction, that is, the proportion of the population for which the event never occurs. The logistic function is used to define the regression model for the surviving fraction.

The article is organized as follows. In Section 2, we define the LβBS distribution. In Section 3, we provide expansions for its moment generating function (mgf) and moments. In Section 4, we propose a LβBS regression model and estimate the model parameters by maximum likelihood. We derive the observed information matrix. Local influence is discussed in Section 5. In Section 6, we propose a LβBS mixture model for survival data with long-term survivors. In Section 7, we show the flexibility, practical relevance and applicability of our regression model by means of two real data sets. Section 8 ends with some concluding remarks.

2. The LβBS distribution

The BS distribution is a very popular model that has been extensively used over the past decades for modeling failure times of fatiguing materials and lifetime data in reliability, engineering and biological studies. Birnbaum and Saunders (1969a,b) define a random variable T having a BS distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, $T \sim \text{BS}(\alpha, \beta)$ say, by $T = \beta[\alpha Z/2 + \{(\alpha Z/2)^2 + 1\}^{1/2}]^2$, where Z is a standard normal random variable. Its cumulative distribution function (cdf) is defined by $G(t) = \Phi(v)$, for $t > 0$, where $v = \alpha^{-1}\rho(t/\beta)$, $\rho(z) = z^{1/2} - z^{-1/2}$ and $\Phi(\cdot)$ is the standard normal distribution function. Since $G(\beta) = \Phi(0) = 1/2$, the parameter β is the median of the distribution. For any $k > 0$, $kT \sim \text{BS}(\alpha, k\beta)$. The probability density function (pdf) of T is then $g(t) = \kappa(\alpha, \beta)t^{-3/2}(t+\beta)\ \exp\{-\tau(t/\beta)/(2\alpha^2)\}$, for $t > 0$, where $\kappa(\alpha, \beta) = \exp(\alpha^{-2})/(2\alpha\sqrt{2\pi\beta})$ and $\tau(z) = z + z^{-1}$. The fractional moments of T are $E(T^p) = \beta^p I(p, \alpha)$, where

$$I(p, \alpha) = \frac{K_{p+1/2}(\alpha^{-2}) + K_{p-1/2}(\alpha^{-2})}{2K_{1/2}(\alpha^{-2})} \tag{1}$$

and the function $K_\nu(z)$ denotes the modified Bessel function of the third kind with ν representing its order and z the argument (see Watson, 1995). Kundu et al. (2008) studied the shape of its hazard function. Results on improved statistical inference for this model are discussed by Wu and Wong (2004) and Lemonte et al. (2007, 2008). Díaz-García and Leiva (2005) proposed a new family of generalized BS distributions based on contoured elliptical distributions, whereas Guiraud et al. (2009) introduced a non-central version of the BS distribution.

The βBS distribution (Cordeiro and Lemonte, 2011), with four parameters $\alpha > 0$, $\beta > 0$, $a > 0$ and $b > 0$, extends the BS distribution and provides more flexibility to fit various types of lifetime data. Its cdf is given by $F(t) = I_{\Phi(v)}(a, b)$, where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$ is the beta function, $\Gamma(\cdot)$ is the gamma function, $I_y(a, b) = B_y(a, b)/B(a, b)$ is the incomplete beta function ratio and $B_y(a, b) = \int_0^y \omega^{a-1}\,(1-\omega)^{b-1}d\omega$ is the incomplete beta function. The density function of T has the form (for $t > 0$)

$$f_T(t) = \frac{\kappa(\alpha, \beta)}{B(a, b)}t^{-3/2}(t + \beta)\ \exp\left\{-\tau(t/\beta)/(2\alpha^2)\right\} \Phi(v)^{a-1}\{1 - \Phi(v)\}^{b-1}. \tag{2}$$

The βBS distribution contains, as special sub-models, the exponentiated BS (EBS), Lehmann type-II BS (LeBS) and BS distributions when $b = 1$, $a = 1$ and $a = b = 1$, respectively. If T is a random variable with density function (2), we write $T \sim \beta\mathrm{BS}(\alpha, \beta, a, b)$. The βBS distribution is easily simulated as follows: if V has a beta distribution with parameters a and b, then $\beta\{\alpha\Phi^{-1}(V)/2+[1+\alpha^2\Phi^{(-1)}(V)^2/4]^{1/2}\}^2$ has the $\beta\mathrm{BS}(\alpha, \beta, a, b)$ distribution. For some structural properties of this distribution, the reader is referred to Cordeiro and Lemonte (2011).

Let T be a random variable having the βBS density function (2). The random variable $W = \log(T)$ has a LβBS distribution. After some algebra, the density function of W, parameterized in terms of $\mu = \log(\beta)$, can be expressed as $f_W(w) = \xi_{01}\exp(-\xi_{02}^2/2)\Phi(\xi_{02})^{a-1}[1-\Phi(\xi_{02})]^{b-1}/\{2\sqrt{2\pi}B(a,b)\}$, $w \in \mathbb{R}$, where $\xi_{01} = 2\alpha^{-1}\cosh((w-\mu)/2)$ and $\xi_{02} = 2\alpha^{-1}\sinh((w-\mu)/2)$. The parameter $\mu \in \mathbb{R}$ is a location parameter and a, b and α are positive shape parameters. The standardized random variable $Z = (W-\mu)/2$ has density function $\pi_Z(z) = 2f_W(2z+\mu)$ given by

$$\pi_Z(z) = \frac{2\cosh(z)}{\sqrt{2\pi}B(a,b)\alpha}\exp\left\{-\frac{2}{\alpha^2}\sinh^2(z)\right\}\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{a-1}\left[1-\Phi\left(\frac{2}{\alpha}\sinh(z)\right)\right]^{b-1}, \tag{3}$$

where $-\infty < z < \infty$. Let $Y = \mu + \sigma Z$, whose density function takes the form

$$f_Y(y) = \frac{\xi_1\exp(-\xi_2^2/2)\Phi(\xi_2)^{a-1}[1-\Phi(\xi_2)]^{b-1}}{\sqrt{2\pi}\sigma B(a,b)}, \quad y \in \mathbb{R}, \tag{4}$$

where

$$\xi_1 = \frac{2}{\alpha}\cosh\left(\frac{y-\mu}{\sigma}\right) \quad \text{and} \quad \xi_2 = \frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right).$$

Here, $\sigma > 0$ acts as a scale parameter. The cdf, survival function and hazard rate function corresponding to (4) (for $y \in \mathbb{R}$) are $F(y) = I_{\Phi(\xi_2)}(a, b)$,

$$S(y) = 1 - I_{\Phi(\xi_2)}(a, b) \tag{5}$$

and $r(y) = \xi_1\exp(-\xi_2^2/2)\Phi(\xi_2)^{a-1}[1-\Phi(\xi_2)]^{b-1}/\{\sqrt{2\pi}\sigma B(a,b)[1-I_{\Phi(\xi_2)}(a,b)]\}$, respectively. If Y is a random variable having density function (4), we write $Y \sim \mathrm{L}\beta\mathrm{BS}(a, b, \alpha, \mu, \sigma)$. Thus, if $T \sim \beta\mathrm{BS}(a, b, \alpha, \beta)$, then $Y = \mu+\sigma[\log(T)-\mu]/2 \sim \mathrm{L}\beta\mathrm{BS}(a, b, \alpha, \mu, \sigma)$. The special case $b = 1$ corresponds to the log-EBS (LEBS) distribution, whereas $a = 1$ gives the log-LeBS (LLeBS) distribution. The basic exemplar is the log-BS (LBS) distribution (Rieck and Nedelman, 1991) when $\sigma = 2$ and $a = b = 1$.

Plots of the density function (4) for selected parameter values are given in Fig. 1. These plots show great flexibility of the new distribution for different values of the shape parameters a, b and α. So, the density function (4) allows for great flexibility and hence it can be very useful in many more practical situations. In fact, it can be symmetric, asymmetric and it can also exhibit bi-modality. The new model is easily simulated as follows: if V is a beta random variable with parameters a and b, then $Y = \mu + \sigma\,\mathrm{arcsinh}(\alpha\Phi^{-1}(V)/2)$ has the $\mathrm{L}\beta\mathrm{BS}(a, b, \alpha, \mu, \sigma)$ distribution. This expression can be rewritten in the form $Y = \mu + \sigma\log\left(\alpha\Phi^{-1}(V)/2 + [1+\alpha^2\Phi^{-1}(V)^2/4]^{1/2}\right)$. This scheme is useful because of the existence of fast generators for beta random variables and the standard normal quantile function is available in most statistical packages.

3. Generating function and moments

We shall obtain the mgf of the standardized LβBS random variable $Z = (W-\mu)/2$, $M_Z(s)$ say, having density function (3). We have the theorem.

Theorem. *If $Z \sim \mathrm{L}\beta\mathrm{BS}(a, b, \alpha)$, then the mgf of Z is given by*

$$M_Z(s) = \sum_{i,r=0}^{\infty} p_{i,r}N_r(s,\alpha),$$

whose coefficients are

$$p_{i,r} = p_{i,r}(a,b,\alpha) = \frac{(-1)^i 2\binom{b-1}{i}s_r(i+a-1)}{\sqrt{2\pi}\alpha B(a,b)}$$

and

$$N_r(s,\alpha) = \exp(\alpha^{-2})\sum_{m=0}^{\infty}\frac{e_{m,r}}{2^{m+2}}\sum_{j=0}^{2m+1}(-1)^j\binom{2m+1}{j}\left[K_{-(m+1-j+s/2)}(1/\alpha^2) + K_{-(m-j+s/2)}(1/\alpha^2)\right].$$

Proof. See Appendix A. □

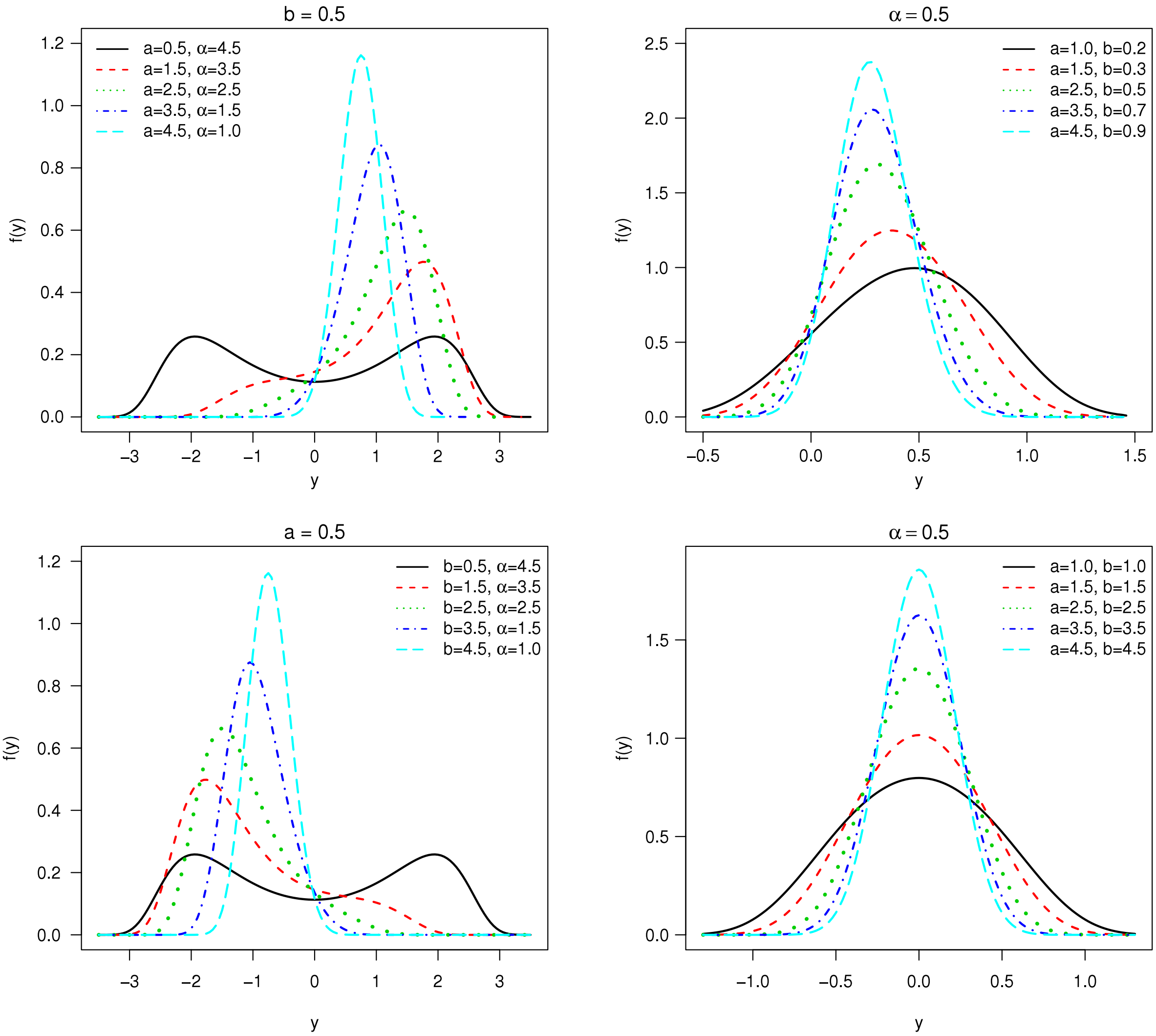

Fig. 1. Plots of the density function (4) for some parameter values: $\mu = 0$ and $\sigma = 1$.

We now derive an expansion for the rth moment of Y. First, the ith moments of T for a real is (Cordeiro and Lemonte, 2011)

$$\mu_i' = E(T^i) = \frac{1}{B(a,b)} \sum_{r=0}^{\infty} (r+1) t_{r+1} \tau_{i,r}. \tag{6}$$

Here, $t_r = \sum_{m=0}^{\infty} (-1)^m \binom{b-1}{m} (a+m)^{-1} s_r(a+m)$ and

$$\tau_{i,r} = \frac{\beta^i}{2^r} \sum_{j=0}^{r} \binom{r}{j} \sum_{k_1,\ldots,k_j=0}^{\infty} A(k_1,\ldots,k_j) \sum_{m=0}^{2s_j+j} (-1)^m \binom{2s_j+j}{m} I\big(i + (2s_j + j - 2m)/2, \alpha\big),$$

where $s_j = k_1 + \cdots + k_j$, $A(k_1,\ldots,k_j) = \alpha^{-2s_j-j} a_{k_1} \cdots a_{k_j}$, $a_k = (-1)^k 2^{(1-2k)/2} \{\sqrt{\pi}(2k+1)k!\}^{-1}$ and $I(i+(2s_j+j-2m)/2, \alpha)$ can be computed from (1) in terms of the modified Bessel function of the third kind.

From a Taylor series expansion of $H(T) = [\log(T)]^r$ around μ_1', we can write

$$E(W^r) = [\log(\mu_1')]^r + \sum_{i=2}^{\infty} \frac{H^{(i)}(\mu_1')\mu_i}{i!},$$

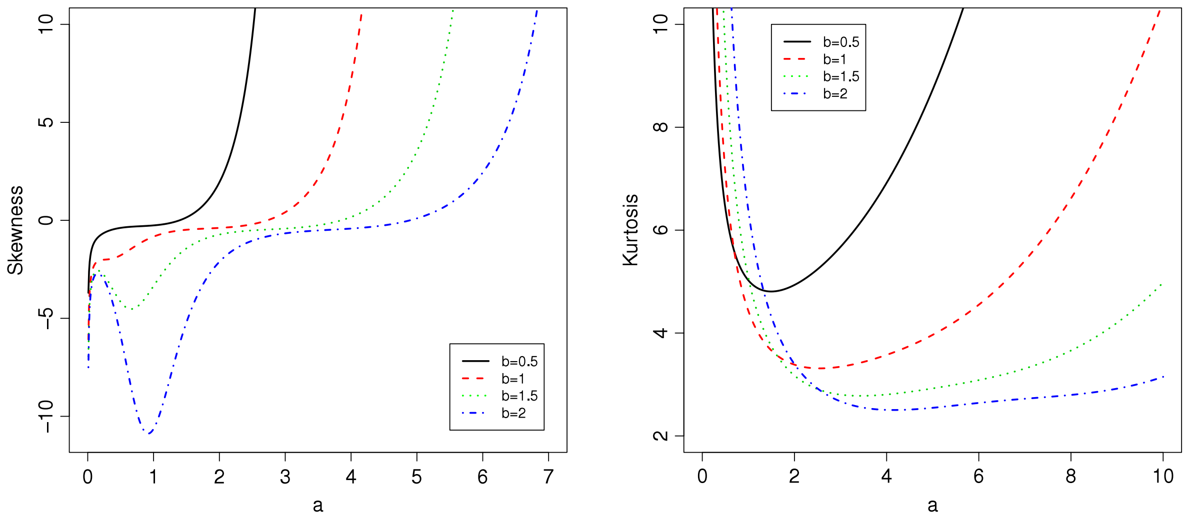

Fig. 2. Skewness and kurtosis of Y in (4) as a function of a for some values of b.

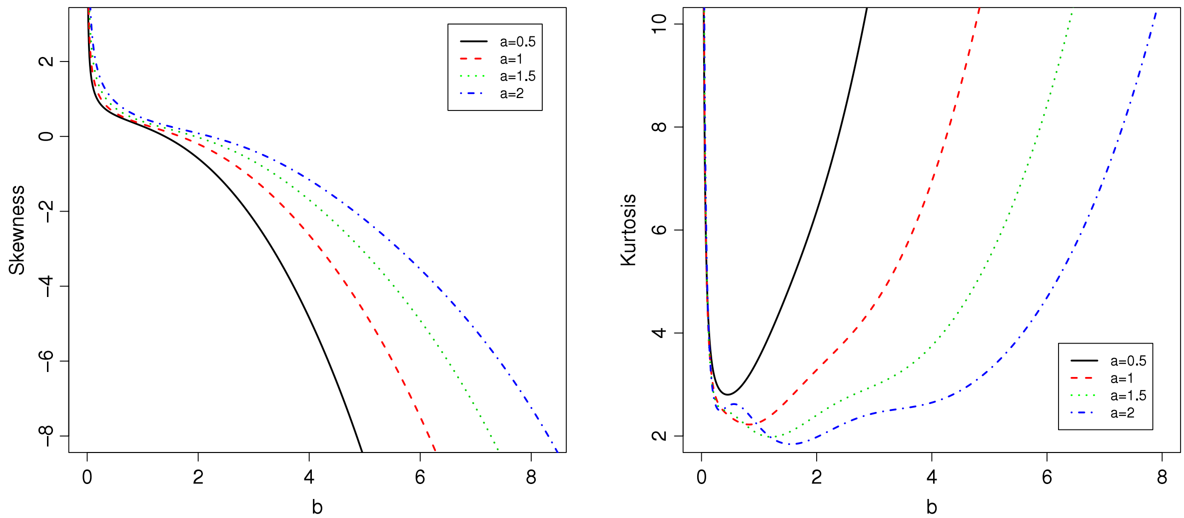

Fig. 3. Skewness and kurtosis of Y in (4) as a function of b for some values of a.

where $H(\mu_1') = [\log(\mu_1')]^r$, $H^{(i)}(\mu_1') = \partial^i H(\mu_1')/\partial \mu_1'^i$ and $\mu_i = \sum_{k=0}^{i}(-1)^k \binom{i}{k} \mu_i'^k \mu_{i-k}'^k$ is the ith central moment of T determined from (6). The ordinary moments of Y are easily obtained from the moments of W by $E(Y^r) = \sum_{i,j=0}^{r}(-1)^{r-i}2^{-j}\mu^{2r-i-j}\sigma^j \binom{r}{i}\binom{r}{j} E(W^i)$.

The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships. Plots of the skewness and kurtosis of Y for selected values of b as function of a, and for selected values of a as function of b, holding $\mu = -1.1$, $\sigma = 1.5$ and $\alpha = 5$ fixed, are shown in Figs. 2 and 3, respectively. These plots immediately reveal that the skewness and kurtosis curves, respectively, as functions of a (b fixed) first decrease and then increase, whereas as functions of b (a fixed), the skewness curve decreases and the kurtosis curve first decreases and then increases, holding the other parameters fixed. Note that both skewness and kurtosis can be quite pronounced.

4. The LβBS regression model

In many practical applications, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure and many others. Parametric models to estimate univariate survival functions and for censored data regression problems are widely used. A parametric model that provides a good fit to lifetime data tends to yield more precise estimates for the quantities of interest. Based on the LβBS distribution, we propose a linear location-scale regression model or log-linear regression model in the form

$$y_i = \mu_i + \sigma z_i, \quad i = 1, \ldots, n, \tag{7}$$

where y_i follows the density function (4), $\mu_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}$ is the location of y_i, $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^\top$ is a vector of known explanatory variables associated with y_i and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a p-vector ($p < n$) of unknown regression parameters. The location parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ of the LβBS model has a linear structure $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$, where $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^\top$ is a known model matrix of full rank, i.e. rank$(\boldsymbol{X}) = p$. The regression model (7) opens new possibilities for fitting many different types of data. It is referred to as the LβBS regression model for censored data, which is an extension of an accelerated failure time model based on the BS distribution for censored data. If $\sigma = 2$ and $a = 1$ in addition to $b = 1$, it coincides with the log-BS regression model for censored data (Leiva et al., 2007).

Let $(y_1, \boldsymbol{x}_1), \ldots, (y_n, \boldsymbol{x}_n)$ be a sample of n independent observations, where the response variable y_i corresponds to the observed log-lifetime or log-censoring time for the ith individual. We consider non-informative censoring and that the observed lifetimes and censoring times are independent. Let D and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (a, b, \alpha, \sigma, \boldsymbol{\beta}^\top)^\top$ from model (7) takes the form $\ell(\boldsymbol{\theta}) = \sum_{i\in D} \ell_i(\boldsymbol{\theta}) + \sum_{i\in C} \ell_i^{(c)}(\boldsymbol{\theta})$, where $\ell_i(\boldsymbol{\theta}) = \log[f(y_i)]$, $\ell_i^{(c)}(\boldsymbol{\theta}) = \log[S(y_i)]$, $f(y_i)$ is the density function (4) and $S(y_i)$ is the survival function (5). The total log-likelihood function for the model parameters $\boldsymbol{\theta} = (a, b, \alpha, \sigma, \boldsymbol{\beta}^\top)^\top$ can be expressed as

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= q \log\left[\frac{(2\pi)^{-1/2}}{B(a,b)\sigma}\right] + \sum_{i\in D} \log(\xi_{i1}) - \frac{1}{2}\sum_{i\in D} \xi_{i2}^2 + (a-1)\sum_{i\in D} \log\big[\Phi(\xi_{i2})\big] \\ &\quad + (b-1)\sum_{i\in D} \log\big[1 - \Phi(\xi_{i2})\big] + \sum_{i\in C} \log\big[1 - I_{\Phi(\xi_{i2})}(a,b)\big],\end{aligned}$$

where q is the observed number of failures and

$$\xi_{i1} = \xi_{i1}(\boldsymbol{\theta}) = \frac{2}{\alpha}\cosh\left(\frac{y_i - \mu_i}{\sigma}\right), \qquad \xi_{i2} = \xi_{i2}(\boldsymbol{\theta}) = \frac{2}{\alpha}\sinh\left(\frac{y_i - \mu_i}{\sigma}\right),$$

for $i = 1, \ldots, n$. The score functions for the parameters a, b, α, σ and $\boldsymbol{\beta}$ are given by

$$U_a(\boldsymbol{\theta}) = q[\psi(a+b) - \psi(a)] + \sum_{i\in D} \log[\Phi(\xi_{i2})] - \sum_{i\in C} \frac{[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_a}{1 - I_{\Phi(\xi_{i2})}(a,b)},$$

$$U_b(\boldsymbol{\theta}) = q[\psi(a+b) - \psi(b)] + \sum_{i\in D} \log[1 - \Phi(\xi_{i2})] - \sum_{i\in C} \frac{[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_b}{1 - I_{\Phi(\xi_{i2})}(a,b)},$$

$$U_\alpha(\boldsymbol{\theta}) = -\frac{q}{\alpha} + \frac{1}{\alpha}\sum_{i\in D} \xi_{i2}^2 - \frac{(a-1)}{\alpha}\sum_{i\in D} \frac{\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \frac{(b-1)}{\alpha}\sum_{i\in D} \frac{\xi_{i2}\phi(\xi_{i2})}{1 - \Phi(\xi_{i2})} - \sum_{i\in C} \frac{[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_\alpha}{1 - I_{\Phi(\xi_{i2})}(a,b)},$$

$$\begin{aligned}U_\sigma(\boldsymbol{\theta}) &= -\frac{q}{\sigma} - \frac{1}{\sigma}\sum_{i\in D} \frac{z_i\xi_{i2}}{\xi_{i1}} + \frac{1}{\sigma}\sum_{i\in D} z_i\xi_{i1}\xi_{i2} - \frac{(a-1)}{\sigma}\sum_{i\in D} \frac{z_i\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} \\ &\quad + \frac{(b-1)}{\sigma}\sum_{i\in D} \frac{z_i\xi_{i1}\phi(\xi_{i2})}{1 - \Phi(\xi_{i2})} - \sum_{i\in C} \frac{[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_\sigma}{1 - I_{\Phi(\xi_{i2})}(a,b)}\end{aligned}$$

and $\boldsymbol{U}_\beta(\boldsymbol{\theta}) = \boldsymbol{X}^\top \boldsymbol{s}$, respectively, where $\boldsymbol{s} = (s_1, \ldots, s_n)^\top$ and

$$s_i = \begin{cases} -\dfrac{\xi_{i2}}{\sigma\xi_{i1}} + \dfrac{\xi_{i1}\xi_{i2}}{\sigma} - \dfrac{(a-1)}{\sigma}\dfrac{\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \dfrac{(b-1)}{\sigma}\dfrac{\xi_{i1}\phi(\xi_{i2})}{[1 - \Phi(\xi_{i2})]}, & i \in D, \\ \dfrac{\xi_{i1}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1 - \Phi(\xi_{i2})]^{b-1}}{\sigma B(a,b)[1 - I_{\Phi(\xi_{i2})}(a,b)]}, & i \in C, \end{cases}$$

$\psi(\cdot)$ is the digamma function, $z_i = (y_i - \mu_i)/\sigma$,

$$[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_a = \bar{I}^{(0)}_{\Phi(\xi_{i2})}(a,b) - [\psi(a) - \psi(a+b)]I_{\Phi(\xi_{i2})}(a,b),$$

$$[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_b = \bar{I}^{(1)}_{\Phi(\xi_{i2})}(a,b) - [\psi(b) - \psi(a+b)]I_{\Phi(\xi_{i2})}(a,b),$$

$$[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_\alpha = -\frac{\xi_{i2}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1 - \Phi(\xi_{i2})]^{b-1}}{\alpha B(a,b)},$$

$$[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_\sigma = -\frac{z_i\xi_{i1}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1 - \Phi(\xi_{i2})]^{b-1}}{\sigma B(a,b)},$$

and

$$\bar{I}^{(k)}_{\Phi(\xi_{i2})}(a,b) = \frac{1}{B(a,b)} \int_0^{\Phi(\xi_{i2})} [\log(w)]^{1-k} [\log(1-w)]^k w^{a-1}(1-w)^{b-1} dw, \quad k = 0, 1.$$

The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}} = (\widehat{a}, \widehat{b}, \widehat{\alpha}, \widehat{\sigma}, \widehat{\boldsymbol{\beta}}^\top)^\top$ of $\boldsymbol{\theta} = (a, b, \alpha, \sigma, \boldsymbol{\beta}^\top)^\top$ can be obtained by solving simultaneously the nonlinear equations $U_a(\boldsymbol{\theta}) = 0$, $U_b(\boldsymbol{\theta}) = 0$, $U_\alpha(\boldsymbol{\theta}) = 0$, $U_\sigma(\boldsymbol{\theta}) = 0$ and $\boldsymbol{U}_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = \mathbf{0}$. These equations cannot be solved analytically and require iterative techniques such as the Newton–Raphson algorithm. After fitting the model (7), the survival function for y_i can be readily estimated by $\widehat{S}(y_i) = 1 - I_{\Phi(\widehat{\xi}_{2i})}(\widehat{a}, \widehat{b})$, where $\widehat{\xi}_{2i} = \xi_{i2}(\widehat{\boldsymbol{\theta}})$, for $i = 1, \ldots, n$.

The normal approximation for the MLE of $\boldsymbol{\theta}$ can be used for constructing approximate confidence intervals and for testing hypotheses on the parameters a, b, α, σ and $\boldsymbol{\beta}$. Under conditions that are fulfilled for the parameters in the interior of the parameter space, we obtain $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{A}{\sim} \mathcal{N}_{p+4}(\mathbf{0}, \boldsymbol{K}_{\boldsymbol{\theta}}^{-1})$, where $\overset{A}{\sim}$ means approximately distributed and $\boldsymbol{K}_{\boldsymbol{\theta}}$ is the unit expected information matrix. The asymptotic result $\boldsymbol{K}_{\boldsymbol{\theta}} = \lim_{n\to\infty} n^{-1}[-\ddot{\boldsymbol{L}}(\boldsymbol{\theta})]$ holds, where $-\ddot{\boldsymbol{L}}(\boldsymbol{\theta})$ is the $(p+4) \times (p+4)$ observed information matrix. The average matrix evaluated at $\widehat{\boldsymbol{\theta}}$, say $-n^{-1}\ddot{\boldsymbol{L}}(\widehat{\boldsymbol{\theta}})$, can estimate $\boldsymbol{K}_{\boldsymbol{\theta}}$. The elements of the matrix $\ddot{\boldsymbol{L}}(\boldsymbol{\theta}) = \partial^2 \ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$ are given in the Appendix B.

The likelihood ratio (LR) statistic can be used to discriminate between the LβBS and LEBS regression models, since they are nested models, by testing the null hypothesis $\mathcal{H}_0 : b = 1$ against the alternative hypothesis $\mathcal{H}_1 : b \neq 1$. In this case, the LR statistic is equal to $w = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\}$, where $\widetilde{\boldsymbol{\theta}} = (\widetilde{a}, 1, \widetilde{\alpha}, \widetilde{\sigma}, \widetilde{\boldsymbol{\beta}}^\top)^\top$ is the MLE of $\boldsymbol{\theta} = (a, b, \alpha, \sigma, \boldsymbol{\beta}^\top)^\top$ under $\mathcal{H}_0$. The null hypothesis is rejected if $w > \chi^2_{1-\eta}(1)$, where $\chi^2_{1-\eta}(1)$ is the quantile of the chi-square distribution with one degree of freedom and η is the significance level.

5. Local influence

Since regression models are sensitive to the underlying model assumptions, generally performing a sensitivity analysis is strongly advisable. Cook (1986) used this idea to motivate his assessment of influence analysis. He suggested that more confidence can be put in a model which is relatively stable under small modifications. The first technique developed to assess the individual impact of cases on the estimation process is based on case-deletion (see, for example, Cook and Weisberg, 1982) in which the effects are studied after removing some observations from the analysis. This is a global influence analysis, since the effect of the case is evaluated by dropping it from the data. The local influence method is recommended when the concern is related to investigate the model sensibility under some minor perturbations in the model. In survival analysis, several authors have investigated the assessment of local influence as, for instance, Pettit and Bin Daud (1989), Escobar and Meeker (1992) and Ortega et al. (2003), among others. Considering the likelihood function for assessing the curvature for influence analysis, other techniques have been proposed to deal with non-standard situations and for various models; see, for example, Fung and Kwan (1997), Kwan and Fung (1998) and Tanaka et al. (2003), among others.

The local influence method is recommended when the concern is related to investigate the model sensitivity under some minor perturbations in the model (or data). Let $\boldsymbol{\omega}$ be a k-dimensional vector of perturbations restricted to some open subset $\boldsymbol{\Omega}$ of $\mathbb{R}^k$. The perturbed log-likelihood function is denoted by $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$. We consider that exists a no perturbation vector $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$ such that $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$, for all $\boldsymbol{\theta}$. The influence of minor perturbations on the MLE $\widehat{\boldsymbol{\theta}}$ can be assessed by using the likelihood displacement $LD_{\boldsymbol{\omega}} = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}$, where $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denotes the maximizer of $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$.

The idea for assessing local influence as advocated by Cook (1986) is essentially the analysis of the local behavior of $LD_{\boldsymbol{\omega}}$ around $\boldsymbol{\omega}_0$ by evaluating the curvature of the plot of $LD_{\boldsymbol{\omega}_0 + a\boldsymbol{d}}$ against a, where $a \in \mathbb{R}$ and $\boldsymbol{d}$ is a unit direction. One of the measures of particular interest is the direction $\boldsymbol{d}_{\max}$ corresponding to the largest curvature $C_{\boldsymbol{d}_{\max}}$. The index plot of $\boldsymbol{d}_{\max}$ may evidence those observations that have considerable influence on $LD_{\boldsymbol{\omega}}$ under minor perturbations. Also, plots of $\boldsymbol{d}_{\max}$ against covariate values may be helpful for identifying atypical patterns. Cook (1986) showed that the normal curvature at the direction $\boldsymbol{d}$ is given by $C_{\boldsymbol{d}}(\boldsymbol{\theta}) = 2|\boldsymbol{d}^\top \boldsymbol{\Delta}^\top \ddot{\boldsymbol{L}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}\boldsymbol{d}|$, where $\boldsymbol{\Delta} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top$, both $\boldsymbol{\Delta}$ and $\ddot{\boldsymbol{L}}(\boldsymbol{\theta})$ are evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. Moreover, $C_{\boldsymbol{d}_{\max}}$ is twice the largest eigenvalue of $\boldsymbol{B} = -\boldsymbol{\Delta}^\top \ddot{\boldsymbol{L}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}$ and $\boldsymbol{d}_{\max}$ is the corresponding eigenvector. The index plot of $\boldsymbol{d}_{\max}$ may reveal how to perturb the model (or data) to obtain large changes in the estimate of $\boldsymbol{\theta}$.

Assume that the parameter vector $\boldsymbol{\theta}$ is partitioned as $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$. The dimensions of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are p_1 and $p - p_1$, respectively. Let

$$\ddot{\boldsymbol{L}}(\boldsymbol{\theta}) = \begin{bmatrix} \ddot{\boldsymbol{L}}_{\theta_1\theta_1} & \ddot{\boldsymbol{L}}_{\theta_1\theta_2} \\ \ddot{\boldsymbol{L}}^\top_{\theta_1\theta_2} & \ddot{\boldsymbol{L}}_{\theta_2\theta_2} \end{bmatrix},$$

where $\ddot{\boldsymbol{L}}_{\theta_1\theta_1} = \partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}_1\partial\boldsymbol{\theta}_1^\top$, $\ddot{\boldsymbol{L}}_{\theta_1\theta_2} = \partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}_1\partial\boldsymbol{\theta}_2^\top$ and $\ddot{\boldsymbol{L}}_{\theta_2\theta_2} = \partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}_2\partial\boldsymbol{\theta}_2^\top$. If the interest lies on $\boldsymbol{\theta}_1$, the normal curvature in the direction of the vector $\boldsymbol{d}$ is $C_{\boldsymbol{d};\boldsymbol{\theta}_1}(\boldsymbol{\theta}) = 2|\boldsymbol{d}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{L}}(\boldsymbol{\theta})^{-1} - \ddot{\boldsymbol{L}}_{22})\boldsymbol{\Delta}\boldsymbol{d}|$, where

$$\ddot{\boldsymbol{L}}_{22} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\boldsymbol{L}}^{-1}_{\theta_2\theta_2} \end{bmatrix}$$

and $\boldsymbol{d}_{\max;\boldsymbol{\theta}_1}$ here is the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{B}_1 = -\boldsymbol{\Delta}^\top(\ddot{\boldsymbol{L}}(\boldsymbol{\theta})^{-1} - \ddot{\boldsymbol{L}}_{22})\boldsymbol{\Delta}$. The index plot of the $\boldsymbol{d}_{\max;\boldsymbol{\theta}_1}$ may reveal those influential elements on $\widehat{\boldsymbol{\theta}}_1$.

In what follows, we derive for three perturbation schemes, the matrix

$$\boldsymbol{\Delta} = \left.\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \begin{bmatrix}\boldsymbol{\Delta}_a^\top & \boldsymbol{\Delta}_b^\top & \boldsymbol{\Delta}_\alpha^\top & \boldsymbol{\Delta}_\sigma^\top & \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top\end{bmatrix}^\top.$$

The quantities evaluated at $\widehat{\boldsymbol{\theta}}$ are written with a circumflex.

Case weight perturbation

A perturbed log-likelihood function, allowing different weights for different observations, can be defined in the form $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i\in D}\omega_i\ell_i(\boldsymbol{\theta}) + \sum_{i\in C}\omega_i\ell_i^{(c)}(\boldsymbol{\theta})$, where $\boldsymbol{\omega} = (\omega_1,\ldots,\omega_n)^\top$ is a n-dimensional vector of weights from the contributions of the components of the log-likelihood function. Also, let $\boldsymbol{\omega}_0 = (1,\ldots,1)^\top$ be the vector of no perturbation such that $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$. After some algebra, we have

$$\boldsymbol{\Delta}_a = (\widehat{k}_{11},\ldots,\widehat{k}_{1n}),\qquad \boldsymbol{\Delta}_b = (\widehat{k}_{21},\ldots,\widehat{k}_{2n}),\qquad \boldsymbol{\Delta}_\alpha = (\widehat{k}_{31},\ldots,\widehat{k}_{3n}),$$
$$\boldsymbol{\Delta}_\sigma = (\widehat{k}_{41},\ldots,\widehat{k}_{4n}),\qquad \boldsymbol{\Delta}_{\boldsymbol{\beta}} = \boldsymbol{X}^\top\widehat{\boldsymbol{S}},$$

where $\boldsymbol{S} = \mathrm{diag}\{s_1,\ldots,s_n\}$,

$$k_{1i} = \begin{cases}\psi(a+b)-\psi(a)+\log[\Phi(\xi_{i2})], & i\in D,\\ \dfrac{-\bar{I}^{(0)}_{\Phi(\xi_{i2})}(a,b)+[\psi(a)-\psi(a+b)]I_{\Phi(\xi_{i2})}(a,b)}{1-I_{\Phi(\xi_{i2})}(a,b)}, & i\in C,\end{cases}$$

$$k_{2i} = \begin{cases}\psi(a+b)-\psi(b)+\log[1-\Phi(\xi_{i2})], & i\in D,\\ \dfrac{-\bar{I}^{(1)}_{\Phi(\xi_{i2})}(a,b)+[\psi(b)-\psi(a+b)]I_{\Phi(\xi_{i2})}(a,b)}{1-I_{\Phi(\xi_{i2})}(a,b)}, & i\in C,\end{cases}$$

$$k_{3i} = \begin{cases}-\dfrac{1}{\alpha}+\dfrac{\xi_{i2}^2}{\alpha}-\dfrac{(a-1)}{\alpha}\dfrac{\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})}+\dfrac{(b-1)}{\alpha}\dfrac{\xi_{i2}\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}, & i\in D,\\ \dfrac{\xi_{i2}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1-\Phi(\xi_{i2})]^{b-1}}{\alpha B(a,b)[1-I_{\Phi(\xi_{i2})}(a,b)]}, & i\in C,\end{cases}$$

$$k_{4i} = \begin{cases}-\dfrac{1}{\sigma}-\dfrac{z_i\xi_{i2}}{\sigma\xi_{i1}}+\dfrac{z_i\xi_{i1}\xi_{i2}}{\sigma}-\dfrac{(a-1)}{\sigma}\dfrac{z_i\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})}+\dfrac{(b-1)}{\sigma}\dfrac{z_i\xi_{i1}\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}, & i\in D,\\ \dfrac{z_i\xi_{i1}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1-\Phi(\xi_{i2})]^{b-1}}{\sigma B(a,b)[1-I_{\Phi(\xi_{i2})}(a,b)]}, & i\in C.\end{cases}$$

Response perturbation

We shall consider here that each y_i is perturbed as $y_{i\omega} = y_i + \omega_i s_y$, where s_y is a scale factor that may be estimated by the standard deviation of $\boldsymbol{y}$. Let $\xi_{i1\omega_1} = \xi_{i1\omega_1}(\boldsymbol{\theta}) = 2\alpha^{-1}\cosh([y_{i\omega}-\mu_i]/\sigma)$, $\xi_{i2\omega_1} = \xi_{i2\omega_1}(\boldsymbol{\theta}) = 2\alpha^{-1}\sinh([y_{i\omega}-\mu_i]/\sigma)$ and $z_{i\omega_1} = (y_{iw}-\mu_i)/\sigma$. Also, let $\boldsymbol{\omega}_0 = (0,\ldots,0)^\top$ be the vector of no perturbations. In this case, we have

$$\boldsymbol{\Delta}_a = (\widehat{m}_{11},\ldots,\widehat{m}_{1n}),\qquad \boldsymbol{\Delta}_b = (\widehat{m}_{21},\ldots,\widehat{m}_{2n}),\qquad \boldsymbol{\Delta}_\alpha = (\widehat{m}_{31},\ldots,\widehat{m}_{3n}),$$
$$\boldsymbol{\Delta}_\sigma = (\widehat{m}_{41},\ldots,\widehat{m}_{4n}),\qquad \boldsymbol{\Delta}_{\boldsymbol{\beta}} = \boldsymbol{X}^\top\widehat{\boldsymbol{N}},$$

where $\boldsymbol{N} = \mathrm{diag}\{N_1,\ldots,N_n\}$,

$$m_{1i} = \begin{cases}\dfrac{s_y\xi_{i1}\phi(\xi_{i2})}{\sigma\Phi(\xi_{i2})}, & i\in D,\\ -\dfrac{\partial}{\partial\omega_i}\left[\dfrac{\bar{I}^{(0)}_{\Phi(\xi_{i2\omega_1})}(a,b)-[\psi(a)-\psi(a+b)]I_{\Phi(\xi_{i2\omega_1})}(a,b)}{1-I_{\Phi(\xi_{i2\omega_1})}}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$m_{2i} = \begin{cases}-\dfrac{s_y\xi_{i1}\phi(\xi_{i2})}{\sigma[1-\Phi(\xi_{i2})]}, & i\in D,\\ -\dfrac{\partial}{\partial\omega_i}\left[\dfrac{\bar{I}^{(1)}_{\Phi(\xi_{i2\omega_1})}(a,b)-[\psi(b)-\psi(a+b)]I_{\Phi(\xi_{i2\omega_1})}(a,b)}{1-I_{\Phi(\xi_{i2\omega_1})}}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$m_{3i}=\begin{cases}\dfrac{2s_y\xi_{i1}\xi_{i2}}{\alpha\sigma}+\dfrac{s_y(a-1)\xi_{i1}\phi(\xi_{i2})}{\alpha\sigma\Phi(\xi_{i2})}\left[\xi_{i2}^2-1+\dfrac{\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\dfrac{s_y(b-1)\xi_{i1}\phi(\xi_{i2})}{\alpha\sigma[1-\Phi(\xi_{i2})]}\left[1-\xi_{i2}^2+\dfrac{\xi_{i2}\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \dfrac{\partial}{\partial\omega_i}\left[\dfrac{\xi_{i2\omega_1}\phi(\xi_{i2\omega_1})\Phi(\xi_{i2\omega_1})^{a-1}[1-\Phi(\xi_{i2\omega_1})]^{b-1}}{\alpha B(a,b)[1-I_{\Phi(\xi_{i2\omega_1})}]}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$m_{4i}=\begin{cases}-\dfrac{s_yz_i}{\sigma^2}-\dfrac{s_y\xi_{i2}}{\sigma^2\xi_{i1}}+\dfrac{s_yz_i\xi_{i2}^2}{\sigma^2\xi_{i1}^2}+\dfrac{s_yz_i\xi_{i1}^2}{\sigma^2}+\dfrac{s_yz_i\xi_{i2}^2}{\sigma^2}+\dfrac{s_yz_i\xi_{i1}\xi_{i2}}{\sigma^2}\\ \quad+\dfrac{s_y(a-1)\phi(\xi_{i2})}{\sigma^2\Phi(\xi_{i2})}\left[z_i\xi_{i1}^2\xi_{i2}-z_i\xi_{i2}-\xi_{i1}+\dfrac{z_i\xi_{i1}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\dfrac{s_y(b-1)\phi(\xi_{i2})}{\sigma^2[1-\Phi(\xi_{i2})]}\left[z_i\xi_{i2}+\xi_{i1}-z_i\xi_{i1}^2\xi_{i2}+\dfrac{z_i\xi_{i1}^2\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \dfrac{\partial}{\partial\omega_i}\left[\dfrac{z_{i\omega_1}\xi_{i1\omega_1}\phi(\xi_{i2\omega_1})\Phi(\xi_{i2\omega_1})^{a-1}[1-\Phi(\xi_{i2\omega_1})]^{b-1}}{\sigma B(a,b)[1-I_{\Phi(\xi_{i2\omega_1})}]}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$N_i=\begin{cases}-\dfrac{s_y}{\sigma^2}+\dfrac{s_y\xi_{i2}^2}{\sigma^2\xi_{i1}^2}+\dfrac{s_y\xi_{i1}^2}{\sigma^2}+\dfrac{s_y\xi_{i2}^2}{\sigma^2}+\dfrac{s_y(a-1)\phi(\xi_{i2})}{\sigma^2\Phi(\xi_{i2})}\left[-\xi_{i2}+\xi_{i1}^2\xi_{i2}+\dfrac{\xi_{i1}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\dfrac{s_y(b-1)\phi(\xi_{i2})}{\sigma^2[1-\Phi(\xi_{i2})]}\left[\xi_{i2}-\xi_{i1}^2\xi_{i2}+\dfrac{\xi_{i1}^2\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \dfrac{\partial}{\partial\omega_i}\left[\dfrac{\xi_{i1\omega_1}\phi(\xi_{i2\omega_1})\Phi(\xi_{i2\omega_1})^{a-1}[1-\Phi(\xi_{i2\omega_1})]^{b-1}}{\sigma B(a,b)[1-I_{\Phi(\xi_{i2\omega_1})}(a,b)]}\right]_{\omega_i=0}, & i\in C.\end{cases}$$

Explanatory variable perturbation

Now, consider an additive perturbation on a particular continuous explanatory variable, say $\boldsymbol{x}_t$, by setting $x_{it\omega}=x_{it}+\omega_i s_x$, where s_x is a scale factor that may be estimated by the standard deviation of $\boldsymbol{x}_t$. Let $\xi_{i1\omega_2}=\xi_{i1\omega_2}(\boldsymbol{\theta})=2\alpha^{-1}\cosh([y_i-\mu_i-\beta_t\omega_i s_x]/\sigma)$, $\xi_{i2\omega_2}=\xi_{i2\omega_2}(\boldsymbol{\theta})=2\alpha^{-1}\sinh([y_i-\mu_i-\beta_t\omega_i s_x]/\sigma)$ and $z_{i\omega_2}=(y_i-\mu_i-\beta_t\omega_i s_x)/\sigma$. Here, $\boldsymbol{\omega}_0=(0,\ldots,0)^\top$ is the vector of no perturbations. Under this perturbation scheme, we have

$$\boldsymbol{\Delta}_a=(\widehat{e}_{11},\ldots,\widehat{e}_{1n}),\qquad \boldsymbol{\Delta}_b=(\widehat{e}_{21},\ldots,\widehat{e}_{2n}),\qquad \boldsymbol{\Delta}_\alpha=(\widehat{e}_{31},\ldots,\widehat{e}_{3n}),\quad \boldsymbol{\Delta}_\sigma=(\widehat{e}_{41},\ldots,\widehat{e}_{4n}),$$

where

$$e_{1i}=\begin{cases}-\dfrac{s_x\beta_t\xi_{i1}\phi(\xi_{i2})}{\sigma\Phi(\xi_{i2})}, & i\in D,\\ -\dfrac{\partial}{\partial\omega_i}\left[\dfrac{\bar{I}^{(0)}_{\Phi(\xi_{i2\omega_2})}(a,b)-[\psi(a)-\psi(a+b)]I_{\Phi(\xi_{i2\omega_2})}(a,b)}{1-I_{\Phi(\xi_{i2\omega_2})}}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$e_{2i}=\begin{cases}\dfrac{s_x\beta_t\xi_{i1}\phi(\xi_{i2})}{\sigma[1-\Phi(\xi_{i2})]}, & i\in D,\\ -\dfrac{\partial}{\partial\omega_i}\left[\dfrac{\bar{I}^{(1)}_{\Phi(\xi_{i2\omega_2})}(a,b)-[\psi(b)-\psi(a+b)]I_{\Phi(\xi_{i2\omega_2})}(a,b)}{1-I_{\Phi(\xi_{i2\omega_2})}}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$e_{3i}=\begin{cases}-\dfrac{2s_x\beta_t\xi_{i1}\xi_{i2}}{\alpha\sigma}+\dfrac{s_x(a-1)\beta_t\xi_{i1}\phi(\xi_{i2})}{\alpha\sigma\Phi(\xi_{i2})}\left[1-\xi_{i2}^2-\dfrac{\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\dfrac{s_x(b-1)\beta_t\xi_{i1}\phi(\xi_{i2})}{\alpha\sigma[1-\Phi(\xi_{i2})]}\left[\xi_{i2}^2-1-\dfrac{\xi_{i2}\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \dfrac{\partial}{\partial\omega_i}\left[\dfrac{\xi_{i2\omega_2}\phi(\xi_{i2\omega_2})\Phi(\xi_{i2\omega_2})^{a-1}[1-\Phi(\xi_{i2\omega_2})]^{b-1}}{\alpha B(a,b)[1-I_{\Phi(\xi_{i2\omega_2})}]}\right]_{\omega_i=0}, & i\in C,\end{cases}$$

$$e_{4i}=\begin{cases}\frac{s_x z_i\beta_t}{\sigma^2}+\frac{s_x\beta_t\xi_{i2}}{\sigma^2\xi_{i1}}-\frac{s_x z_i\beta_t\xi_{i2}^2}{\sigma^2\xi_{i1}^2}-\frac{s_x z_i\beta_t\xi_{i1}^2}{\sigma^2}-\frac{s_x z_i\beta_t\xi_{i2}^2}{\sigma^2}-\frac{s_x z_i\beta_t\xi_{i1}\xi_{i2}}{\sigma^2}\\ \quad+\frac{s_x(a-1)\beta_t\phi(\xi_{i2})}{\sigma^2\Phi(\xi_{i2})}\left[z_i\xi_{i2}+\xi_{i1}-z_i\xi_{i1}^2\xi_{i2}-\frac{z_i\xi_{i1}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\frac{s_x(b-1)\beta_t\phi(\xi_{i2})}{\sigma^2[1-\Phi(\xi_{i2})]}\left[z_i\xi_{i1}^2\xi_{i2}-z_i\xi_{i2}-\xi_{i1}-\frac{z_i\xi_{i1}^2\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \frac{\partial}{\partial\omega_i}\left[\frac{z_{i\omega_2}\xi_{i1\omega_2}\phi(\xi_{i2\omega_2})\Phi(\xi_{i2\omega_2})^{a-1}[1-\Phi(\xi_{i2\omega_2})]^{b-1}}{\sigma B(a,b)[1-I_{\Phi(\xi_{i2\omega_2})}]}\right]_{\omega_i=0}, & i\in C.\end{cases}$$

The matrix $\mathbf{\Delta}_{\beta}=\{\widehat{\delta}_{ji}\}$ of dimension $p\times n$ ($j=1,\dots,p$ and $i=1,\dots,n$) has elements when $j\neq t$ in the form

$$\delta_{ji}=\begin{cases}\frac{s_x\beta_t x_{ij}}{\sigma^2}-\frac{s_x\beta_t x_{ij}\xi_{i2}^2}{\sigma^2\xi_{i1}^2}-\frac{s_x\beta_t x_{ij}\xi_{i1}^2}{\sigma^2}-\frac{s_x\beta_t x_{ij}\xi_{i2}^2}{\sigma^2}+\frac{s_x(a-1)\beta_t x_{ij}\phi(\xi_{i2})}{\sigma^2\Phi(\xi_{i2})}\left[\xi_{i2}-\xi_{i1}^2\xi_{i2}-\frac{\xi_{i1}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\frac{s_x(b-1)\beta_t x_{ij}\phi(\xi_{i2})}{\sigma^2[1-\Phi(\xi_{i2})]}\left[\xi_{i1}^2\xi_{i2}-\xi_{i2}-\frac{\xi_{i1}^2\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \frac{\partial}{\partial\omega_i}\left[\frac{\xi_{i1\omega_1}\phi(\xi_{i2\omega_1})\Phi(\xi_{i2\omega_1})^{a-1}[1-\Phi(\xi_{i2\omega_1})]^{b-1}}{\sigma B(a,b)[1-I_{\Phi(\xi_{i2\omega_1})}(a,b)]}\right]_{\omega_i=0}, & i\in C.\end{cases}$$

For $j=t$, we have

$$\delta_{ti}=\begin{cases}\frac{s_x\beta_t x_{it}}{\sigma^2}-\frac{s_x\beta_t x_{it}\xi_{i2}^2}{\sigma^2\xi_{i1}^2}-\frac{s_x\beta_t x_{it}\xi_{i1}^2}{\sigma^2}-\frac{s_x\beta_t x_{it}\xi_{i2}^2}{\sigma^2}-\frac{s_x\xi_{i2}}{\sigma\xi_{i1}}+\frac{s_x\xi_{i1}\xi_{i2}}{\sigma}-\frac{s_x(a-1)\xi_{i1}\phi(\xi_{i2})}{\sigma\Phi(\xi_{i2})}\\ \quad+\frac{s_x(b-1)\xi_{i1}\phi(\xi_{i2})}{\sigma[1-\Phi(\xi_{i2})]}+\frac{s_x(a-1)\phi(\xi_{i2})}{\sigma^2\Phi(\xi_{i2})}\left[\xi_{i2}-\xi_{i1}^2\xi_{i2}-\frac{\xi_{i1}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})}\right]\\ \quad+\frac{s_x(b-1)\beta_t x_{it}\phi(\xi_{i2})}{\sigma^2[1-\Phi(\xi_{i2})]}\left[\xi_{i1}^2\xi_{i2}-\xi_{i2}-\frac{\xi_{i1}^2\phi(\xi_{i2})}{[1-\Phi(\xi_{i2})]}\right], & i\in D,\\ \frac{\partial}{\partial\omega_i}\left[\frac{\xi_{i1\omega_1}\phi(\xi_{i2\omega_1})\Phi(\xi_{i2\omega_1})^{a-1}[1-\Phi(\xi_{i2\omega_1})]^{b-1}}{\sigma B(a,b)[1-I_{\Phi(\xi_{i2\omega_1})}(a,b)]}\right]_{\omega_i=0}\\ \quad+\frac{s_x\xi_{i1}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1-\Phi(\xi_{i2})]^{b-1}}{\sigma B(a,b)}, & i\in C.\end{cases}$$

6. The LβBS mixture model for cure fraction

In population-based cancer studies, cure is said to occur when mortality in the group of cancer patients returns to the same level as that expected in the general population. The cure fraction is of interest to patients as well as a useful measure when analyzing trends in cancer patients survival. Models for survival analysis typically assume that every subject in the study population is susceptible to the event under study and will eventually experience such event if follow-up is sufficiently long. However, there are situations when a fraction of individuals are not expected to experience the event of interest, that is, those individuals are cured or not susceptible. Cure rate models have been used for modeling time-to-event data for various types of cancers, including breast cancer, non-Hodgkins lymphoma, leukemia, prostate cancer and melanoma. Perhaps, the most popular cure rate models are the mixture models (MMs) introduced by Boag (1949), Berkson and Gage (1952) and Farewell (1982). Additionally, MMs allow both the cure fraction and the survival function of uncured patients (latency distribution) to depend on covariates. Further, Longini and Halloran (1996) and Price and Manatunga (2001) have introduced frailty to MMs for individual survival data. Recently, Peng and Dear (2000) investigated a nonparametric mixture model for cure estimation, Sy and Taylor (2000) considered estimation in a proportional hazard cure model, Yu and Peng (2008) have extended MMs to bivariate survival data by modeling marginal distributions and Ortega et al. (2009c) proposed the generalized log-gamma mixture model with covariates. Benerjee and Carlin (2004) extended multivariate cure rate models to allow for spatial correlation as well as interval censoring and used a Bayesian approach, where posterior summaries are obtained via the hybrid Markov Chain Monte Carlo algorithm. Li et al. (2005) considered MMs in the presence of dependent censoring, from the perspective of competing risks and model the dependence between the censoring time and the survival time using a class of Archimedean copula models and Zeng et al. (2006) proposed a class of transformation models for survival data with a cure fraction. This class of transformation models was motivated by biological considerations, and it includes both the proportional hazards and proportional odds cure models as two special cases.

To formulate the LβBS mixture (LβBSM) model, we consider that the studied population is a mixture of susceptible (uncured) individuals, who may experience the event of interest, and non-susceptible (cured) individuals, who will experience it (Maller and Zhou, 1996). This approach allows to estimate simultaneously whether the event of interest will occur, which is called *incidence*, and when it will occur, given that it can occur, which is called *latency*. Let N_i $(i = 1, \ldots, n)$ be the indicator denoting that the ith individual is susceptible $(N_i = 1)$ or non-susceptible $(N_i = 0)$. The mixture model is given by

$$S_{\text{pop}}(y_i|\boldsymbol{x}_i) = \pi(\boldsymbol{x}_i) + \big[1 - \pi(\boldsymbol{x}_i)\big] S(y_i|N_i = 1), \tag{8}$$

where $S_{\text{pop}}(y_i|\boldsymbol{x}_i)$ is the unconditional survival function of y_i for the entire population, $S(y_i|N_i = 1)$ is the survival function for susceptible individuals and $\pi(\boldsymbol{x}_i) = P(N_i = 0|\boldsymbol{x}_i)$ is the probability of cure variation from individual to individual given a covariates vector $\boldsymbol{x}_i^\top = (x_{i1}, \ldots, x_{ip})$. We shall use a logistic link to the covariates, so that the probability that individual i is cured is modeled by

$$\pi(\boldsymbol{x}_i) = \frac{\exp\big(\boldsymbol{x}_i^\top \boldsymbol{\gamma}\big)}{1 + \exp\big(\boldsymbol{x}_i^\top \boldsymbol{\gamma}\big)}, \tag{9}$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)^\top$ indicates the long-term effects. The LβBSM model is defined assuming that the survival function for susceptible individuals in (8) is given by $S(y_i|N_i = 1) = 1 - I_{\Phi(\xi_2)}(a, b)$, where ξ_2 is defined in Section 2. For this model, the parameters of interest are $\boldsymbol{\theta} = (a, b, \alpha, \mu, \sigma, \boldsymbol{\gamma}^\top)^\top$. The L$\beta$BSM model, when $\pi(\boldsymbol{x}_i) = 0$ for all $\boldsymbol{x}_i$, $a = b = 1$ and $\sigma = 2$ reduces to the log-BS regression model. The identifiability between the parameters in the cure fraction and those in the latency distribution for the mixture model has been discussed by Li et al. (2005). The mixture model is not identifiable when the cure fraction $\pi(x)$ is a constant π, but is identifiable when $\pi(x)$ is modeled by a logistic regression with non-constant covariates (Li et al., 2005). So, it is necessary to include some covariates in the cure fraction to ensure identifiability. The LβBSM model contains, as special sub-models, the log-exponentiated BS mixture (LEBSM) model and log-BS mixture (LBSM) model when $b = 1$ and $a = b = 1$, respectively.

Consider data in the form $(y_i, \boldsymbol{x}_i)$, where the response variable y_i corresponds to the observed log-lifetime or log-censoring time for the ith individual and $\boldsymbol{x}_i$ is a covariate vector, for $i = 1, \ldots, n$. Under this assumption, the contribution of an individual that failed at y_i to the likelihood function is $\big[1 - \pi(\boldsymbol{x}_i)\big]\xi_{i1}^* \exp(-\xi_{i2}^{*2}/2)\Phi(\xi_{i2}^*)^{a-1}[1 - \Phi(\xi_{i2}^*)]^{b-1}/\{\sqrt{2\pi}\sigma B(a, b)\}$ and the contribution of an individual that is at risk at y_i is $\pi(\boldsymbol{x}_i) + [1 - \pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a, b)]$, where

$$\xi_{i1}^* = \xi_{i1}^*(\boldsymbol{\theta}) = \frac{2}{\alpha}\cosh\left(\frac{y_i - \mu}{\sigma}\right) \quad \text{and} \quad \xi_{i2}^* = \xi_{i2}^*(\boldsymbol{\theta}) = \frac{2}{\alpha}\sinh\left(\frac{y_i - \mu}{\sigma}\right).$$

The total log-likelihood function for the parameter vector $\boldsymbol{\theta} = (a, b, \alpha, \sigma, \mu, \boldsymbol{\gamma}^\top)^\top$ is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= q \log\left[\frac{(2\pi)^{-1/2}}{\sigma B(a, b)}\right] + \sum_{i\in D} \log(\xi_{i1}^*) - \frac{1}{2}\sum_{i\in D} \xi_{i2}^{*2} + (a - 1)\sum_{i\in D} \log\big[\Phi(\xi_{i2}^*)\big] \\ &\quad + (b - 1)\sum_{i\in D} \log\big[1 - \Phi(\xi_{i2}^*)\big] + \sum_{i\in D} \log[1 - \pi(\boldsymbol{x}_i)] + \sum_{i\in C} \log\Big\{\pi(\boldsymbol{x}_i) + \big[1 - \pi(\boldsymbol{x}_i)\big]\big[1 - I_{\Phi(\xi_{i2}^*)}(a, b)\big]\Big\},\end{aligned}$$

where q is the observed number of failures and D and C denote the sets of individuals corresponding to the log-lifetime and log-censoring time, respectively. The score functions for the parameters $a, b, \alpha, \sigma, \mu$ and $\boldsymbol{\gamma}$ are given by

$$U_a(\boldsymbol{\theta}) = q[\psi(a + b) - \psi(a)] + \sum_{i\in D} \log[\Phi(\xi_{i2}^*)] - \sum_{i\in C} [\dot{I}_{\Phi(\xi_{i2}^*)}(a, b)]_a^*,$$

$$U_b(\boldsymbol{\theta}) = q[\psi(a + b) - \psi(b)] + \sum_{i\in D} \log[1 - \Phi(\xi_{i2}^*)] - \sum_{i\in C} [\dot{I}_{\Phi(\xi_{i2}^*)}(a, b)]_b^*,$$

$$U_\alpha(\boldsymbol{\theta}) = -\frac{q}{\alpha} + \frac{1}{\alpha}\sum_{i\in D} \xi_{i2}^{*2} - \frac{(a - 1)}{\alpha}\sum_{i\in D} \frac{\xi_{i2}^*\phi(\xi_{i2}^*)}{\Phi(\xi_{i2}^*)} + \frac{(b - 1)}{\alpha}\sum_{i\in D} \frac{\xi_{i2}^*\phi(\xi_{i2}^*)}{1 - \Phi(\xi_{i2}^*)} + \sum_{i\in C} [\dot{I}_{\Phi(\xi_{i2}^*)}(a, b)]_\alpha^*,$$

$$\begin{aligned}U_\sigma(\boldsymbol{\theta}) &= -\frac{q}{\sigma} - \frac{1}{\sigma}\sum_{i\in D} \frac{z_i^*\xi_{i2}^*}{\xi_{i1}^*} + \frac{1}{\sigma}\sum_{i\in D} z_i^*\xi_{i1}^*\xi_{i2}^* - \frac{(a - 1)}{\sigma}\sum_{i\in D} \frac{z_i^*\,\xi_{i1}^*\phi(\xi_{i2}^*)}{\Phi(\xi_{i2}^*)} \\ &\quad + \frac{(b - 1)}{\sigma}\sum_{i\in D} \frac{z_i^*\,\xi_{i1}^*\phi(\xi_{i2}^*)}{1 - \Phi(\xi_{i2}^*)} + \sum_{i\in C} [\dot{I}_{\Phi(\xi_{i2}^*)}(a, b)]_\sigma^*,\end{aligned}$$

$$U_\mu(\boldsymbol{\theta}) = -\frac{1}{\sigma}\sum_{i\in D} \frac{\xi_{i2}^*}{\xi_{i1}^*} + \frac{1}{\sigma}\sum_{i\in D} \xi_{i1}^*\xi_{i2}^* - \frac{(a - 1)}{\sigma}\sum_{i\in D} \frac{\xi_{i1}^*\phi(\xi_{i2}^*)}{\Phi(\xi_{i2}^*)} + \frac{(b - 1)}{\sigma}\sum_{i\in D} \frac{\xi_{i1}^*\phi(\xi_{i2}^*)}{1 - \Phi(\xi_{i2}^*)} + \sum_{i\in C} [\dot{I}_{\Phi(\xi_{i2}^*)}(a, b)]_\mu^*$$

and $\boldsymbol{U}_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = \boldsymbol{X}^\top \boldsymbol{s}^*$, respectively. Here, $z_i^* = (y_i - \mu)/\sigma$, $\boldsymbol{s}^* = (s_1^*, \ldots, s_n^*)^\top$ with

$$s_i^* = \begin{cases} -\pi(\boldsymbol{x}_i) & i \in D, \\ \dfrac{\pi(\boldsymbol{x}_i)[1-\pi(\boldsymbol{x}_i)]I_{\Phi(\xi_{i2}^*)}(a,b)}{\pi(\boldsymbol{x}_i) + [1-\pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a,b)]}, & i \in C, \end{cases}$$

$$[\dot{I}_{\Phi(\xi_{i2}^*)}(a,b)]_a^* = \frac{[1-\pi(\boldsymbol{x}_i)]\bar{I}^{(0)}_{\Phi(\xi_{i2}^*)}(a,b) - [\psi(a) - \psi(a+b)]I_{\Phi(\xi_{i2}^*)}(a,b)}{\pi(\boldsymbol{x}_i) + [1-\pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a,b)]},$$

$$[\dot{I}_{\Phi(\xi_{i2}^*)}(a,b)]_b = \frac{[1-\pi(\boldsymbol{x}_i)]\bar{I}^{(1)}_{\Phi(\xi_{i2}^*)}(a,b) - [\psi(b) - \psi(a+b)]I_{\Phi(\xi_{i2}^*)}(a,b)}{\pi(\boldsymbol{x}_i) + [1-\pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a,b)]},$$

$$[\dot{I}_{\Phi(\xi_{i2}^*)}(a,b)]_\alpha^* = \frac{[1-\pi(\boldsymbol{x}_i)]\xi_{i2}^*\phi(\xi_{i2}^*)\Phi(\xi_{i2}^*)^{a-1}[1-\Phi(\xi_{i2}^*)]^{b-1}}{\alpha B(a,b)\{\pi(\boldsymbol{x}_i) + [1-\pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a,b)]\}},$$

$$[\dot{I}_{\Phi(\xi_{i2}^*)}(a,b)]_\sigma^* = \frac{[1-\pi(\boldsymbol{x}_i)]z_i^*\xi_{i1}^*\phi(\xi_{i2}^*)\Phi(\xi_{i2}^*)^{a-1}[1-\Phi(\xi_{i2}^*)]^{b-1}}{\sigma B(a,b)\{\pi(\boldsymbol{x}_i) + [1-\pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a,b)]\}},$$

$$[\dot{I}_{\Phi(\xi_{i2}^*)}(a,b)]_\mu^* = \frac{[1-\pi(\boldsymbol{x}_i)]\xi_{i1}^*\phi(\xi_{i2}^*)\Phi(\xi_{i2}^*)^{a-1}[1-\Phi(\xi_{i2}^*)]^{b-1}}{\sigma B(a,b)\{\pi(\boldsymbol{x}_i) + [1-\pi(\boldsymbol{x}_i)][1 - I_{\Phi(\xi_{i2}^*)}(a,b)]\}},$$

and

$$\bar{I}^{(k)}_{\Phi(\xi_{i2}^*)}(a,b) = \frac{1}{B(a,b)}\int_0^{\Phi(\xi_{i2}^*)} [\log(w)]^{1-k}[\log(1-w)]^k w^{a-1}(1-w)^{b-1}dw, \quad k = 0, 1.$$

The MLEs of the parameters in $\boldsymbol{\theta}$ can be obtained by solving simultaneously the nonlinear equations $U_a(\boldsymbol{\theta}) = 0$, $U_b(\boldsymbol{\theta}) = 0$, $U_\alpha(\boldsymbol{\theta}) = 0$, $U_\sigma(\boldsymbol{\theta}) = 0$, $U_\mu(\boldsymbol{\theta}) = 0$ and $\boldsymbol{U}_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = \boldsymbol{0}$. The covariances of the MLEs in $\widehat{\boldsymbol{\theta}}$ can also be obtained using the Hessian matrix. Under standard regularity conditions, confidence intervals and hypothesis tests can be conducted based on the large sample distribution of the MLE, which is multivariate normal with covariance matrix given by the inverse of the expected information matrix, i.e. $\widehat{\boldsymbol{\theta}} \sim \mathcal{N}_{p+5}(\boldsymbol{\theta}, \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1})$, where the asymptotic covariance matrix is given by $\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}$, $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = -E(\ddot{\boldsymbol{L}}_{\theta\theta})$ and $\ddot{\boldsymbol{L}}_{\theta\theta} = \partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$. Since it is not possible to compute the expected information matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ due to the censored observations (censoring is random and noninformative), we can use the matrix of second derivatives $-\ddot{\boldsymbol{L}}_{\theta\theta}$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ to estimate $\boldsymbol{\Sigma}(\boldsymbol{\theta})$.

More recently, several authors have adopted only a regression structure for the cure probability in long-term survivor models. See, for example, de Castro et al. (2009, 2010) and Rodrigues et al. (2011), among others. Here, we consider the same approach for the LβBS regression model with cure fractions. However, as suggested by a referee, a future research can be conducted to include covariates for the cure probability, i.e. $\pi(\boldsymbol{x}_i) = \exp(\boldsymbol{x}_i^\top\boldsymbol{\gamma})/\{1 + \exp(\boldsymbol{x}_i^\top\boldsymbol{\gamma})\}$, and possibly other covariates and a vector of new parameters, say $\boldsymbol{\zeta}$, for the logarithm of the survival time such as

$$\xi_{i1}^* = \xi_{i1}^*(\boldsymbol{\theta}) = \frac{2}{\alpha}\cosh\left(\frac{y_i - \boldsymbol{z}_i^\top\boldsymbol{\zeta}}{\sigma}\right) \quad \text{and} \quad \xi_{i2}^* = \xi_{i2}^*(\boldsymbol{\theta}) = \frac{2}{\alpha}\sinh\left(\frac{y_i - \boldsymbol{z}_i^\top\boldsymbol{\zeta}}{\sigma}\right).$$

7. Applications

7.1. First application: the LβBS regression model

In this section, we use a real data set to show the flexibility and applicability of the LβBS regression model. We compare the results from the fits of the LβBS, LEBS, LLeBS and LBS regression models. All the computations were done using the Ox matrix programming language (Doornik, 2006). The Ox program is freely distributed for academic purposes and available at http://www.doornik.com.

We shall consider the real data set given by Hirose (1993) as the results of an accelerated life-test on polyethylene terephthalate (PET) film (used in electrical insulation) in SF_6 gas insulated transformers. The accelerated life test was performed at four levels of voltage: $v = 5, 7, 10$ and 15, with 10, 15, 10 and 9 observations for each level, respectively. Three censored values were observed at $v = 5$. The data are listed in Table 1. They have also been considered by Wang and Kececioglu (2000) as an illustration of the log-linear Weibull model to accelerated life-test.

The aim of the study is to relate the resistance times of insulating films (t) with the levels of voltage (v). We consider the following regression model:

$$y_i = \beta_0 + \beta_1 v_i + \sigma z_i,$$

Table 1
Resistance times of insulating films.

Voltage (kV)	Failure or censoring time (h)					
5	7131	8482	8559	8762	9026	9034
	9104	9104.25[a]	9104.25[a]	9104.25[a]		
7	50.25	87.75	87.76	87.77	92.90	92.91
	95.96	108.3	108.3	117.9	123.9	124.3
	129.7	135.6	135.6			
10	15.17	19.87	20.18	21.50	21.88	22.23
	23.02	23.90	28.17	29.70		
15	2.40	2.42	3.17	3.75	4.65	4.95
	6.23	6.68	7.30			

[a] Indicates censored data.

Table 2
MLEs of the parameters (standard errors in parentheses and p-values in [·]) and the AIC, BIC and HQIC measures.

Model	a	b	α	σ	β_0	β_1	AIC	BIC	HQIC
LβBS	0.6614	1.4563	135.9833	0.4147	9.3643	−0.4077	91.68	102.38	95.65
	(0.842)	(1.414)	(131.946)	(0.055)	(0.166)	(0.017)			
					[<0.01]	[<0.01]			
LEBS	0.4143		102.9348	0.4165	9.3605	−0.4071	89.83	98.76	93.14
	(0.131)		(72.866)	(0.053)	(0.165)	(0.016)			
					[<0.01]	[<0.01]			
LLeBS		1.9841	167.5896	0.4148	9.3623	−0.4080	89.77	98.69	93.08
		(0.383)	(115.231)	(0.056)	(0.158)	(0.016)			
					[<0.01]	[<0.01]			
LBS			246.1849	0.3695	9.1815	−0.4051	98.72	105.86	101.37
			(180.666)	(0.046)	(0.138)	(0.016)			
					[<0.01]	[<0.01]			

Table 3
LR statistics.

Model	Hypotheses	w	p-value
LβBS vs. LEBS	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	0.155	0.694
LβBS vs. LLeBS	$\mathcal{H}_0 : a = 1$ vs. $\mathcal{H}_1 : a \neq 1$	0.094	0.760
LβBS vs. LBS	$\mathcal{H}_0 : a = b = 1$ vs. $\mathcal{H}_1 : \mathcal{H}_0$ is false	11.039	0.004
LEBS vs. LBS	$\mathcal{H}_0 : a = 1$ vs. $\mathcal{H}_1 : a \neq 1$	10.884	0.001
LLeBS vs. LBS	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	10.946	0.001

where y_i has the LβBS distribution (4), for $i = 1, \ldots, 44$. Table 2 lists the MLEs (standard errors in parentheses) of the model parameters of the LβBS, LEBS, LLeBS and LBS regression models fitted to the data and the statistics: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and HQIC (Hannan–Quinn Information Criterion). Notice that the MLEs of σ, β_0 and β_1 (and their respective standard errors) are approximately the same for all the models. Also, note that the MLEs of α (and the standard errors) are very different for these models. Although the MLEs of a and b are approximately equal to these models, their corresponding standard errors are different. The MLEs of the parameters a, b and α are different for each model because they are shape parameters.

The figures in Table 2 indicate that the LLeBS (new) regression model has the lowest AIC, BIC and HQIC values among those of the fitted models, and so it could be chosen as the best model. The LβBS and LEBS regression models also outperform the LBS model according to these statistics. In summary, the LβBS, LEBS and LLeBS regression models outperform the LBS model irrespective of the criteria and they can be effectively used in the analysis of these data. For the fitted regression models, note that β_1 is marginally significant at the level of 1% and then there is a significant difference among the levels of the voltage for the resistance times of insulating films.

A comparison of the LβBS regression model with some of its sub-models using LR statistics is performed in Table 3. The figures in this table, specially the p-values, indicate that the LβBS regression model gives the same fit to the current data than those of the LEBS and LLeBS regression models. Additionally, these models yield better fits to the data than the LBS regression model. A graphical comparison among the LβBS, LEBS, LLeBS and LBS models is explored in Fig. 4. These plots provide the Kaplan–Meier (KM) estimate and the estimated survival functions of the LβBS, LEBS, LLeBS and LBS regression models. Based on these plots, it is evident that these models fit the current data better than the LBS model. As expected, the curves for the LβBS model is very similar to the curves of the LEBS and LLeBS models.

In what follows, we shall apply the local influence method for the purpose of identifying influential observations in the LβBS, LEBS, LLeBS and LBS regression models fitted to the data. Fig. 5 gives the influence index plot for these models based

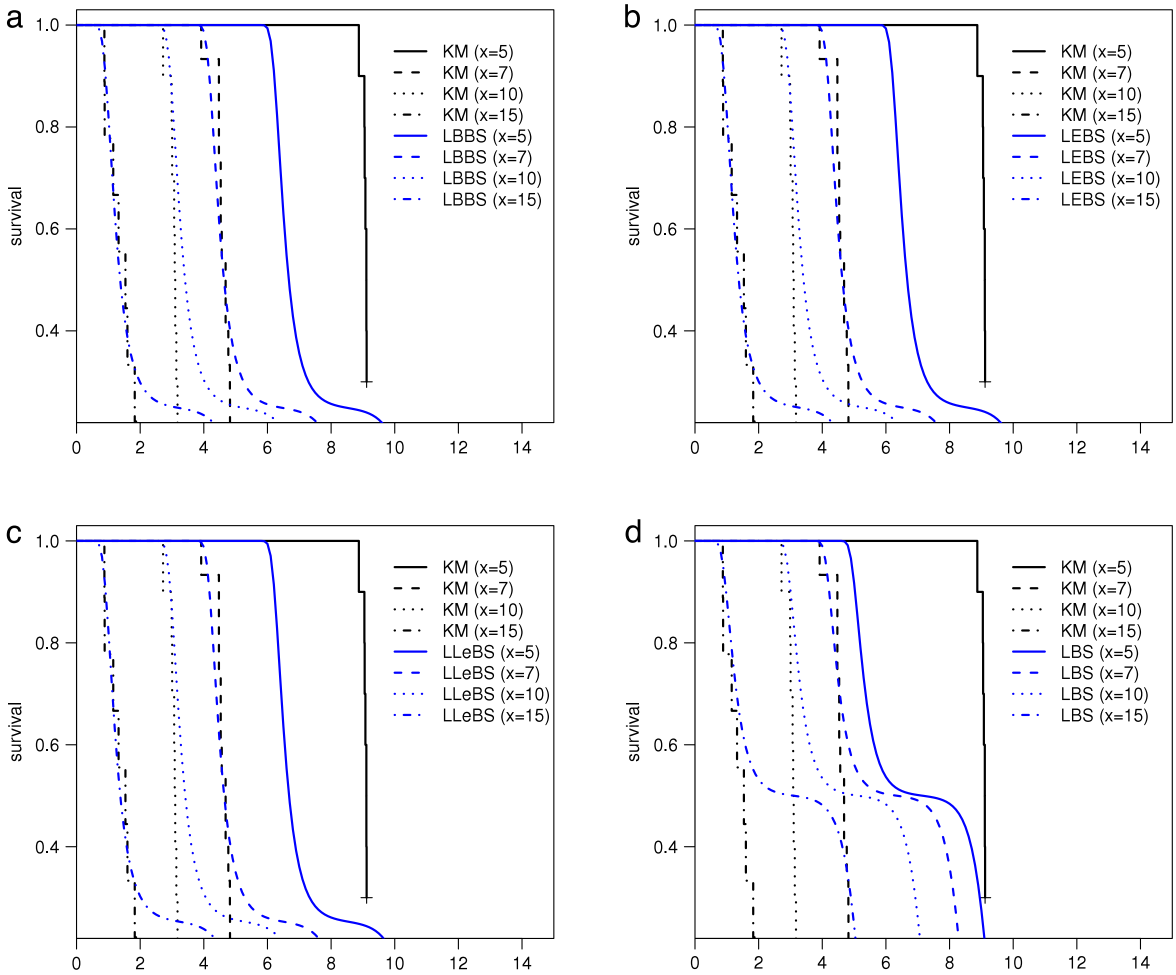

Fig. 4. Estimated survival functions and the empirical survival: (a) LβBS regression model versus KM; (b) LEBS regression model versus KM; (c) LLeBS regression model versus KM; (d) LBS regression model versus KM.

Table 4
Relative changes (%) dropping the labeled cases.

Model	Dropping											
	#11						#26					
	$\widehat{a}$	$\widehat{b}$	$\widehat{\alpha}$	$\widehat{\sigma}$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{a}$	$\widehat{b}$	$\widehat{\alpha}$	$\widehat{\sigma}$	$\widehat{\beta}_1$	$\widehat{\beta}_2$
LβBS	24.6	17.7	65.5	12.1	2.4	7.6	8.3	4.6	33.9	5.7	0.1	0.2
LEBS	0.6		44.4	11.8	2.5	7.7	2.3		29.4	5.7	0.1	0.2
LLeBS		0.0	48.1	12.1	2.4	7.5		1.2	28.7	5.7	0.1	0.2
LBS			36.2	10.2	2.7	7.6			28.3	5.1	0.1	0.2

on the case weight perturbation. An inspection of these plots reveal that the cases #11 and #26 have more pronounced influence on the MLEs than the other observations. They correspond to the smallest observations for the levels of voltage 7 and 10, respectively. Based on Fig. 5, we eliminated those most influential observations and refitted the LβBS, LEBS, LLeBS and LBS regression models. The relative change (RC), in percentage, of each parameter estimate is used to evaluate the effect of the potentially influential case. The RC is defined by $\mathrm{RC}_\theta(i) = |\widehat{\theta} - \widehat{\theta}_{(i)}/\widehat{\theta}| \times 100\%$, where $\widehat{\theta}_{(i)}$ denotes the MLE of θ after removing the ith observation. The results are listed in Table 4. This table indicates that the relative changes for the MLE of the parameter α for the four models are very pronounced, mainly for the observation #11. However, the inferences do not change at the significance level of 1%, i.e., the significance of the covariable is not influenced by these observations.

As pointed out by an anonymous referee, it would be interesting to investigate the effect of influential cases on the LR statistics. In Table 5, we present the LR statistics when the influential cases are excluded. Note that we arrive at the same conclusion when all observations are considered; compare the figures of this table with the figures in Table 3.

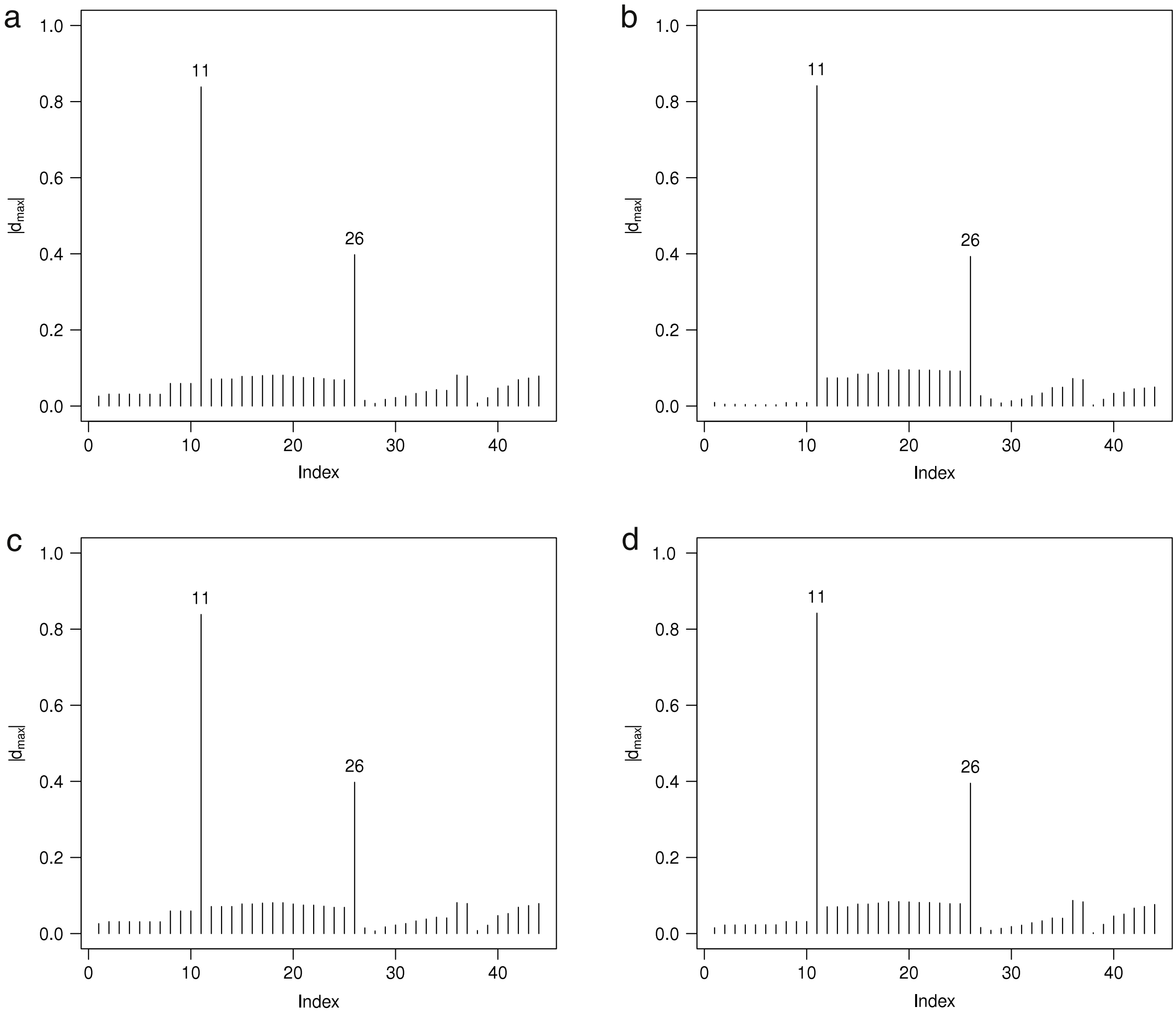

Fig. 5. Influence index plots (case weight perturbation) for the LβBS (a), LEBS (b), LLeBS (c) and LBS (d) models.

Table 5
LR statistics dropping the influential cases.

Model	Hypotheses	w	p-value
Dropping #11			
LβBS vs. LEBS	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	0.184	0.668
LβBS vs. LLeBS	$\mathcal{H}_0 : a = 1$ vs. $\mathcal{H}_1 : a \neq 1$	0.016	0.898
LβBS vs. LBS	$\mathcal{H}_0 : a = b = 1$ vs. $\mathcal{H}_1 : \mathcal{H}_0$ is false	10.850	0.004
LEBS vs. LBS	$\mathcal{H}_0 : a = 1$ vs. $\mathcal{H}_1 : a \neq 1$	10.666	0.001
LLeBS vs. LBS	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	10.833	0.001
Dropping #26			
LβBS vs. LEBS	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	0.160	0.690
LβBS vs. LLeBS	$\mathcal{H}_0 : a = 1$ vs. $\mathcal{H}_1 : a \neq 1$	0.053	0.818
LβBS vs. LBS	$\mathcal{H}_0 : a = b = 1$ vs. $\mathcal{H}_1 : \mathcal{H}_0$ is false	10.381	0.006
LEBS vs. LBS	$\mathcal{H}_0 : a = 1$ vs. $\mathcal{H}_1 : a \neq 1$	10.221	0.001
LLeBS vs. LBS	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	10.328	0.001

In summary, the proposed LβBS, LEBS and LLeBS regression models produce better fit for the current data than the LBS regression model (Rieck and Nedelman, 1991). In this case, the LLeBS regression model could be chosen since it has less parameters to be estimated and according to the LR statistic (see Table 3), it presents a similar fit to that of the LβBS regression model. Also, this regression model gives the lowest AIC, BIC and HQIC values (see Table 2). For example, we may

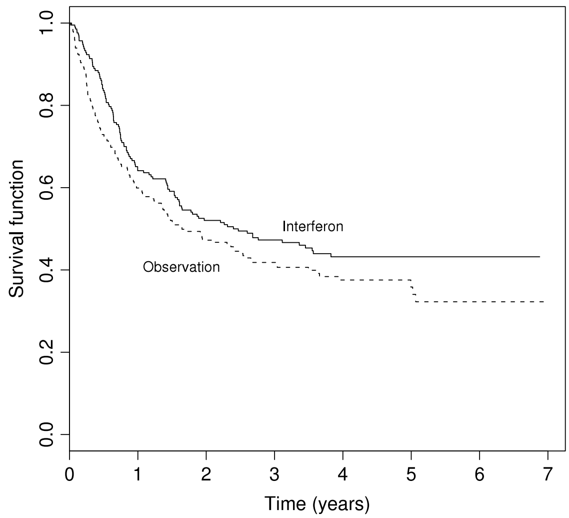

Fig. 6. Kaplan–Meier curves of high-dose interferon and observation groups in cutaneous melanoma data.

interpret the estimated coefficient of the LLeBS model as follows. The expected resistance time of insulating films should decrease approximately 34%[$(1 - e^{-0.4080}) \times 100\%$] as the level of voltage increases one unity.

7.2. *Second application: the* LβBS *mixture model*

The data are part of a study on cutaneous melanoma (a type of malignant cancer) for the evaluation of postoperative treatment performance with a high dose of a certain drug (interferon alfa-2b) in order to prevent recurrence. Patients were included in the study from 1991 to 1995, and follow-up was conducted until 1998. The data are collected by Ibrahim et al. (2001) and represent the survival times, T, as the time until the patient's death. The original sample size was $n = 427$ patients, 10 of whom did not present a value for explanatory variable tumor thickness. When such cases were removed, a sample of size $n = 417$ patients was retained. The percentage of censored observations was 56%. The following variables are associated with each participant ($i = 1, \ldots, 417$): y_i: observed time (in years); x_{i1}: treatment (0: observation, 1: interferon); x_{i2}: age (in years); x_{i3}: nodule (nodule category: 1–4); x_{i4}: sex (0: male, 1: female); x_{i5}: p.s. (performance status-patient's functional capacity scale as regards his daily activities—0: fully active, 1: other) and x_{i6}: tumor (tumor thickness in mm).

Fig. 6 shows the estimated survival curves for interferon and the observation groups. An obvious plateau can be observed after about a 5 years' follow-up, which offers empirical evidence for a cure possibility in cutaneous melanoma data.

Firstly, we consider the following LβBSM model described in Section 6

$$y_i = \mu + \sigma z_i, \quad i = 1, \ldots, 417,$$

where y_i has the LβBS distribution (4) and

$$\pi(\boldsymbol{x}_i) = \frac{\exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma})}{1 + \exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma})},$$

where $\boldsymbol{x}_i^\top \boldsymbol{\gamma} = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2} + \gamma_3 x_{i3} + \gamma_4 x_{i4} + \gamma_5 x_{i5} + \gamma_6 x_{i6}$.

To obtain the MLEs of the parameters in the LβBSM model, we used the procedure `NLMixed` in SAS, whose results are listed in Table 6. We note that the covariate *nodule* is significant (at 5%) for the cure fraction. Further, the predictors *age* and *tumor* are significant for the cure fraction (at 10%).

A summary of the values of the AIC, CAIC and BIC statistics to compare the LβBSM model with some of its sub-models is given in Table 7. The LβBSM model yields the best fit according to these criteria.

A comparison of the LβBS regression model with some of its sub-models using LR statistics is performed in Table 8. The figures in this table, specially the p-values, indicate that the LβBSM model provides a better representation of the data than the LEBSM and LBSM models.

Next, we turn to a simplified model retaining only nodule category as an explanatory variable. The estimates for the LβBSM regression model with long-term survivors fitted to the cutaneous melanoma data are listed in Table 9.

Finally, we estimate the proportion of cured individuals, using Eq. (9), by

$$\widehat{\pi}_i = \frac{\exp(1.1171 - 0.4878 x_{i3})}{1 + \exp(1.1171 - 0.4878 x_{i3})} \quad \text{and} \quad \widehat{\pi} = \frac{1}{417} \sum_{i=1}^{417} \widehat{\pi}_i.$$

Table 6
MLEs for the parameters of the LβBSM model with long-term survivors fitted to the cutaneous melanoma data set.

Parameter	Estimate	S.E.	p-value
a	1.3587	0.3010	–
b	12.4847	0.1287	–
α	8.7935	3.9743	–
μ	9.5969	3.0135	<0.001
σ	3.4191	0.3403	–
γ_0	4.3551	1.7167	0.0115
γ_1	−0.2245	0.3948	0.5700
γ_2	−0.0281	0.0162	0.0831
γ_3	−1.3671	0.5859	0.0201
γ_4	0.2726	0.3919	0.4871
γ_5	−0.0665	0.6053	0.9126
γ_6	−0.2542	0.1371	0.0645

Table 7
Some statistics for comparing the LβBSM model with some of its sub-models.

Model	AIC	CAIC	BIC
LβBSM	905.9	906.6	954.3
LEBSM	941.3	942.0	985.7
LBSM	943.5	944.0	983.8

Table 8
LR statistics.

Model	Hypotheses	w	p-value
LβBSM vs. LEBSM	$\mathcal{H}_0 : b = 1$ vs. $\mathcal{H}_1 : b \neq 1$	37.4	<0.001
LβBSM vs. LBSM	$\mathcal{H}_0 : a = b = 1$ vs. $\mathcal{H}_1 : \mathcal{H}_0$ is false	41.6	<0.001

Table 9
MLEs for the LβBSM model fitted to the cutaneous melanoma data.

Parameter	Estimate	S.E.	p-value
a	1.6727	0.0402	–
b	191.78	0.3336	–
α	0.8169	0.0892	–
μ	8.2832	3.0057	0.006
σ	8.7371	4.3878	–
γ_0	1.1171	0.2791	<0.001
γ_3	−0.4878	0.1105	<0.001

The mean cure fraction estimated was $\widehat{\pi} = 0.4955$. Estimates of the cure rate patients stratified by nodule category are $\widehat{\pi}_j$, for $j = 1, \ldots, 4$. The estimates of the surviving fraction of patients stratified by nodule category from 1 to 4 are 0.6523, 0.5353, 0.4143 and 0.3028, respectively.

8. Concluding remarks

For the first time, we study the called log-β-Birnbaum–Saunders (LβBS) distribution. We derive explicit expressions for the moment generating function and moments. Based on this distribution, we propose a LβBS regression model very suitable for modeling censored and uncensored lifetime data. The new regression model serves as a good alternative for lifetime data analysis and it is much more flexible than the log-Birnbaum–Saunders regression model (Rieck and Nedelman, 1991) in many practical situations. The parameter estimation is approached by maximum likelihood and the observed information matrix is derived. We also discuss influence diagnostics in the LβBS regression model fitted to censored data. We also propose a LβBS mixture model for survival data with long-term survivors. The usefulness of the new regression model is illustrated by means of two real data sets. Our formulas related with the LβBS regression model are manageable, and with the use of modern computer resources with analytic and numerical capabilities, may turn into adequate tools comprising the arsenal of applied statisticians. In other words, the proposed methodology can be implemented straightforwardly and runs immediately in some statistical packages. We hope that the proposed regression model may attract wider applications in survival analysis and fatigue life modeling.

Acknowledgments

We gratefully acknowledge grants from CNPq and FAPESP (Brazil). We are also grateful to two referees and the editor for helpful comments and suggestions.

Appendix A

The proof of the theorem is as follows. First, if a is a positive real non-integer, we can expand $\Phi(v)^a$ as

$$\Phi(v)^a = \sum_{r=0}^{\infty} s_r(a)\Phi(v)^r, \tag{10}$$

where $s_r(x) = \sum_{j=r}^{\infty}(-1)^{r+j}\binom{x}{j}\binom{j}{r}$. We can write from the binomial expansion

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{a-1}\left[1-\Phi\left(\frac{2}{\alpha}\sinh(z)\right)\right]^{b-1} = \sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{i+a-1}$$

and then using (10)

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{a-1}\left[1-\Phi\left(\frac{2}{\alpha}\sinh(z)\right)\right]^{b-1} = \sum_{i,r=0}^{\infty}(-1)^i\binom{b-1}{i}s_r(i+a-1)\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^r.$$

Hence,

$$M_Z(s) = \sum_{i,r=0}^{\infty} p_{i,r}\int_{-\infty}^{\infty}\exp(sz)\cosh(z)\exp\{-2\sinh^2(z)/\alpha^2\}\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^r dz,$$

where $p_{i,r} = p_{i,r}(a, b, \alpha)$ is defined above.

We require the following results for the error function erf$(\cdot)$ to calculate the last integral, say $N_r(s,\alpha)$: $\Phi(x) = [1 + \mathrm{erf}(x/\sqrt{2})]/2$ and $\mathrm{erf}(x) = (2/\sqrt{\pi})\int_0^x \exp(-y^2)dy$. If $b_m = (-1)^m[(2m+1)2^{m/2}m!\sqrt{\pi}]^{-1}$, we can write the power series $\mathrm{erf}(x/\sqrt{2}) = \sum_{m=0}^{\infty} b_m x^{2m+1}$. We use the equation $\left(\sum_{i=0}^{\infty} a_i x^i\right)^j = \sum_{i=0}^{\infty} c_{j,i}x^i$ for a power series raised to a positive integer j (Gradshteyn and Ryzhik, 2007, Section 0.314), whose coefficients $c_{j,i}$ (for $i = 1, 2, \ldots$) are determined from the recurrence equation

$$c_{j,i} = (ia_0)^{-1}\sum_{m=1}^{i}(jm - i + m)a_m c_{j,i-m} \tag{11}$$

and $c_{j,0} = a_0^j$. Hence, the coefficients $c_{j,i}$ can be calculated directly from $c_{j,0}, \ldots, c_{j,i-1}$ and, therefore, from $a_0, \ldots, a_i$. We have

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^r = \frac{1}{2^r}\left\{1+\sum_{m=0}^{\infty} d_m \sinh(z)^{2m+1}\right\}^r,$$

where $d_m = 2^{2m+1}\alpha^{-(2m+1)}b_m$. Thus, using (11), we can obtain

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^r = \frac{1}{2^r}\sum_{k=0}^{r}\binom{r}{k}\left(\sum_{m=0}^{\infty} d_m\sinh(z)^{2m+1}\right)^k = \sum_{m=0}^{\infty} e_{m,r}\sinh(z)^{2m+1},$$

where $e_{m,r} = 2^{-r}\sum_{k=0}^{r}\binom{r}{k}g_{k,m}$, $g_{k,0} = d_0^j$ and $g_{k,m} = (i\,d_0)^{-1}\sum_{\ell=1}^{m}(k\ell - m + \ell)d_\ell g_{k,m-\ell}$. Further,

$$N_r(s,\alpha) = \sum_{m=0}^{\infty} e_{m,r}\int_{-\infty}^{\infty}\exp(sz)\cosh(z)\sinh(z)^{2m+1}\exp\{-2\sinh^2(z)/\alpha^2\}\,dz.$$

From the identity $\cosh(2z) = 2\sinh^2(z) + 1$ and the definition of $\sinh(z)$ and $\cosh(z)$, by expanding the binomial term, we obtain after some algebra

$$N_r(s,\alpha) = \exp(1/\alpha^2)\sum_{m=0}^{\infty}\frac{e_{m,r}}{2^{m+3}}\sum_{j=0}^{2m+1}(-1)^j\binom{2m+1}{j}$$
$$\times\int_{-\infty}^{\infty}\{\exp[(m+1-j+s/2)x] + \exp[(m-j+s/2)x]\}\,\exp\{-\cosh(x)/\alpha^2\}dx.$$

From the integral representation $K_\nu(z) = 0.5\int_{-\infty}^{\infty} \exp\{-z\cosh(x) - \nu x\}dx$, it follows that

$$N_r(s,\alpha) = e^{\alpha^{-2}} \sum_{m=0}^{\infty} \frac{e_{m,r}}{2^{m+2}} \sum_{j=0}^{2m+1} (-1)^j \binom{2m+1}{j} \left[K_{-(m+1-j+s/2)}(1/\alpha^2) + K_{-(m-j+s/2)}(1/\alpha^2)\right]. \tag{12}$$

Hence, the LβBS generating function takes the form $M_Z(s) = \sum_{i,r=0}^{\infty} p_{i,r} N_r(s,\alpha)$, where $N_r(s,\alpha)$ is calculated from (12).

Appendix B

The elements of the Hessian matrix

$$\ddot{\boldsymbol{L}}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} = \begin{pmatrix} \ddot{L}_{aa} & \ddot{L}_{ab} & \ddot{L}_{a\alpha} & \ddot{L}_{a\sigma} & \ddot{L}_{a\beta_j} \\ \cdot & \ddot{L}_{bb} & \ddot{L}_{b\alpha} & \ddot{L}_{b\sigma} & \ddot{L}_{b\beta_j} \\ \cdot & \cdot & \ddot{L}_{\alpha\alpha} & \ddot{L}_{\alpha\sigma} & \ddot{L}_{\alpha\beta_j} \\ \cdot & \cdot & \cdot & \ddot{L}_{\sigma\sigma} & \ddot{L}_{\sigma\beta_j} \\ \cdot & \cdot & \cdot & \cdot & \ddot{L}_{\beta_j\beta_s} \end{pmatrix},$$

are determined. After extensive algebraic manipulations, we obtain

$$\ddot{L}_{aa} = q[\psi'(a+b) - \psi'(a)] + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{aa},$$

$$\ddot{L}_{ab} = q\psi'(a+b) + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{ab},$$

$$\ddot{L}_{a\alpha} = -\frac{1}{\alpha}\sum_{i\in D}\frac{\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{a\alpha},$$

$$\ddot{L}_{a\sigma} = -\frac{1}{\sigma}\sum_{i\in D}\frac{z_i\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{a\sigma},$$

$$\ddot{L}_{a\beta_j} = -\frac{1}{\sigma}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{a\beta_j},$$

$$\ddot{L}_{bb} = q[\psi'(a+b) - \psi'(b)] + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{bb},$$

$$\ddot{L}_{b\alpha} = \frac{1}{\alpha}\sum_{i\in D}\frac{\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{b\alpha},$$

$$\ddot{L}_{b\sigma} = \frac{1}{\sigma}\sum_{i\in D}\frac{z_i\xi_{i1}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{b\sigma},$$

$$\ddot{L}_{b\beta_j} = \frac{1}{\sigma}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{b\beta_j},$$

$$\begin{aligned}\ddot{L}_{\alpha\alpha} = {} & \frac{q}{\alpha^2} - \frac{3}{\alpha^2}\sum_{i\in D}\xi_{i2}^2 + \frac{2(a-1)}{\alpha^2}\sum_{i\in D}\frac{\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} - \frac{(a-1)}{\alpha^2}\sum_{i\in D}\frac{\xi_{i2}^3\phi(\xi_{i2})}{\Phi(\xi_{i2})} - \frac{(a-1)}{\alpha^2}\sum_{i\in D}\frac{\xi_{i2}^2\phi(\xi_{i2})^2}{\Phi(\xi_{i2})^2} \\ & - \frac{2(b-1)}{\alpha^2}\sum_{i\in D}\frac{\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} + \frac{(b-1)}{\alpha^2}\sum_{i\in D}\frac{\xi_{i2}^3\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} \\ & - \frac{(b-1)}{\alpha^2}\sum_{i\in D}\frac{\xi_{i2}^2\phi(\xi_{i2})^2}{[1-\Phi(\xi_{i2})]^2} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{\alpha\alpha},\end{aligned}$$

$$\begin{aligned}\ddot{L}_{\alpha\sigma} = {} & -\frac{2}{\sigma\alpha}\sum_{i\in D} z_i\xi_{i1}\xi_{i2} + \frac{(a-1)}{\sigma\alpha}\sum_{i\in D}\frac{z_i\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} - \frac{(a-1)}{\sigma\alpha}\sum_{i\in D}\frac{z_i\xi_{i1}\xi_{i2}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})} \\ & - \frac{(a-1)}{\sigma\alpha}\sum_{i\in D}\frac{z_i\xi_{i1}\xi_{i2}\phi(\xi_{i2})^2}{\Phi(\xi_{i2})^2} - \frac{(b-1)}{\sigma\alpha}\sum_{i\in D}\frac{z_i\xi_{i1}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} + \frac{(b-1)}{\sigma\alpha}\sum_{i\in D}\frac{z_i\xi_{i1}\xi_{i2}^2\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} \\ & - \frac{(b-1)}{\sigma\alpha}\sum_{i\in D}\frac{z_i\xi_{i1}\xi_{i2}^2\phi(\xi_{i2})^2}{[1-\Phi(\xi_{i2})]^2} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{\alpha\sigma},\end{aligned}$$

$$\begin{aligned}
\ddot{L}_{\alpha\beta_j} = & -\frac{2}{\sigma\alpha}\sum_{i\in D} x_{ij}\xi_{i1}\xi_{i2} + \frac{(a-1)}{\sigma\alpha}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} - \frac{(a-1)}{\sigma\alpha}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\xi_{i2}^2\phi(\xi_{i2})}{\Phi(\xi_{i2})} \\
& - \frac{(a-1)}{\sigma\alpha}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\xi_{i2}\phi(\xi_{i2})^2}{\Phi(\xi_{i2})^2} - \frac{(b-1)}{\sigma\alpha}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} + \frac{(b-1)}{\sigma\alpha}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\xi_{i2}^2\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} \\
& - \frac{(b-1)}{\sigma\alpha}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\xi_{i2}\phi(\xi_{i2})^2}{[1-\Phi(\xi_{i2})]^2} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{\alpha\beta_j},
\end{aligned}$$

$$\begin{aligned}
\ddot{L}_{\sigma\sigma} = & \frac{q}{\sigma^2} + \frac{1}{\sigma^2}\sum_{i\in D} z_i^2 + \frac{2}{\sigma^2}\sum_{i\in D}\frac{z_i\xi_{i2}}{\xi_{i1}} - \frac{1}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i2}^2}{\xi_{i1}^2} - \frac{1}{\sigma^2}\sum_{i\in D} z_i^2\xi_{i1}^2 - \frac{1}{\sigma^2}\sum_{i\in D} z_i^2\xi_{i2}^2 \\
& - \frac{2}{\sigma^2}\sum_{i\in D} z_i\xi_{i1}\xi_{i2} + \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \frac{2(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} \\
& - \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i1}^2\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} - \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i1}^2\phi(\xi_{i2})^2}{\Phi(\xi_{i2})^2} \\
& - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} - \frac{2(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i\xi_{i1}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} \\
& + \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i1}^2\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i^2\xi_{i1}^2\phi(\xi_{i2})^2}{[1-\Phi(\xi_{i2})]^2} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{\sigma\sigma},
\end{aligned}$$

$$\begin{aligned}
\ddot{L}_{\sigma\beta_j} = & \frac{1}{\sigma^2}\sum_{i\in D} z_i x_{ij} + \frac{1}{\sigma^2}\sum_{i\in D}\frac{x_{ij}\xi_{i2}}{\xi_{i1}} - \frac{1}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i2}^2}{\xi_{i1}^2} - \frac{1}{\sigma^2}\sum_{i\in D} z_i x_{ij}\xi_{i1}^2 - \frac{1}{\sigma^2}\sum_{i\in D} z_i x_{ij}\xi_{i2}^2 \\
& - \frac{1}{\sigma^2}\sum_{i\in D} x_{ij}\xi_{i1}\xi_{i2} + \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\phi(\xi_{i2})}{\Phi(\xi_{i2})} \\
& - \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i1}^2\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} - \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i1}^2\phi(\xi_{i2})^2}{\Phi(\xi_{i2})^2} \\
& - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}\xi_{i1}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} \\
& + \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i1}^2\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{z_i x_{ij}\xi_{i1}^2\phi(\xi_{i2})^2}{[1-\Phi(\xi_{i2})]^2} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{\sigma\beta_j},
\end{aligned}$$

$$\begin{aligned}
\ddot{L}_{\beta_j\beta_s} = & \frac{1}{\sigma^2}\sum_{i\in D} x_{ij}x_{is} - \frac{1}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i2}^2}{\xi_{i1}^2} - \frac{1}{\sigma^2}\sum_{i\in D} x_{ij}x_{is}(\xi_{i1}^2+\xi_{i2}^2) + \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} \\
& - \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i1}^2\xi_{i2}\phi(\xi_{i2})}{\Phi(\xi_{i2})} + \frac{(a-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i1}^2\phi(\xi_{i2})^2}{\Phi(\xi_{i2})^2} - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} \\
& + \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i1}^2\xi_{i2}\phi(\xi_{i2})}{1-\Phi(\xi_{i2})} - \frac{(b-1)}{\sigma^2}\sum_{i\in D}\frac{x_{ij}x_{is}\xi_{i1}^2\phi(\xi_{i2})^2}{[1-\Phi(\xi_{i2})]^2} + \sum_{i\in C}\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{\beta_j\beta_s},
\end{aligned}$$

where $j, s = 1, \ldots, p$, $\left[\ddot{I}_{\Phi(\xi_{i2})}(a,b)\right]_{km} = \partial\left(\left[\dot{I}_{\Phi(\xi_{i2})}(a,b)\right]_k/[1 - I_{\Phi(\xi_{i2})}(a,b)]\right)/\partial m$ and $\psi'(\cdot)$ is the trigamma function. Here, $[\dot{I}_{\Phi(\xi_{i2})}(a,b)]_{\beta_j} = -x_{ij}\xi_{i1}\phi(\xi_{i2})\Phi(\xi_{i2})^{a-1}[1-\Phi(\xi_{i2})]^{b-1}/[\sigma B(a,b)]$ and all the others quantities were defined before.

References

Barros, M., Paula, G.A., Leiva, V., 2008. A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. Lifetime Data Analysis 14, 316–332.

Benerjee, S., Carlin, B.P., 2004. Parametric spatial cure rate models for interval-censored time-to-relapse data. Biometrics 60, 268–275.

Berkson, J., Gage, R.P., 1952. Survival curve for cancer patients following treatment. Journal of the American Statistical Association 88, 1412–1418.

Birnbaum, Z.W., Saunders, S.C., 1969a. A new family of life distributions. Journal of Applied Probability 6, 319–327.

Birnbaum, Z.W., Saunders, S.C., 1969b. Estimation for a family of life distributions with applications to fatigue. Journal of Applied Probability 6, 328–377.

Boag, J., 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. Journal of the Royal Statistical Society. Series B 11, 15–44.

Carrasco, J.M.F., Ortega, E.M.M., Paula, G.A., 2008. Log-modified Weibull regression models with censored data: sensitivity and residual analysis. Computational Statistics and Data Analysis 52, 4021–4039.

Cook, R.D., 1986. Assessment of local influence (with discussion). Journal of the Royal Statistical Society. Series B 48, 133–169.

Cook, R.D., Weisberg, S., 1982. Residuals and Influence in Regression. Chapman and Hall, London.

Cordeiro, G.M., Lemonte, A.J., 2011. The β-Birnbaum–Saunders distribution: an improved distribution for fatigue life modeling. Computational Statistics and Data Analysis 55, 1445–1461.

de Castro, M., Cancho, V.G., Rodrigues, J., 2009. A Bayesian long-term survival model parametrized in the cured fraction. Biometrical Journal 51, 443–455.

de Castro, M., Cancho, V.G., Rodrigues, J., 2010. A hands-on approach for fitting long-term survival models under the GAMLSS framework. Computer Methods and Programs in Biomedicine 97, 168–177.

Desmond, A.F., 1985. Stochastic models of failure in random environments. Canadian Journal of Statistics 13, 171–183.

Desmond, A.F., 1986. On the relationship between two fatigue-life models. IEEE Transactions on Reliability 35, 167–169.
Díaz-García, J.A., Leiva, V., 2005. A new family of life distributions based on the elliptically contoured distributions. Journal of Statistical Planning and Inference 128, 445–457.
Doornik, J.A., 2006. An Object-Oriented Matrix Language—Ox 4, 5th ed. Timberlake Consultants Press, London.
Escobar, L.A., Meeker, W.Q., 1992. Assessing influence in regression analysis with censored data. Biometrics 48, 507–528.
Farewell, V.T., 1982. The use of mixture models for the analysis of survival data with long-term survivors. Biometrics 38, 1041–1046.
Fung, W.K., Kwan, C.W., 1997. A note on local influence based on the normal curvature. Journal of the Royal Statistical Society. Series B 59, 839–843.
Gradshteyn, I.S., Ryzhik, I.M., 2007. Table of Integrals, Series, and Products. Academic Press, New York.
Guiraud, P., Leiva, V., Fierro, R., 2009. A non-central version of the Birnbaum–Saunders distribution for reliability analysis. IEEE Transactions on Reliability 58, 152–160.
Hashimoto, E.M., Ortega, E.M.M., Cancho, V.G., Cordeiro, G.M., 2010. The log-exponentiated Weibull regression model for interval-censored data. Computational Statistics and Data Analysis 54, 1017–1035.
Hirose, H., 1993. Estimation of threshold stress in accelerated life-testing. IEEE Transactions on Reliability 42, 650–657.
Ibrahim, J.G., Chen, M.H., Sinha, D., 2001. Bayesian Survival Analysis. Springer-Verlag, New York.
Kundu, D., Kannan, N., Balakrishnan, N., 2008. On the hazard function of Birnbaum–Saunders distribution and associated inference. Computational Statistics and Data Analysis 52, 2692–2702.
Kwan, C.W., Fung, W.K., 1998. Assessing local influence for specific restricted likelihood: application to factor analysis. Psychometrika 63, 35–46.
Lawless, J.F., 2003. Statistical Models and Methods for Lifetime Data. Wiley, New York.
Leiva, V., Barros, M., Paula, G.A., Galea, M., 2007. Influence diagnostics in log-Birnbaum–Saunders regression models with censored data. Computational Statistics and Data Analysis 51, 5694–5707.
Lemonte, A.J., 2011. Covariance matrix formula for Birnbaum–Saunders regression models. Journal of Statistical Computation and Simulation 81, 899–908.
Lemonte, A.J., Cordeiro, G.M., 2009. Birnbaum–Saunders nonlinear regression models. Computational Statistics and Data Analysis 53, 4441–4452.
Lemonte, A.J., Cordeiro, G.M., 2010. Asymptotic skewness in Birnbaum–Saunders nonlinear regression models. Statistics and Probability Letters 80, 892–898.
Lemonte, A.J., Cribari-Neto, F., Vasconcellos, K.L.P., 2007. Improved statistical inference for the two-parameter Birnbaum–Saunders distribution. Computational Statistics and Data Analysis 51, 4656–4681.
Lemonte, A.J., Ferrari, S.L.P., 2011a. Size and power properties of some tests in the Birnbaum–Saunders regression model. Computational Statistics and Data Analysis 55, 1109–1117.
Lemonte, A.J., Ferrari, S.L.P., 2011b. Small-sample corrections for score tests in Birnbaum–Saunders regressions. Communications in Statistics—Theory and Methods 40, 232–243.
Lemonte, A.J., Ferrari, S.L.P., 2011c. Signed likelihood ratio tests in the Birnbaum–Saunders regression model. Journal of Statistical Planning and Inference 141, 1031–1040.
Lemonte, A.J., Ferrari, S.L.P., Cribari-Neto, F., 2010. Improved likelihood inference in Birnbaum–Saunders regressions. Computational Statistics and Data Analysis 54, 1307–1316.
Lemonte, A.J., Patriota, A.G., 2011. Influence diagnostics in Birnbaum–Saunders nonlinear regression models. Journal of Applied Statistics 38, 871–884.
Lemonte, A.J., Simas, A.B., Cribari-Neto, F., 2008. Bootstrap-based improved estimators for the two-parameter Birnbaum–Saunders distribution. Journal of Statistical Computation and Simulation 78, 37–49.
Li, Y., Tiwari, R.C., Guha, S., 2005. Mixture cure survival models with dependent censoring. Harvard University Biostatistics Working Paper Series, vol. 41, pp. 211–224.
Longini, I.M., Halloran, M.E., 1996. A frailty mixture model for estimating vaccine efficacy. Applied Statistics 45, 165–173.
Maller, R., Zhou, X., 1996. Survival Analysis with Long-term Survivors. Wiley, New York.
Ortega, E.M.M., Bolfarine, H., Paula, G.A., 2003. Influence diagnostics in generalized log-gamma regression models. Computational Statistics and Data Analysis 42, 165–186.
Ortega, E.M.M., Cancho, V.G., Lachos, V.H., 2009a. Generalized log-gamma mixture model with long-term survivors: sensitivity analysis. Sankhyā Series B 71, 1–29.
Ortega, E.M.M., Cancho, V.G., Paula, G.A., 2009b. Generalized log-gamma regression models with cure fraction. Lifetime Data Analysis 15, 79–106.
Ortega, E.M.M., Rizzato, F.F., Demétrio, C.G.B., 2009c. The generalized log-gamma mixture model model with covariates: local influence and residual analysis. Statistical Methods and Applications 18, 305–331.
Peng, Y., Dear, K., 2000. A nonparametric mixture model for cure rate estimation. Biometrics 56, 237–243.
Pettit, A.N., Bin Daud, I., 1989. Case-weight measures of influence for proportional hazards regression. Journal of the Royal Statistical Society. Series C 38, 51–67.
Price, D.L., Manatunga, A.K., 2001. Modelling survival data with a cured fraction using frailty models. Statistics in Medicine 20, 1515–1527.
Rieck, J.R., Nedelman, J.R., 1991. A log-linear model for the Birnbaum–Saunders distribution. Technometrics 33, 51–60.
Rodrigues, J., de Castro, M., Balakrisnhan, N., Cancho, V.G., 2011. Destructive weighted Poisson cure rate models. Lifetime Data Analysis 17, 333–346.
Silva, G.O., Ortega, E.M.M., Cancho, V.G., 2010. Log-Weibull extended regression model: estimation, sensitivity and residual analysis. Statistical Methodology 7, 614–631.
Silva, G.O., Ortega, E.M.M., Cancho, V.G., Barreto, M.L., 2008. Log-Burr XII regression models with censored data. Computational Statistics and Data Analysis 52, 3820–3842.
Silva, G.O., Ortega, E.M.M., Cordeiro, G.M., 2009. A log-extended Weibull regression model. Computational Statistics and Data Analysis 53, 4482–4489.
Sy, J.P., Taylor, M.M.G., 2000. Estimation in a proportional hazards cure model. Biometrics 56, 227–336.
Tanaka, Y., Zhang, F., Mori, Y., 2003. Local influence in principal component analysis: relationship between the local influence and influence function approaches revisited. Computational Statistics and Data Analysis 44, 143–160.
Wang, W., Kececioglu, D.B., 2000. Fitting the Weibull log-linear model to accelerated life-test data. IEEE Transactions on Reliability 49, 217–223.
Watson, G.N., 1995. A Treatise on the Theory of Bessel Functions, 2nd ed. Cambridge University Press.
Wu, J., Wong, A.C.M., 2004. Improved interval estimation for the two-parameter Birnbaum–Saunders distribution. Computational Statistics and Data Analysis 47, 809–821.
Xie, F.C., Wei, B.C., 2007. Diagnostics analysis in censored generalized Poisson regression model. Journal of Statistical Computation and Simulation 77, 695–708.
Yu, B., Peng, Y., 2008. Mixture cure models for multivariate survival data. Computational Statistics & Data Analysis 52, 1524–1532.
Zeng, D., Yin, G., Ibrahim, G.I., 2006. Semiparametric transformation models for survival data with a cure fraction data. Journal of the American Statistical Association 101, 670–684.