Lifetime Data Anal
DOI 10.1007/s10985-008-9085-1

A new class of survival regression models with heavy-tailed errors: robustness and diagnostics

Michelli Barros · Gilberto A. Paula · Víctor Leiva

Received: 19 June 2007 / Accepted: 28 February 2008

Abstract Birnbaum-Saunders models have largely been applied in material fatigue studies and reliability analyses to relate the total time until failure with some type of cumulative damage. In many problems related to the medical field, such as chronic cardiac diseases and different types of cancer, a cumulative damage caused by several risk factors might cause some degradation that leads to a fatigue process. In these cases, BS models can be suitable for describing the propagation lifetime. However, since the cumulative damage is assumed to be normally distributed in the BS distribution, the parameter estimates from this model can be sensitive to outlying observations. In order to attenuate this influence, we present in this paper BS models, in which a Student-t distribution is assumed to explain the cumulative damage. In particular, we show that the maximum likelihood estimates of the Student-t log-BS models attribute smaller weights to outlying observations, which produce robust parameter estimates. Also, some inferential results are presented. In addition, based on local influence and deviance component and martingale-type residuals, a diagnostics analysis is derived. Finally, a motivating example from the medical field is analyzed using log-BS regression models. Since the parameter estimates appear to be very sensitive to outlying and influential observations, the Student-t log-BS regression model should attenuate such

M. Barros
Department of Mathematics and Statistics, Universidade Federal de Campina Grande, Campina Grande, Brazil
e-mail: michelli@dme.ufcg.edu.br; michelli.karinne@gmail.com

G. A. Paula (✉)
Department of Statistics, Universidade de São Paulo, São Paulo, Brazil
e-mail: giapaula@ime.usp.br

V. Leiva
Department of Statistics, Universidad de Valparaíso, Casilla 5030, Valparaíso, Chile
e-mail: victor.leiva@uv.cl; victor.leiva@yahoo.com

influences. The model checking methodologies developed in this paper are used to compare the fitted models.

Keywords Generalized Birnbaum-Saunders distribution · Likelihood methods · Local influence · Log-linear models · Residual analysis · Robustness · Sinh-normal distribution

1 Introduction

An important probability model originating from a physical problem related to material fatigue is the one derived by Birnbaum and Saunders (1969). The Birnbaum-Saunders (BS) distribution has a close relation to the normal distribution and has applications in a wide variety of fields. For details about old and new applications of the BS distributions, including the medical field, see Johnson et al. (1995, p. 651), Balakrishnan et al. (2007) and Leiva et al. (2007, 2008a,b,c).

Desmond (1985) presented the following proposals related to the BS distribution: (i) he provided a more general derivation of this distribution based on Cramér's biological model (Cramér 1999, p. 219); (ii) he demonstrated that this distribution describes the total time that passes until some type of cumulative damage produced by the development and growth of a dominant crack surpasses a threshold and causes a failure; and (iii) he strengthened the physical justification for the use of this distribution by relaxing some assumptions made earlier by Birnbaum and Saunders (1969).

In many medical problems, such as chronic cardiac diseases and different types of cancer, a cumulative damage caused by several risk factors is presented. This degradation leads to a fatigue process, whose propagation lifetimes can be adequately modeled by the BS distribution. Recently, Leiva et al. (2007) applied the classical version of the BS distribution (BS model generated from the normal law) for modeling survival times in patients with multiple myeloma by using a number of prognostic variables with censored data.

Similarly to normal models, maximum likelihood estimates (MLE) from BS models are sensitive to outlying observations. Lange et al. (1989) proposed the Student-t distribution (which we will call simply t distribution or t model) as a robust alternative to the normal case, since it has greater kurtosis than the normal model. Thus, cases which might considered as outlying under normality, might not under the t model, producing robust parameter estimates. Recently, Díaz-García and Leiva (2005) obtained a generalization of the BS distribution. The main motivation for the use of the generalized Birnbaum-Saunders (GBS) distribution is to make the kurtosis flexible (compared to the BS model). This is achieved when the normal distribution used in the derivation of the BS model is replaced by a general class of standard symmetrical distributions on $\mathbb{R}$, the real line. This class of distributions admits different degrees of kurtosis and the t model is one of these distributions. Moreover, since the BS distribution is a particular case of the GBS distribution, several properties of the classical BS distribution are transferred to its generalized version. More recently, Sanhueza et al. (2008) have presented a complete compilation of results related to the GBS model.

Statistical modeling based on the GBS distribution has not received much attention. However, for the classical BS distribution, some efforts can be found in the works by Rieck and Nedelman (1991), Owen and Padgett (1999, 2000) and Tsionas (2001). In Galea et al. (2004), Leiva et al. (2007) and Xie and Wei (2007), aspects related to influence diagnostics in log-Birnbaum-Saunders (log-BS) regression models with non-censored and censored data have been studied. Robust estimation for the BS distribution has also been discussed in Dupuis and Mills (1998).

The main objective of this paper is to propose a new class of lifetime regression models for which the errors follow the GBS distribution based on the t model (GBS-t or simply BS-t). Robustness aspects of the MLEs are discussed and some diagnostics procedures are derived. A data set from the medical field is used to compare the sensitivity of the parameter estimates from BS and BS-t models.

The article is organized as follows. A preliminary notion of the GBS and log-GBS models, mainly based on the t distribution, is presented in Sect. 2. In Sect. 3, we carry out a statistical analysis for the BS-t log-linear regression model. Specifically, we compute the maximum likelihood estimating equations by assuming no informative censoring. We show that the MLEs are down-weighted so that smaller weights are attributed to outlying observations. Also, some asymptotic inferential results are presented. Model checking methodologies such as local influence and residual analysis are described in Sect. 4. In particular, we discuss the relationship between the deviance component (d-c) and a martingale-type (m-g) residuals with the sinh residual. In Sect. 5, a motivating example is analyzed using BS and BS-t models. A model checking analysis is performed and some comparisons are made. Finally, a brief discussion is given in Sect. 6.

2 Log-BS-t models

The GBS distribution is defined in terms of standard symmetrical distributions on $\mathbb{R}$, the real line (also known as spherically contoured univariate distributions); see Fang et al. (1990). When a random variable (r.v.) T follows the GBS distribution, the notation $T \sim \text{GBS}(\alpha, \beta; f)$ is used, where α is the shape parameter, β is the scale parameter as well as the median and f is the probability density function (pdf) of the associated symmetrical distribution. Specifically, if the variate $T = \frac{\beta}{4}[\alpha Z + \sqrt{\alpha^2 Z^2 + 4}]^2 \sim \text{GBS}(\alpha, \beta; f)$, then $Z = \frac{1}{\alpha}[\sqrt{T/\beta} - \sqrt{\beta/T}]$ follows a standard symmetrical distribution on $\mathbb{R}$, the real line, which is denoted by $Z \sim \text{S}(f)$. Some properties of the GBS model are: $c\,T \sim \text{GBS}(\alpha, c\beta; f)$, with $c > 0$, and $T^{-1} \sim \text{GBS}(\alpha, \beta^{-1}; f)$. These properties establish that the GBS distribution belongs to the scale (proportionality) and of random variables closed under reciprocation (Saunders 1974) families.

In particular, if Z has the t distribution with $\nu > 0$ degrees of freedom (d.f.), denoted by $Z \sim t_\nu$, then $T = \frac{\beta}{4}[\alpha Z + \sqrt{\alpha^2 Z^2 + 4}]^2 \sim \text{GBS}(\alpha, \beta; t_\nu)$.

Rieck and Nedelman (1991) developed the sinh-normal (SN) distribution by means of the transformation $Y = \gamma + \sigma \operatorname{arcsinh}(\alpha Z/2)$, where $Z \sim \text{N}(0, 1)$, which is denoted by $Y \sim \text{SN}(\alpha, \gamma, \sigma)$. The SN distribution has as a particular case the log-BS distribution when $\gamma = \log(\beta)$ and $\sigma = 2$. The notation $Y \sim \text{log-BS}(\alpha, \gamma)$ is used in this case.

In a similar way to the SN distribution, we can define the sinh-t distribution, which will be denoted by $Y \sim \mathrm{S}t(\alpha, \gamma, \sigma, \nu)$, with pdf and cdf, respectively, given by

$$f_Y(y) = \phi_t\left(\frac{2\sinh\left(\frac{y-\gamma}{\sigma}\right)}{\alpha}\right)\frac{2\cosh\left(\frac{y-\gamma}{\sigma}\right)}{\sigma\alpha} = \frac{2\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\frac{\cosh\left(\frac{y-\gamma}{\sigma}\right)}{\sigma\alpha} \times\left[1+\frac{4\sinh^2\left(\frac{y-\gamma}{\sigma}\right)}{\nu\alpha^2}\right]^{-\frac{(\nu+1)}{2}} \tag{1}$$

and

$$F_Y(y) = \Phi_t\left(\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right) = \frac{1}{2}\left[1+\mathrm{I}_{\frac{4\sinh^2\left(\frac{y-\gamma}{\sigma}\right)}{4\sinh^2\left(\frac{y-\gamma}{\sigma}\right)+\nu\alpha^2}}\left(\frac{1}{2},\frac{\nu}{2}\right)\right];\quad y\in\mathbb{R}, \tag{2}$$

where $\mathrm{I}_x(a,b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1}\,dt}{\int_0^1 t^{a-1}(1-t)^{b-1}\,dt}$ is the incomplete beta ratio function; see Johnson et al. (1995, p. 364) and $\phi_t(\cdot)$ and $\Phi_t(\cdot)$ denoting the pdf and cdf of the t distribution. The survival (s.f.) and hazard (h.f.) functions of Y are, respectively, given by

$$S_Y(y) = \Phi_t\left(-\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right) \text{ and}$$
$$h_Y(y) = \frac{2\phi_t\left(\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right)\cosh\left(\frac{y-\gamma}{\sigma}\right)}{\Phi_t\left(-\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right)\sigma\alpha};\quad y\in\mathbb{R}. \tag{3}$$

Analogously to the case of the SN model, the St distribution has as a particular case the log-BS-t distribution when $\gamma = \log(\beta)$ and $\sigma = 2$, which will be denoted by $Y \sim$ log-GBS$(\alpha, \gamma; t_\nu)$, so that $T = \exp(Y) \sim$ GBS$(\alpha, \beta; t_\nu)$.

Next, we present a brief graphical analysis for the log-BS-t model. In Fig. 1, pdf plots are shown for a number of different choices of α and ν, being $\gamma = \log(\beta) = 0$ and $\sigma = 2$. Based on Fig. 1, we note that the log-BS-t distribution is very flexible for modeling the kurtosis. Figure 1 also establishes comparisons of the log-BS-t distribution and the log-BS, normal and t distributions. The parameters α and γ are the shape and location (mean) parameters, respectively, whereas the parameter ν of the t distribution makes it possible to handle the kurtosis. In addition, α is also related to the modality.

Consider the regression model given by

$$y_i = \boldsymbol{x}_i^\top\boldsymbol{\beta} + \varepsilon_i;\quad i = 1,\ldots,n, \tag{4}$$

where y_i is the observed log-lifetime or log-censoring time for the ith individual, $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^\top$ is a vector of unknown parameters to be estimated, $\boldsymbol{x}_i^\top = (x_{i1},\ldots,x_{ip})$

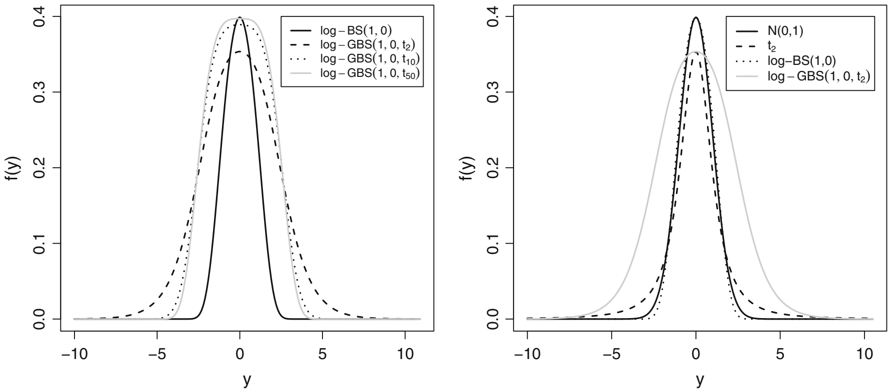

Fig. 1 Pdf graphs of log-BS, log-BS-t, normal and t models for the indicated values

contains the values of the explanatory variables and $\varepsilon_i \sim$ log-GBS$(\alpha, 0; t_\nu)$. We assume no informative censoring and independence of the observed lifetime and censoring time. We denote the sets of individuals for which y_i is the log-lifetime or log-censoring by D and C, respectively. The log-likelihood function of the model given in Eq. 4 for the parameter $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$ takes the form $l(\boldsymbol{\theta}) \propto \sum_{i\in \mathrm{D}} l_i(\boldsymbol{\theta}) + \sum_{i\in \mathrm{C}} l_i^{(c)}(\boldsymbol{\theta})$, where $l_i(\boldsymbol{\theta}) = \log(f_Y(y_i))$ and $l_i^{(c)}(\boldsymbol{\theta}) = \log(S_Y(y_i))$, with $f_Y(\cdot)$ and $S_Y(\cdot)$ being the pdf and s.f. of Y given in Eqs. 1 and 3, respectively. Thus, the log-likelihood function for $\boldsymbol{\theta}$ is

$$l(\boldsymbol{\theta}) \propto \sum_{i\in \mathrm{D}} \left[\log(\xi_{i1}) - \left\{\frac{\nu+1}{2}\right\} \log(\nu + \xi_{i2}^2)\right] + \sum_{i\in \mathrm{C}} \log\left(\Phi_t(-\xi_{i2})\right), \tag{5}$$

where $\xi_{i1} = \frac{2}{\alpha}\cosh\left(\frac{y_i-\mu_i}{2}\right)$, $\xi_{i2} = \frac{2}{\alpha}\sinh\left(\frac{y_i-\mu_i}{2}\right)$ and $\mu_i = \boldsymbol{x}_i^\top\boldsymbol{\beta}$, with $i = 1, \ldots, n$.

The score vector is $\dot{\mathbf{L}}_\theta = \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_1}, \ldots, \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_p}, \frac{\partial l(\boldsymbol{\theta})}{\partial \alpha}\right)^\top = (U_{\beta_1}, \ldots, U_{\beta_p}, U_\alpha)^\top$, where

$$U_{\beta_j} = \sum_{i\in \mathrm{D}} \left[\frac{x_{ij}}{\alpha^2}\sinh(y_i-\mu_i)w(\xi_{i2}^2) - \frac{x_{ij}}{2}\tanh\left(\frac{y_i-\mu_i}{2}\right)\right] + \frac{1}{2}\sum_{i\in \mathrm{C}} x_{ij}\xi_{i1}h(\xi_{i2}), \tag{6}$$

for $j = 1, \ldots, p$, and

$$U_\alpha = \frac{1}{\alpha}\left[\sum_{i\in \mathrm{D}}\left\{w(\xi_{i2}^2)\xi_{i2}^2 - 1\right\} + \sum_{i\in \mathrm{C}} \xi_{i2}h(\xi_{i2})\right], \tag{7}$$

with $w(\xi_{i2}^2) = \frac{\nu+1}{(\nu+\xi_{i2}^2)}$ and $h(\cdot) = h_Y(\cdot)$ being the h.f. of Y given in Eq. 3. Note that as $\nu \to \infty$, one has $w(\xi_{i2}^2) \to 1$, $\forall\, i = 1, \ldots, n$, and the scores given in Eqs. 6 and 7 are

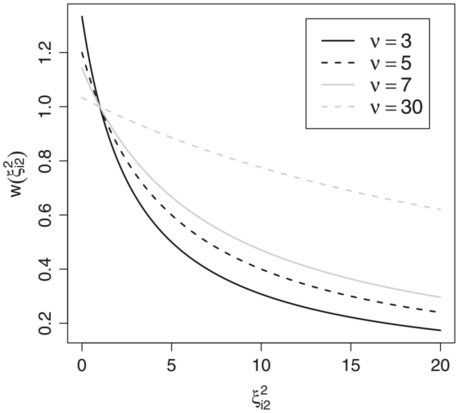

Fig. 2 Behavior of $w(\xi_{i2}^2)$ against ξ_{i2}^2 for some values of the d.f. of the t distribution

reduced to the ones from the log-BS regression, as expected; see Galea et al. (2004) and Leiva et al. (2007). Thus, the quantity $w(\xi_{i2}^2)$ that appears for non-censored cases may be interpreted as a kind of weight in the log-BS-t model and since this is inversely proportional to ξ_{i2}^2, cases with larger values for ξ_{i2}^2 should have smaller weights (less than one). Figure 2 shows graphs of $w(\xi_{i2}^2)$ against ξ_{i2}^2 for some values of ν.

Remark As it is well known, the Student-t distribution can be expressed as a scale mixture of normal distributions, which facilitates for instance Bayesian analysis for the model or maximum likelihood estimation via the EM algorithm. Similarly, we can assume that Z has a scale mixture of normal distribution so that T will follow a GBS distribution based on scale mixture of normals. In this case, maximum likelihood equations given as in Eqs. 6 and 7 can be obtained as well as applications of the EM algorithm and Bayesian analysis can be performed.

Notice that the weights decrease as ν becomes smaller. These results indicate the robustness of the MLE $\hat{\alpha}$ and $\hat{\boldsymbol{\beta}}$ against extreme non-censored cases in the sense of the quantity ξ_{i2}. These robustness aspects are extended in Sect. 3 to d-c and m-g residuals usually applied in lifetime regression models.

The observed Fisher information matrix is obtained by $-\ddot{\mathbf{L}}^{-1}$ evaluated at $\hat{\theta}$, where

$$\ddot{\mathbf{L}} = \begin{bmatrix} \ddot{\mathbf{L}}_{\beta\beta} & \ddot{\mathbf{L}}_{\beta\alpha} \\ \ddot{\mathbf{L}}_{\alpha\beta} & \ddot{\mathbf{L}}_{\alpha\alpha} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top} l(\boldsymbol{\theta}) & \frac{\partial^2}{\partial\boldsymbol{\beta}\partial\alpha} l(\boldsymbol{\theta}) \\ \frac{\partial^2}{\partial\alpha\partial\boldsymbol{\beta}} l(\boldsymbol{\theta}) & \frac{\partial^2}{\partial\alpha^2} l(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^\top \boldsymbol{V} \boldsymbol{X} & \boldsymbol{X}^\top \boldsymbol{K} \\ \boldsymbol{K}^\top \boldsymbol{X} & \mathrm{tr}(\boldsymbol{G}) \end{bmatrix}, \quad (8)$$

with $\boldsymbol{V} = \mathrm{diag}\{v_1(\boldsymbol{\theta}), \ldots, v_n(\boldsymbol{\theta})\}$, $\boldsymbol{K} = (k_1(\boldsymbol{\theta}), \ldots, k_n(\boldsymbol{\theta}))^\top$, $\boldsymbol{G} = \mathrm{diag}\{g_1(\boldsymbol{\theta}), \ldots, g_n(\boldsymbol{\theta})\}$,

$$v_i(\boldsymbol{\theta}) = \begin{cases} \frac{1}{4}\mathrm{sech}^2\left(\frac{y_i - \mu_i}{2}\right) - \frac{1}{\alpha^2}\cosh(y_i - \mu_i)\, w(\xi_{i2}^2) \\ \quad - \frac{2}{\alpha^4}\sinh^2(y_i - \mu_i)\, w'(\xi_{i2}^2), & i \in \mathrm{D}; \\ -\frac{1}{4}\xi_{i2}\, h(\xi_{i2}) - \frac{1}{4}\xi_{i1}^2\, h'(\xi_{i2}), & i \in \mathrm{C}; \end{cases}$$

$$k_i(\boldsymbol{\theta}) = \begin{cases} -\frac{2}{\alpha^3}\sinh(y_i - \mu_i)\left\{w(\xi_{i2}^2) + \xi_{i2}^2 w'(\xi_{i2}^2)\right\}, & i \in \mathrm{D}; \\ -\frac{1}{2\alpha}\xi_{i1}h(\xi_{i2}) - \frac{1}{\alpha^3}\sinh(y_i - \mu_i)h'(\xi_{i2}), & i \in \mathrm{C}; \end{cases}$$

and

$$g_i(\boldsymbol{\theta}) = \begin{cases} \frac{1}{\alpha^2} - \frac{3\xi_{i2}^2}{\alpha^2}\, w(\xi_{i2}^2) - \frac{2}{\alpha^2}\xi_{i2}^4\, w'(\xi_{i2}^2), & i \in \mathrm{D}; \\ -\frac{2}{\alpha^2}\xi_{i2}\, h(\xi_{i2}) - \frac{1}{\alpha^2}\xi_{i2}^2\, h'(\xi_{i2}), & i \in \mathrm{C}; \end{cases}$$

where $w'(x) = \frac{d}{dx}w(x)$ and $h'(x) = \frac{d}{dx}h(x)$.

The MLEs of the regression coefficients and shape parameter are the solutions of the likelihood equations $U_{\beta_j} = 0$, with $j = 1, \ldots, p$, and $U_\alpha = 0$. However, as in the log-BS case, the equations do not present analytical solutions necessitating the use of iterative methods (Leiva et al. 2007). Asymptotic inference for the parameter $\boldsymbol{\theta}$ can be based on the normal approximation of the MLEs given by $\hat{\boldsymbol{\theta}} \sim \mathrm{N}_{p+1}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\hat{\theta}})$, where $\boldsymbol{\Sigma}_{\hat{\theta}}$ is the variance-covariance matrix of $\hat{\boldsymbol{\theta}}$, which can be approximated by $-\ddot{\mathbf{L}}^{-1}$, with $-\ddot{\mathbf{L}}$ being the observed information matrix evaluated at $\hat{\boldsymbol{\theta}}$ and obtained from $\ddot{\mathbf{L}}$ given in Eq. 8.

In order to construct a confidence region for the parameter $\boldsymbol{\theta}$, we can use the fact that $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}_{\hat{\theta}}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \stackrel{\cdot}{\sim} \chi^2(p+1)$, which is obtained from the asymptotic normality of the MLEs. Therefore, an approximate $100(1-\gamma)\%$ confidence region, with $0 < \gamma < 1$, for $\boldsymbol{\theta}$ is given by $\mathcal{R} = \{\boldsymbol{\theta} \in \mathbb{R}^{p+1} : (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}_{\hat{\theta}}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \le \chi^2_{1-\gamma}(p+1)\}$, where $\chi^2_{1-\gamma}(p+1)$ denotes the $(1-\gamma)$th percentile of the chi-square distribution with $p+1$ d.f.

3 Model checking

In order to assess the robustness aspects of the MLEs, an analysis of normal curvatures of local influence (Cook 1986) may be carried out for some common perturbation schemes. If the likelihood displacement $\mathrm{LD}(\boldsymbol{\omega}) = 2[l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_\omega)]$ is used as the influence measure, where $\hat{\boldsymbol{\theta}}_\omega$ denotes the MLE of $\boldsymbol{\theta}$ under the perturbed model $l(\boldsymbol{\theta}|\boldsymbol{\omega})$, the normal curvature at the unitary direction $\boldsymbol{\ell}$ assumes the form $\mathrm{C}_\ell(\boldsymbol{\theta}) = 2|\boldsymbol{\ell}^\top \boldsymbol{\Delta}^\top \ddot{\mathbf{L}}^{-1} \boldsymbol{\Delta}\boldsymbol{\ell}|$, which is evaluated at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$ (the no perturbation vector). Here, $\boldsymbol{\Delta}$ is a $(p+1) \times n$ perturbation matrix with elements $\Delta_{ji} = \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top} l(\boldsymbol{\theta}|\boldsymbol{\omega})$, for $j = 1, \ldots, p+1$ and $i = 1, \ldots, n$. The elements of the matrix $\boldsymbol{\Delta}$ are derived in the Appendix for some common perturbation schemes. Diagnostics graphs may be constructed from the normal curvature $\mathrm{C}_i(\boldsymbol{\theta}) = \mathrm{C}_{\ell_i}(\boldsymbol{\theta})$ (Lesaffre and Verbeke 1998), where $\boldsymbol{\ell}_i$ is an $n \times 1$ vector of zeros with one at the ith position, or by considering the direction $\boldsymbol{\ell}_{\max}$ corresponding to the largest curvature $\mathrm{C}_{\ell_{\max}}(\boldsymbol{\theta})$ (Cook 1986; Poon and Poon 1999). Previous works in which local influence curvatures are derived for

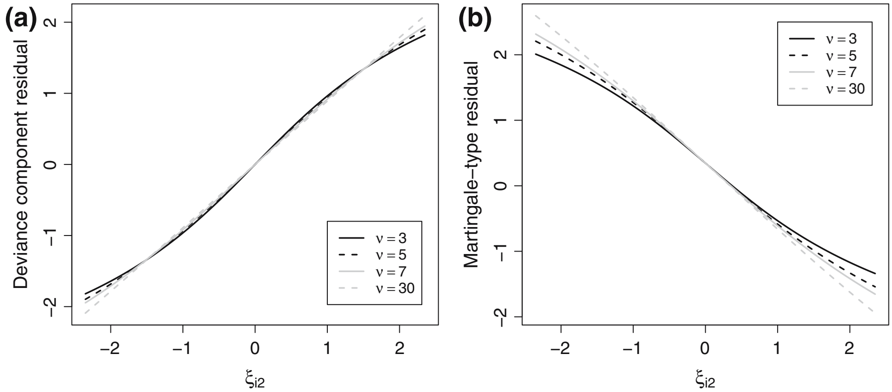

Fig. 3 Behavior of residuals r_{DC_i} (**a**) and r_{MD_i} (**b**) again sinh residual for non-censored cases

regressing modeling with censored data are attributed to Escobar and Meeker (1992), Ortega et al. (2003) and Leiva et al. (2007).

For a residual analysis, we suggest working with the d-c residual (Davison and Gigli 1989) given by

$$r_{DC_i} = \begin{cases} \operatorname{sign}(\hat{\xi}_{i2})\sqrt{2}\left[\frac{\nu+1}{2}\log(\nu+\hat{\xi}_{i2}^2) - \frac{\nu+1}{2}\log(\nu) \right. & \\ \left. -\frac{1}{2}\log\left(1+\frac{\hat{\alpha}^2\hat{\xi}_{i2}^2}{4}\right)\right]^{\frac{1}{2}}, & i \in \mathrm{D}; \\ \operatorname{sign}(\hat{\xi}_{i2})\left[-2\log\left(\Phi_t(-\hat{\xi}_{i2})\right)\right]^{\frac{1}{2}}, & i \in \mathrm{C}; \end{cases}$$

and the martingale-type residual defined as $r_{M_i} = \delta_i + \log(\Phi_t(-\xi_{i2}))$ (Klein and Moeschberger 1997), with $\delta_i = 0$ and 1 indicating that the observation is censored or uncensored, respectively. However, since r_{M_i} is skewness, with range between $-\infty$ and $+1$, some authors (e.g., Collett 2003, p. 116) have proposed the following transformation to attenuate the skewness: $r_{MD_i} = \operatorname{sign}(r_{M_i})[-2\{r_{M_i} + \delta_i \log(\delta_i - r_{M_i})\}]^{\frac{1}{2}}$. This transformation leads to the d-c residual for the Cox's proportional hazard model with no time-dependent variable (Therneau et al. 1990). Leiva et al. (2007) have investigated the empirical distributions of r_{DC_i} and r_{MD_i} for log-BS regression models by using different sample sizes and censoring proportions. They found a very good agreement with the standard normal distribution as the sample size increases and the censoring proportion decreases. Figure 3 describes the behavior of r_{MD_i} and r_{DC_i} against the sinh residual ξ_{i2} for non-censored observations, for some values of ν and $\alpha = 1$. We notice in both graphs that non-censored observations with large values for r_{DC_i} and r_{MD_i} will have smaller weights. Thus, the robustness aspects that were pointed out in Sect. 2 are also extend to the residuals r_{DC_i} and r_{MD_i}.

4 Application

We consider the data set presented in Kalbfleisch and Prentice (2002, p. 378) as a motivating example, where male patients with advanced inoperable lung cancer were randomized to either standard or test chemotherapy. One of the objectives of this study was to explain the survival time (T, in days) of 137 patients with lung cancer (9 of which were censored) by using a regression model and considering the following explanatory variables: a measure, at randomization, of the patient's performance status—Karnofsky rating– (x_1), where 10–30 is completely hospitalized, 40-60 is partial confinement and 70–90 is able to take care of self; time in months from diagnosis to randomization (x_2); age in years (x_3); prior therapy (x_4), a dichotomous variable taking the value 10 for yes and 0 for no; histological type of tumor, which has the categories squamous, small cell, adeno and large cell, making necessary the use of dummy variables given by $x_5 = 1$, $x_6 = 1$ and $x_7 = 1$ if the type of cancer cell is squamous, small and adeno, respectively, and 0 otherwise; and type of treatment (x_8), which takes the value 0 for standard chemotherapy and 1 for test chemotherapy.

The data considered here has been analyzed by means of regression models assuming different error distributions, such as the exponential, generalized gamma, log-logistic, lognormal and Weibull (Lee and Wang 2003). In spite of this, an important aspect that should be considered is the criterion through which the lifetime distribution is justified. Thus, a theoretical argument is often used for describing the mechanism of death or failure. For example, the following arguments can be considered: "wear-out of phase-type", with exponential lifetimes for each phase, for the gamma distribution; "first passage time" for the inverse Gaussian distribution; "multiplicative degradation" for the lognormal distribution; and "type extreme values" for the Weibull distribution. An argument for the use of the BS distribution is the possibility of relating the propagation lifetimes that lead to a fatigue process with some cumulative damage. Thus, based on this argument, we propose the BS model to analyze the data for this motivating example.

Different reasons justify the relationship between a distribution and its associated logarithmic distribution (log-distribution). For instance, in lifetime regression analysis, the lifetime and its associated covariates are generally log-linearly related. Thus, if the lifetime regression model is linearized, then a log-distribution for the random error is required. This model is called an accelerated failure time regression model; see Lawless (2002, Chapter 6) and Meeker and Escobar (1998, Chapter 18). In addition, the parameter estimates, moments and random numbers of a distribution could be obtained more efficiently from its log-distribution. For these reasons, it is necessary to know the log-distribution associated with a distribution.

Based on the theoretical argument given above, we initially fit the data by assuming the log-BS model for censored data that was proposed in Leiva et al. (2007).

4.1 Analysis under the log-BS model

Firstly, the model considered is given by

$$\log(t_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_8 x_{i8} + \varepsilon_i, \quad i = 1, \ldots, 137, \tag{9}$$

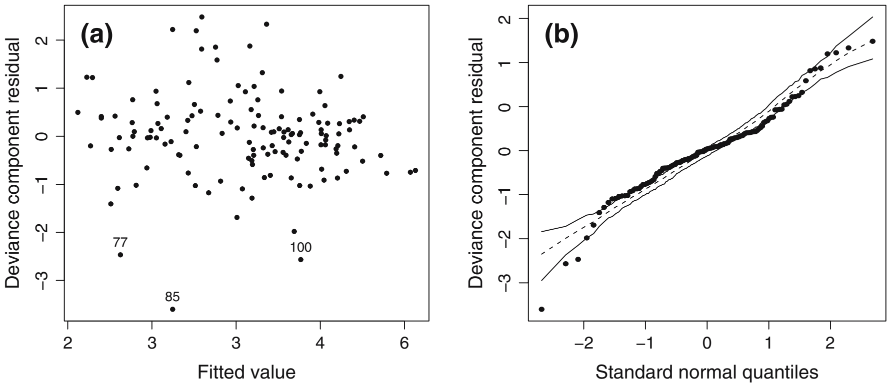

Fig. 4 d-c residual again the fitted value (**a**) and normal probability plot with envelope (**b**)

assuming that the ε_i's are independent and identically distributed (i.i.d.) variates, where $\varepsilon_i \sim$ log-BS$(\alpha, 0)$. The response $y_i = \log(t_i)$ denotes the logarithm of the survival or censoring time. Since $\mathbb{E}[\varepsilon_i] = 0$, the regression is performed on $\mathbb{E}[\log(T_i)]$. The values of the MLEs (approximate standard error in parenthesis) are: $\hat{\beta}_0 = 1.142$ (0.638), $\hat{\beta}_1 = 0.040$ (0.005), $\hat{\beta}_2 = -0.003$ (0.009), $\hat{\beta}_3 = 0.022$ (0.008), $\hat{\beta}_4 = -0.002$ (0.023), $\hat{\beta}_5 = -0.280$ (0.303), $\hat{\beta}_6 = -0.705$ (0.301), $\hat{\beta}_7 = -0.691$ (0.367), $\hat{\beta}_8 = -0.383$ (0.193) and $\hat{\alpha} = 1.262$ (0.079). In this case, the predictors x_2, x_4, x_5, x_7 and x_8 are not marginally significant at 5%.

A diagnostics analysis based on the d-c residual (Leiva et al. 2007) highlights strongly the observation #85; see Fig. 4a. This case corresponds to a 35-year-old patient whose survival time was one day and who had waited for 7 months until randomization. His performance was partial confinement, he did not have any prior therapy, the histological tumor type was squamous and he received the test chemotherapy treatment. The normal probability plot with generated envelope for the d-c residual (see Fig. 4b) presents some indication that a heavy-tailed error distribution might be more appropriate. Furthermore, three observations fall outside the envelope, which correspond to patients #77, #85 and #100. These are pointed out in Fig. 4a. Graphical analyses of total local influence under the case-weight perturbation scheme are displayed in Fig. 5. Also, by using the cut-off point proposed in Verbeke and Molenberghs (2000, Subsect. 11.3), the three observations mentioned above (cases #77, #85 and #100) appear to have an outstanding influence.

Table 1 shows the relative changes (RC) in the estimates after dropping one of the three cases with outstanding influence and also when all of them are dropped at once (represented by the set I $= \{77, 85, 100\}$). The p-values for the new estimates are given in parentheses in Table 1. The RC (in percentage) of each estimated parameter are defined by: $\mathrm{RC}_{\theta_j} = \left|[\hat{\theta}_j - \hat{\theta}_{j(\mathrm{I})}]/\hat{\theta}_j\right| \times 100\%$, where $\hat{\theta}_{j(\mathrm{I})}$ denotes the MLE of θ_j, with $j = 1, \ldots, 10$, after the set I of observations has been removed. In general, the RC are large, even for those estimates that were significant at 5%. In addition, we notice that there are changes in the inference for some coefficients. Particularly, β_3 and β_8 are not significant at 5% after eliminating the three observations.

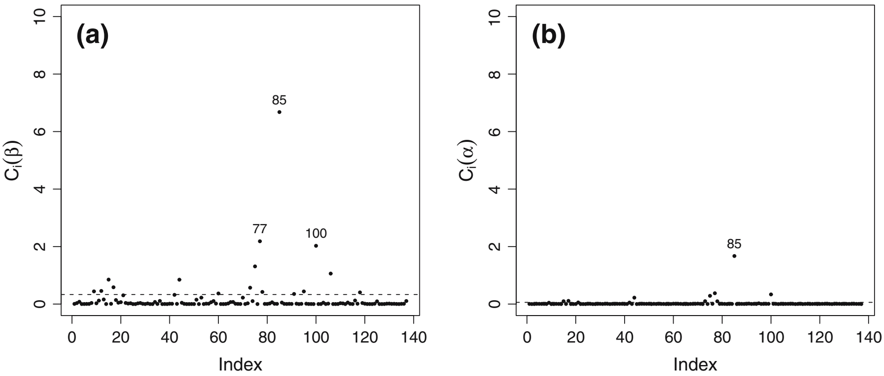

Fig. 5 Index plots of $C_i(\boldsymbol{\beta})$ **(a)** and $C_i(\alpha)$ **(b)**, under case-weight perturbation

Table 1 RC (in %) and the corresponding p-values in parentheses

Estimated coefficient	Dropped observation				
	None	77	85	100	Set I
$\hat{\beta}_0$	–	−3	−104	35	−81
	(0.073)	(0.053)	(0.000)	(0.243)	(0.000)
$\hat{\beta}_1$	–	11	8	−8	14
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\beta}_2$	–	−37	−15	−40	−108
	(0.750)	(0.659)	(0.688)	(0.651)	(0.439)
$\hat{\beta}_3$	–	−16	77	−14	51
	(0.006)	(0.001)	(0.549)	(0.001)	(0.158)
$\hat{\beta}_4$	–	556	−1261	−155	−756
	(0.927)	(0.673)	(0.200)	(0.810)	(0.382)
$\hat{\beta}_5$	–	28	77	−2	119
	(0.356)	(0.489)	(0.817)	(0.361)	(0.833)
$\hat{\beta}_6$	–	−4	−2	20	11
	(0.019)	(0.011)	(0.008)	(0.053)	(0.011)
$\hat{\beta}_7$	–	−9	−17	−4	−28
	(0.060)	(0.039)	(0.014)	(0.042)	(0.002)
$\hat{\beta}_8$	–	18	48	36	101
	(0.047)	(0.092)	(0.268)	(0.210)	(0.981)
$\hat{\alpha}$	–	3	9	4	17

Although we can choose a final model like the one mentioned above, clearly the MLEs present lack of robustness when the outlying observations are considered in the data and the assumption for the distribution of the error term appears to be unsuitable. We will reanalyze these data in the sequel using a heavy-tailed error model.

4.2 Analysis under the log-BS-t model

The data set analyzed in Subsect. 4.1 under the log-BS regression model is reanalyzed here under the log-BS-t regression model, that is, under the model given in Eq. 9

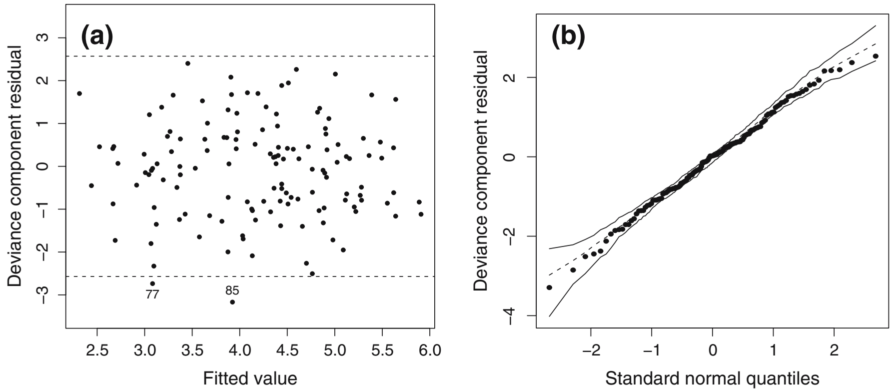

Fig. 6 d-c residual against the fitted value (**a**) and normal probability plot with envelope (**b**)

assuming that the ε_i's are i.i.d. r.v., where $\varepsilon_i \sim \text{log-GBS}(\alpha, 0; t_\nu)$. An important point to consider under t models is related to the estimation of the d.f. ν. Several authors have dealt with this topic (Lange et al. 1989; Berkane et al. 1994; Fernandez and Steel 1999; Taylor and Verbyla 2004; Leiva et al. 2008b) and pointed out difficulties in estimating ν due to problems of unbounded and local maximum in the likelihood function. This parameter can be fixed previously as recommended by Lange et al. (1989) and Berkane et al. (1994). They suggest considering $\nu = 4$ or otherwise to get information for ν from the data set.

In order to estimate $\boldsymbol{\beta}$ and α of the log-BS-t regression model, we fix different values for ν and consider the MLEs of $\boldsymbol{\beta}$ and α in the log-BS model as starting values for the numerical procedure. We choose the value of ν that maximizes the likelihood function over several values of $\nu \in [2, 50]$, obtaining $\nu = 3$. The values of the MLEs correspond to: $\hat{\beta}_0 = 2.065(0.636)$, $\hat{\beta}_1 = 0.036(0.004)$, $\hat{\beta}_2 = 0.004(0.010)$, $\hat{\beta}_3 = 0.008(0.009)$, $\hat{\beta}_4 = -0.010(0.021)$, $\hat{\beta}_5 = -0.004(0.267)$, $\hat{\beta}_6 = -0.723(0.242)$, $\hat{\beta}_7 = -0.746$ (0.256), $\hat{\beta}_8 = -0.076(0.176)$ and $\hat{\alpha} = 0.816(0.073)$. From this analysis, the explanatory variables x_2, x_3, x_4, x_5 and x_8 are not marginally significant at 5%. However, the patient's performance status and histological tumor type were statistically significant when the log-BS-t model is taken into consideration.

By performing the diagnostics analysis, we created the graphs of the d-c residuals against the fitted values and the normal probability plot with generated envelope in Fig. 6a and b, respectively. In Fig. 6a, we notice some large negative residuals (patients #77 and #85), but from Fig. 6b the assumption of log-BS-t error seems to be suitable, since there are no observations falling outside the envelope.

Index plots of $|\boldsymbol{\ell}_{\max}|$ for $C_\ell(\boldsymbol{\beta})$ and $C_\ell(\alpha)$ under case-weight perturbation (which are not shown here) indicate that observations #12, #95 and #106 are revealed as potentially influential. The patients that appear in these two graphs have large values in common for the time from the diagnosis to randomization. However, in Fig. 7, where the index plots of $C_i(\boldsymbol{\beta})$ –7(a)– and $C_i(\alpha)$ –7(b)– under the case-weight perturbation scheme are displayed, few observations appear as potentially influential.

Similar to the analysis performed in Subsect. 4.1, the four observations pointed out in Fig. 7a and b (cases #12, #77, #95 and #106) were dropped. Then, the relative

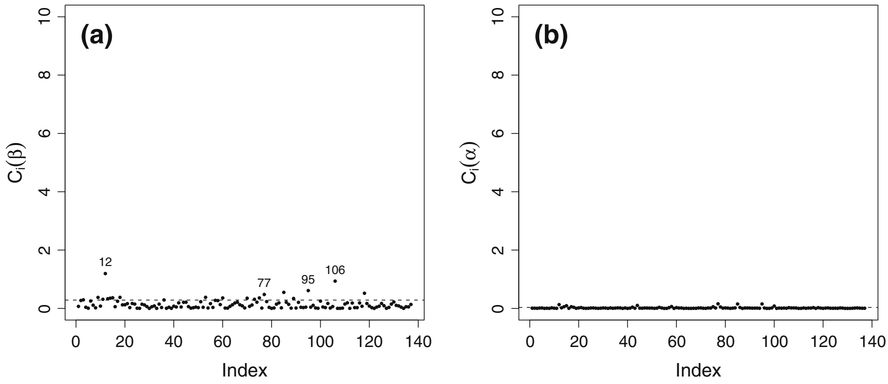

Fig. 7 Index plots of $C_i(\boldsymbol{\beta})$ (**a**) and $C_i(\alpha)$ (**b**), under case-weight perturbation

Table 2 RC (in %) and the corresponding p-values in parentheses

Estimated coefficient	Dropped observation					
	None	12	77	95	106	Set I
$\hat{\beta}_0$	–	3	−1	−4	−8	−5
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\hat{\beta}_1$	–	1	3	0	1	4
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\beta}_2$	–	−171	−23	−89	471	−113
	(0.730)	(0.274)	(0.657)	(0.463)	(0.318)	(0.569)
$\hat{\beta}_3$	–	−10	−3	25	2	13
	(0.372)	(0.318)	(0.353)	(0.497)	(0.384)	(0.422)
$\hat{\beta}_4$	–	−7	31	23	42	60
	(0.634)	(0.604)	(0.740)	(0.710)	(0.787)	(0.843)
$\hat{\beta}_5$	–	1018	1216	87	1146	2642
	(0.988)	(0.882)	(0.857)	(0.998)	(0.868)	(0.681)
$\hat{\beta}_6$	–	−2	−1	6	−7	1
	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)
$\hat{\beta}_7$	–	2	0	1	−3	3
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
$\hat{\beta}_8$	–	−46	21	46	−47	6
	(0.669)	(0.527)	(0.732)	(0.813)	(0.533)	(0.677)
$\hat{\alpha}$	–	2	2	2	0	6

changes after the set I of observations was removed were computed as well as the new p-values (see Table 2). Large variations are only noticed for the non-significant parameters and α, but inferential changes are not observed. Figure 8 displays the estimated weight $\hat{\xi}_{i2}$ against the d-c residual r_{DC_i} and, as was expected, larger residuals have smaller weights, where some of them were detected in the first analysis given in Subsect. 4.1.

Based on the analysis here, we conclude that the log-BS-t regression model is more appropriate for fitting this data set than the log-BS model. Thus, the final selected model in our analysis is

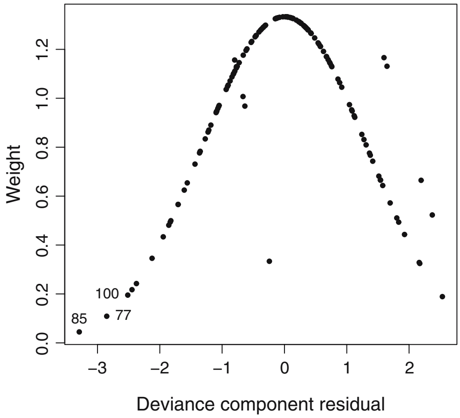

Fig. 8 Estimated weight from the log-BS-t regression model against the d-c residual

$$\log(t_i) = \beta_0 + \beta_1 x_{i1} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \varepsilon_i, \quad i = 1, \ldots, 137, \qquad (10)$$

where the values of the MLEs (approximate standard error in parenthesis) are: $\hat{\beta}_0 = 2.462\,(0.336)$, $\hat{\beta}_1 = 0.036\,(0.004)$, $\hat{\beta}_5 = 0.048\,(0.260)$, $\hat{\beta}_6 = -0.664\,(0.234)$, $\hat{\beta}_7 = -0.738\ (0.253)$, and $\hat{\alpha} = 0.817\ (0.073)$, which may be interpreted in the following manner. The survival time is expected to increase with the performance status and no significance appears between the squamous and large types; however, the expected survival time should decrease 94% and 109% for the small and adeno tumor types with respect to the large one, respectively, assuming that the performance status is fixed.

5 Concluding remarks

In this paper, we proposed a new class of lifetime regression models called log-BS-t models, for which the maximum likelihood estimates appear to be robust against outlying observations in the sense of two common residuals. In order to study the robustness aspects of the maximum likelihood estimates against perturbations in the model/data, the local influence curvatures are derived under various perturbation schemes. Finally, a motivating example previously analyzed by considering other distributions, such as the exponential, generalized gamma, log-logistic, lognormal and Weibull distributions, was reanalyzed based on the BS and BS-t models. The medical justification allowed the use of the BS distribution, but this model, as well as those previously analyzed, was shown to be sensitive to extreme data. We showed that the BS-t model seems to be more appropriate for fitting the data set than the BS model.

Appendix: matrix $\boldsymbol{\Delta}$ calculations

Here, we compute the elements of the matrix $\boldsymbol{\Delta}$ for each perturbation scheme considering the model given in Eq. 4. This is partitioned as $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta, \boldsymbol{\Delta}_\alpha)^\top$, where $\boldsymbol{\Delta}_\beta$ is a $p \times n$ matrix and $\boldsymbol{\Delta}_\alpha$ is a $n \times 1$ vector.

[*A1*] *Case-weight perturbation.* Here $\boldsymbol{\Delta}_\beta = \boldsymbol{X}^\top \mathrm{diag}\{\hat{b}_1, \ldots, \hat{b}_n\}$ and $\boldsymbol{\Delta}_\alpha = (\hat{a}_1, \ldots, \hat{a}_n)$, where

$$\hat{b}_i = \begin{cases} \frac{1}{2}\left[\hat{\xi}_{i1}\hat{\xi}_{i2}\, w(\hat{\xi}_{i2}^2) - \frac{\hat{\xi}_{i2}}{\hat{\xi}_{i1}}\right], & i \in \mathrm{D}; \\ \frac{\hat{\xi}_{i1}}{2}\, h(\hat{\xi}_{i2}), & i \in \mathrm{C}; \end{cases} \quad \text{and} \quad \hat{a}_i = \begin{cases} \frac{1}{\hat{\alpha}}\left[\hat{\xi}_{i2}^2\, w(\hat{\xi}_{i2}^2) - 1\right], & i \in \mathrm{D}; \\ \frac{\hat{\xi}_{i2}}{\hat{\alpha}}\, h(\hat{\xi}_{i2}), & i \in \mathrm{C}. \end{cases}$$

[*A2*] *Response perturbation.* In this case, we have $\boldsymbol{\Delta}_\beta = \boldsymbol{X}^\top \mathrm{diag}\{\hat{d}_1, \ldots, \hat{d}_n\}$, where

$$\hat{d}_i = \begin{cases} S_y\left[\frac{1}{\hat{\alpha}^2}\cosh(y_i - \hat{\mu}_i)\, w(\hat{\xi}_{i2}^2) \right. \\ \quad \left. -\frac{1}{4}\mathrm{sech}^2\left(\frac{y_i - \hat{\mu}_i}{2}\right) - \frac{2}{\hat{\alpha}^4}\sinh(y_i - \hat{\mu}_i)\, w'(\hat{\xi}_{i2}^2)\right], & i \in \mathrm{D}; \\ \frac{S_y}{4}\left[\hat{\xi}_{i2}\, h(\hat{\xi}_{i2}) + \hat{\xi}_{i1}^2 h'(\hat{\xi}_{i2})\right], & i \in \mathrm{C}; \end{cases}$$

and $\boldsymbol{\Delta}_\alpha = (\hat{c}_1, \ldots, \hat{c}_n)$, where

$$\hat{c}_i = \begin{cases} \frac{S_y}{\hat{\alpha}}\left[\hat{\xi}_{i1}\hat{\xi}_{i2}\, w(\hat{\xi}_{i2}^2) + \hat{\xi}_{i1}\hat{\xi}_{i2}^3\, w'(\hat{\xi}_{i2}^2)\right], & i \in \mathrm{D}; \\ \frac{S_y}{2\hat{\alpha}}\left[\hat{\xi}_{i1}\, h(\hat{\xi}_{i2}) + \hat{\xi}_{i1}\hat{\xi}_{i2} h'(\hat{\xi}_{i2})\right], & i \in \mathrm{C}. \end{cases}$$

[A3] Explanatory variable perturbation. Here $\boldsymbol{\Delta}_\beta$ is formed by the elements $\boldsymbol{\Delta}_{\beta_{ij}}$, which assume for $j \neq t$ the form

$$\Delta_{\beta_{ij}} = \begin{cases} S_x\hat{\beta}_t x_{ij}\left[\frac{1}{4}\mathrm{sech}^2(\frac{y_i - \hat{\mu}_i}{2}) - \frac{1}{\hat{\alpha}^2}\cosh(y_i - \hat{\mu}_i)\, w(\hat{\xi}_{i2}^2) \right. \\ \quad \left. -\frac{2}{\hat{\alpha}^2}\sinh^2(y_i - \hat{\mu}_i)\, w'(\hat{\xi}_{i2}^2)\right], & i \in \mathrm{D}; \\ -\frac{S_x\hat{\beta}_t x_{ij}}{4}\left[\hat{\xi}_{i2}\, h(\hat{\xi}_{i2}) + (\hat{\xi}_{i1})^2\, h'(\hat{\xi}_{i2})\right], & i \in \mathrm{C}; \end{cases}$$

and for $j = t$ the form

$$\Delta_{\beta_{it}} = \begin{cases} S_x\hat{\beta}_t x_{it}\left[\frac{1}{4}\mathrm{sech}^2(\frac{y_i - \hat{\mu}_i}{2}) - \frac{1}{\hat{\alpha}^2}\cosh(y_i - \hat{\mu}_i)\, w(\hat{\xi}_{i2}^2) \right. \\ \quad \left. -\frac{2}{\hat{\alpha}^2}\sinh^2(y_i - \hat{\mu}_i)\, w'(\hat{\xi}_{i2}^2)\right] \\ \quad -S_x\left[\frac{1}{\hat{\alpha}^2}\sinh(y_i - \hat{\mu}_i)\, w(\hat{\xi}_{i2}^2) - \frac{1}{2}\tanh(\frac{y_i - \hat{\mu}_i}{2})\right], & i \in \mathrm{D}; \\ -\frac{S_x\hat{\beta}_t x_{it}}{4}\left[\hat{\xi}_{i2}\, h(\hat{\xi}_{i2}) + (\hat{\xi}_{i1})^2 h'(\hat{\xi}_{i2})\right] + \frac{S_x}{2}\hat{\xi}_{i1}\, h(\hat{\xi}_{i2}), & i \in \mathrm{C}; \end{cases}$$

and $\boldsymbol{\Delta}_\alpha = (\hat{\phi}_1, \ldots, \hat{\phi}_n)$, where

$$\hat{\phi}_i = \begin{cases} -\frac{2}{\hat{\alpha}^3} S_x\hat{\beta}_t \sinh(y_i - \hat{\mu}_i)\left[w(\hat{\xi}_{i2}^2) + \hat{\xi}_{i2}^2\, w'(\hat{\xi}_{i2}^2)\right], & i \in \mathrm{D}; \\ -\frac{\hat{\beta}_t S_x}{2\hat{\alpha}}\left[\hat{\xi}_{i1}\, h(\hat{\xi}_{i2}) + \frac{2}{\hat{\alpha}^2}\sinh(y_i - \hat{\mu}_i) h'(\hat{\xi}_{i2})\right], & i \in \mathrm{C}. \end{cases}$$

[A4] Non-censoring case-weight perturbation. In this case, the quantities $\hat{b}_i$ and $\hat{a}_i$ given in [A1] reduce to

$$\hat{b}_i = \begin{cases} \frac{1}{2}\left[\hat{\xi}_{i1}\hat{\xi}_{i2}\, w(\hat{\xi}_{i2}^2) - \frac{\hat{\xi}_{i2}}{\hat{\xi}_{i1}}\right], & i \in \text{D}; \\ 0, & i \in \text{C}; \end{cases} \quad \text{and} \quad \hat{a}_i = \begin{cases} \frac{1}{\hat{\alpha}}\left[\hat{\xi}_{i2}^2\, w(\hat{\xi}_{i2}^2) - 1\right], & i \in \text{D}; \\ 0, & i \in \text{C}. \end{cases}$$

[A5] Censoring case-weight perturbation. For this case, the quantities $\hat{b}_i$ and $\hat{a}_i$ given in [A1] reduce to

$$\hat{b}_i = \begin{cases} 0, & i \in \text{D}; \\ \frac{\hat{\xi}_{i1}}{2}\, h(\hat{\xi}_{i2}), & i \in \text{C}; \end{cases} \quad \text{and} \quad \hat{a}_i = \begin{cases} 0, & i \in \text{D}; \\ \frac{\hat{\xi}_{i2}}{\hat{\alpha}}\, h(\hat{\xi}_{i2}), & i \in \text{C}. \end{cases}$$

Acknowledgements The authors would like to thank the editor and referees for their helpful comments aided in improving this article. This study was supported by FONDECYT grant 1080326, Chile, and CNPq and FAPESP, Brazil.

References

Balakrishnan N, Leiva V, López J (2007) Acceptance sampling plans from truncated life tests from generalized Birnbaum-Saunders distribution. Commun Stat Simul Comput 36:643–656

Berkane M, Kano Y, Bentler PM (1994) Pseudo maximum likelihood estimation in elliptical theory: effects of misspecification. Comput Stat Data Anal 18:255–267

Birnbaum ZW, Saunders SC (1969) A new family of life distributions. J Appl Prob 6:319–327

Collett D (2003) Modelling survival data in medical research. Chapman and Hall, London

Cook RD (1986) Assessment of local influence. J Roy Stat Soc B 48:133–169

Cramér H (1999) Mathematical methods of statistics. Princeton University Press, Princeton

Davison AC, Gigli A (1989) Deviance residuals and normal scores plots. Biometrika 76:211–221

Desmond A (1985) Stochastic models of failure in random environments. Can J Stat 13:171–183

Díaz-García JA, Leiva V (2005) A new family of life distributions based on elliptically contoured distributions. J Stat Plan Infer 128:445–457 (Erratum: J Stat Plan Infer 137:1512–1513)

Dupuis DJ, Mills JE (1998) Robust estimation of the Birnbaum-Saunders distribution. IEEE Trans Rel 47:88–95

Escobar LA, Meeker WQ (1992) Assessing influence in regression analysis with censored data. Biometrics 48:507–528

Fang KT, Kotz S, Ng KW (1990) Symmetric multivariate and related distribution. Chapman and Hall, London

Fernandez C, Steel M (1999) Multivariate Student-*t* regression models: pitfalls and inference. Biometrika 86:153–167

Galea M, Leiva V, Paula GA (2004) Influence diagnostics in log-Birnbaum-Saunders regression models. J Appl Stat 31:1049–1064

Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions, vol 2. Wiley, New York

Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data. Wiley, New York

Klein JP, Moeschberger ML (1997) Survival analysis: techniques for censored and truncated data. Springer, New York

Lange KL, Little JA, Taylor MG (1989) Robust statistical modelling using the *t* distribution. J Amer Stat Soc 84:881–896

Lawless JF (2002) Statistical models and methods for lifetime data, 2nd edn. Wiley, New York

Lee ET, Wang JW (2003) Statistical methods for survival data analysis. Wiley, New York

Leiva V, Barros M, Paula GA, Galea M (2007) Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. Comput Stat Data Anal 51:5694–5707

Leiva V, Barros M, Paula GA, Sanhueza A (2008a) Generalized Birnbaum-Saunders distribution applied to air pollutant concentration. Environmetrics 19 (in press) (doi:10.1002/env.861)

Leiva V, Riquelme M, Balakrishnan N, Sanhueza A (2008b) Lifetime analysis based on the generalized Birnbaum-Saunders distribution. Comput Stat Data Anal 52:2079–2097

Leiva V, Sanhueza A, Angulo JM (2008c) A length-biased version of the Birnbaum-Saunders distribution with application in water quality. Stoch Environ Res Risk Assess (in press) (doi:10.1007/s00477-008-0215-9)

Lesaffre E, Verbeke G (1998) Local influence in linear mixed models. Biometrics 54:570–582

Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. Wiley, New York

Ortega EMM, Bolfarine H, Paula GA (2003) Influence diagnostic in generalized log-gamma regression models. Comput Stat Data Anal 40:165–186

Owen WJ, Padgett WJ (1999) Accelerated test models for system strength based on Birnbaum-Saunders distribution. Lifetime Data Anal 5:133–147

Owen WJ, Padgett WJ (2000) A Birnbaum-Saunders accelerated life model. IEEE Trans Rel 49:224–229

Poon WY, Poon YS (1999) Conformal normal curvature and assessment of local influence. J Roy Stat Soc B 61:51–61

Rieck JR, Nedelman JR (1991) A log-linear model for the Birnbaum-Saunders distribution. Technometrics 33:51–60

Sanhueza A, Leiva V, Balakrishnan N (2008) The generalized Birnbaum-Saunders distribution and its theory, methodology and application. Commun Stat Theor Meth 37:645–670

Saunders SC (1974) A family of random variables closed under reciprocation. J Amer Stat Soc 6:319–327

Taylor J, Verbyla A (2004) Joint modeling of location and scale parameters of *t* distribution. Stat Model 4:91–112

Therneau TM, Grambsch PM, Fleming TR (1990) Martingale-based residuals for survival models. Biometrika 77:147–160

Tsionas EG (2001) Bayesian inference in Birnbaum-Saunders regression. Commun Stat Theor Meth 30:179–193

Verbeke G, Molenberghs G (2000) Linear mixed models for longitudinal data. Springer, New York

Xie FC, Wei BC (2007) Diagnostics analysis for log-Birnbaum-Saunders regression models. Comput Stat Data Anal 51:4692–4706