Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Computational Statistics & Data Analysis 51 (2007) 5694 – 5707
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

# Influence diagnostics in log-Birnbaum–Saunders regression models with censored data

Víctor Leiva[a,*], Michelli Barros[b], Gilberto A. Paula[b], Manuel Galea[a]

[a]Departamento de Estadística, CIMFAV, Universidad de Valparaíso, Casilla de correos 5030, Valparaíso, Chile
[b]Departamento de Estatística, Universidade de São Paulo, Brazil

Received 1 June 2006; received in revised form 17 September 2006; accepted 19 September 2006
Available online 13 October 2006

**Abstract**

In this paper we discuss log-Birnbaum–Saunders regression models with censored observations. This kind of model has been largely applied to study material lifetime subject to failure or stress. The score functions and observed Fisher information matrix are given as well as the process for estimating the regression coefficients and shape parameter is discussed. The normal curvatures of local influence are derived under various perturbation schemes and two deviance-type residuals are proposed to assess departures from the log-Birnbaum–Saunders error assumption as well as to detect outlying observations. Finally, a data set from the medical area is analyzed under log-Birnbaum–Saunders regression models. A diagnostic analysis is performed in order to select an appropriate model.


*Keywords:* Birnbaum–Saunders distribution; Generalized leverage; Influence diagnostic; Lifetime distributions; Local influence; Log-linear model; Maximum likelihood estimator; Residual analysis; Sinh-normal distribution

## 1. Introduction

In some two-parameter probability models that more popularly have been proposed to describe lifetime, we find these distributions: gamma, lognormal and Weibull, all of which fit with precision into the central zone of the lifetime distributions. Nevertheless, it is important to concern ourselves with the low or high percentiles of the distributions, zones precisely where little data are usually found. This in turn leads to a poor fit in these zones of the models previously mentioned.

An important lifetime model originated from a problem of material fatigue, is the one developed by Birnbaum and Saunders (1969). This model fits well within the extremes of the distribution, even when there is little lifetime data, due to the physical justification that originated it. The Birnbaum–Saunders (B–S) distribution is a lifetime model for fatigue failure caused under cyclic loading and assumed that the failure is due to the development and growth of a dominant crack. A more general derivation was provided by Desmond (1985) based on a biological model. Later, Rieck and Nedelman (1991) derived the log-Birnbaum–Saunders (log-BS) distribution based on an exponential regression model. The log-BS distribution is a particular case of the sinh-normal (SN) distribution of Rieck and Nedelman (1991).

* Corresponding author. Tel.: +56 32 2508273; fax: +56 32 2508322.
*E-mail addresses:* victor.leiva@uv.cl, victor.leiva@yahoo.com (V. Leiva).

0167-9473/$ - see front matter 
doi:10.1016/j.csda.2006.09.020

The SN distribution is symmetrical, presents greater and smaller degrees of kurtosis than the normal model and also has bi-modality.

Generally, the analysis with symmetrical non-normal distributions considers the Student-$t$ model, since it has greater kurtosis than the normal model. Thus, observations which are considered as outlying under normality, are not outlying under the $t$ law, producing more robust parameter estimates. Nevertheless, in genetics and risk analysis it is also important to emphasize atypical cases. Thus, symmetrical distributions whose kurtosis are smaller than the normal one are also of interest and the SN distribution appears as an interesting alternative to the $t$ and normal distributions.

An efficient way to detect influential observations is the diagnostic analysis. The techniques of case elimination based on different types of residuals, at present known as global influence methods, have been discussed widely (see, for example, Cook and Weisberg, 1982). Nevertheless, Cook (1986) proposed a diagnostic approach named local influence to assess the effect of small perturbations in the model and/or data on the parameter estimates. Several authors have applied the local influence methodology in regression models more general than the normal regression model (see, for example, Paula, 1993; Galea et al., 1997, 2000 and Díaz-García et al., 2003). Also, some authors have investigated the assessment of local influence in survival analysis models: for instance, Pettit and Bin Daud (1989) investigate local influence in proportional hazard regression models, Escobar and Meeker (1992) adapt local influence methods to regression analysis with censoring, Ortega et al. (2003) consider the problem of assessing local influence in generalized log-gamma regression models with censored observations and more recently Galea et al. (2004) derive the curvature calculations under various perturbation schemes in log-BS linear regression models with uncensored data. Besides taking the likelihood function for assessing the curvature for influence analysis, other suggestions 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.

In this article, we present diagnostic methods based on local influence and residual analysis for linear regression models under the log-BS distribution for the errors and censored observations. In Section 2, we present the B–S distribution and its relationship with the SN distribution, together with a brief analysis of this distribution. In Section 3, log-BS linear regression models for censoring observations are defined. The score functions and observed Fisher information matrix are given as well as the process for estimating the regression coefficients and the shape parameter is discussed. In Section 4, the normal curvatures of local influence are derived under various perturbation schemes and two kinds of deviance-type residuals are proposed to assess departures from the log-BS error assumption as well as outlying observations. An application with real data, that have not been analyzed from a perspective of diagnostic, is discussed in Section 5. Finally, some concluding remarks are made in Section 6.

## 2. The sinh-normal distribution

The B–S distribution is defined in terms of the normal distribution by means of the random variable (r.v.) $T = \beta[\alpha Z/2 + \sqrt{(\alpha Z/2)^2 + 1}]^2$, where $Z \sim N(0, 1)$, $\alpha > 0$ is the shape parameter and $\beta > 0$ is the scale parameter and the median. This is denoted by $T \sim BS(\alpha, \beta)$. The probability density function (pdf) of $T$ is

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2}\left[\frac{t}{\beta} + \frac{\beta}{t} - 2\right]\right) \frac{t^{-3/2}(t+\beta)}{2\alpha\sqrt{\beta}}, \quad t > 0. \tag{1}$$

If $T \sim BS(\alpha, \beta)$, then $Z = \alpha^{-1}(\sqrt{T/\beta} - \sqrt{\beta/T}) \sim N(0, 1)$. Furthermore, $cT \sim BS(\alpha, c\beta)$, with $c > 0$, $T^{-1} \sim BS(\alpha, \beta^{-1})$, and the cumulative distribution function (cdf) of $T$ is $F_T(t) = \Phi(\alpha^{-1}(\sqrt{t/\beta} - \sqrt{\beta/t}))$, where $\Phi(\cdot)$ is the standard normal cdf.

Rieck and Nedelman (1991) analyzed the exponential regression model $T = \mathrm{e}^{X\theta}\varepsilon$, where $T \sim BS(\alpha, \beta)$. Thus, using the proportionality property of the B–S distribution, we have that $\varepsilon \sim BS(\alpha, \mathrm{e}^{-X\theta}\beta)$. Then, the associated log-linear model is $Y = \log T = X\theta + \log \varepsilon$. This problem requires knowing the distribution of $Y = \log T$, which is a particular case of the SN distribution.

The SN distribution is defined in terms of the normal distribution by means of the r.v. $Y = \mathrm{arcsinh}(\alpha Z/2)\sigma + \gamma$, where $Z \sim N(0, 1)$, $\alpha > 0$ is the shape parameter, $\gamma \in \mathbb{R}$ is the location parameter and $\sigma > 0$ is the scale parameter.

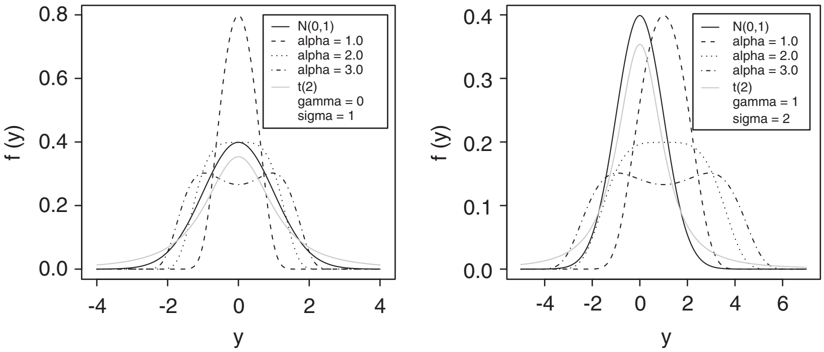


Fig. 1. Pdf graphs of the SN distribution for the indicated parameters.

This is denoted by $Y \sim SN(\alpha, \gamma, \sigma)$. The pdf of $Y$ is

$$f_Y(y) = \left(\frac{2}{\alpha\sigma\sqrt{2\pi}}\right)\cosh\left(\frac{y-\gamma}{\sigma}\right)\exp\left(-2\alpha^{-2}\sinh^2\left[\frac{y-\gamma}{\sigma}\right]\right), \quad y \in \mathbb{R}. \tag{2}$$

Thus, if $Y \sim SN(\alpha, \gamma, \sigma)$, then $Z = 2\alpha^{-1}\ \sinh((Y-\gamma)/\sigma) \sim N(0, 1)$. The cdf of $Y$ is

$$F_Y(y) = \Phi\left(\frac{2}{\alpha}\ \sinh\left[\frac{y-\gamma}{\sigma}\right]\right), \tag{3}$$

where $\Phi(\cdot)$ is the standard normal cdf.

Rieck (1989) notes that the $SN(\alpha, \gamma, \sigma)$ distribution is symmetrical around $\gamma$, strongly unimodal for $\alpha \leqslant 2$ and bimodal for $\alpha > 2$. Furthermore, if $Y \sim SN(\alpha, \gamma, \sigma)$, then $U = 2\,\alpha^{-1}(Y-\gamma)/\sigma$ converges to the standard normal distribution as $\alpha \to 0$. The mean and variance can be obtained using the moment generating function given by

$$m(s) = \exp(\mu s)\left[\frac{K_a(\delta^{-2}) + K_b(\delta^{-2})}{2K_{1/2}(\delta^{-2})}\right],$$

where $a = (\sigma s + 1)/2$, $b = (\sigma s - 1)/2$ and $K_\lambda(\cdot)$ is the modified Bessel function of the third kind given by $K_\lambda(w) = (\frac{1}{2})(w/2)^\lambda \int_0^\infty y^{-\lambda-1}\mathrm{e}^{-y-(w^2/4y)}\,\mathrm{d}y$ (see Gradshteyn and Randzhik, 2000, p. 907).

Rieck and Nedelman (1991) developed the SN distribution and proved that if $T \sim BS(\alpha, \beta)$, then $Y = \log\, T \sim SN(\alpha, \gamma, \sigma = 2)$, where $\gamma = \log\, \beta$. For this reason, the sinh-normal distribution is also named the log-BS distribution.

In general, if $Y \sim SN(\alpha, \gamma, \sigma)$, when $\alpha$ increases, the kurtosis of the SN distribution also increases. In particular, for $\alpha \leqslant 2$, one has unimodality and smaller kurtosis than the normal case. Nevertheless, for $\alpha > 2$, when $\alpha$ increases, the SN distribution begins to emphasize the bimodality, has modes that are more separated and the kurtosis is greater than the normal case. The parameter $\gamma$ modifies the location and the parameter $\sigma$ the scale of the distribution, as it can be seen in Fig. 1.

## 3. Log-Birnbaum–Saunders regression models

### 3.1. Model and estimation

We consider the following log-BS regression models:

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

where $y_i$ is the observed log-lifetime or log-censoring time for the $i$th individual, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^{\mathrm{T}}$ is a vector of unknown parameters to be estimated, $\boldsymbol{x}_i^{\mathrm{T}} = (x_{i1}, \ldots, x_{ip})$ contains values of explanatory variables and $\varepsilon_i \sim SN(\alpha, 0, 2)$. We will assume non-informative censoring and that the observed lifetime and censoring time are independent. The sets

$D$ and $C$ will denote the sets of individuals for which $y_i$ is the log-lifetime or log-censoring, respectively. The total log-likelihood function of the model (4) for $\boldsymbol{\theta}=(\boldsymbol{\beta}^{\mathrm{T}},\alpha)^{\mathrm{T}}$ is given by

$$l(\boldsymbol{\theta}) \propto \sum_{i\in D} l_i(\boldsymbol{\theta}) + \sum_{i\in C} l_{ic}(\boldsymbol{\theta}),$$

where $l_i(\boldsymbol{\theta})=\log\ f_Y(y_i;\boldsymbol{\theta})$, $l_{ic}(\boldsymbol{\theta})=\log\ S_Y(y_i;\boldsymbol{\theta})$ and $S_Y(y;\boldsymbol{\theta})=1-F_Y(y;\boldsymbol{\theta})$ is the survival function, with $f_Y(y;\boldsymbol{\theta})$ given in (2) and $F_Y(\cdot)$ in (3). Then, the log-likelihood function for $\boldsymbol{\theta}$ can be expressed as

$$l(\boldsymbol{\theta}) \propto \sum_{i\in D}\left[\log(\xi_{i1}) - \frac{1}{2}\left\{\log(8\pi)+\xi_{i2}^2\right\}\right] + \sum_{i\in C}\log(1-\Phi[\xi_{i2}]), \tag{5}$$

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

The score functions for $\alpha$ and $\beta_j$ $(j=1,\ldots,p)$ are, respectively, given by

$$U_\alpha(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial\alpha} = \frac{1}{\alpha}\left[\sum_{i\in D}(\xi_{i2}^2-1) + \sum_{i\in C}\xi_{i2}h(\xi_{i2})\right] \tag{6}$$

and

$$\begin{aligned} U_{\beta_j}(\boldsymbol{\theta}) &= \frac{\partial l(\boldsymbol{\theta})}{\partial\beta_j} \\ &= \sum_{i\in D}\left[\frac{x_{ij}}{\alpha^2}\ \sinh(y_i-\mu_i) - \frac{x_{ij}}{2}\ \tanh\left(\frac{y_i-\mu_i}{2}\right)\right] + \frac{1}{2}\sum_{i\in C} x_{ij}\xi_{i1}h(\xi_{i2}), \end{aligned} \tag{7}$$

where $h(\xi_{i2})=\phi(\xi_{i2})/(1-\Phi(\xi_{i2}))$ with $\phi(\cdot)$ and $\Phi(\cdot)$ denoting the pdf and the cdf of a standard normal distribution, respectively.

The MLEs of the regression coefficients and the shape parameter are solutions of the equations $U_\alpha(\boldsymbol{\theta})=0$ and $U_{\beta_j}(\boldsymbol{\theta})=0$ $(j=1,\ldots,p)$. However, such equations do not present analytical solutions and the use of iterative methods is necessary to find the roots. The BFGS method (see, for example, Press et al., 1992) has been used by the authors for maximizing the total log-likelihood function $l(\boldsymbol{\theta})$. As initial values we have considered the estimates under the uncensored case for $\alpha$ (see details in Galea et al., 2004) and for $\boldsymbol{\beta}$ the least-squares estimate. The asymptotic inference for the parameter vector, $\boldsymbol{\theta}$, can be based on the normal approximation of the MLE of $\hat{\boldsymbol{\theta}}$ given by

$$\hat{\boldsymbol{\theta}} \dot{\sim} N_{(p+1)}(\boldsymbol{\theta},\Sigma_\theta),$$

where $\Sigma_\theta$ denotes the asymptotic variance–covariance matrix for $\hat{\boldsymbol{\theta}}$, which may be approximated by $-\ddot{\boldsymbol{L}}_{\theta\theta}^{-1}$, where $-\ddot{\boldsymbol{L}}_{\theta\theta}$ is the $(p+1)\times(p+1)$ observed information matrix, obtained from

$$\ddot{\boldsymbol{L}}_{\theta\theta} = \begin{bmatrix}\ddot{\boldsymbol{L}}_{\beta\beta} & \ddot{\boldsymbol{L}}_{\beta\alpha}\\ \ddot{\boldsymbol{L}}_{\alpha\beta} & \ddot{\boldsymbol{L}}_{\alpha\alpha}\end{bmatrix} = \begin{bmatrix}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{X} & \boldsymbol{X}^{\mathrm{T}}\boldsymbol{k}\\ \boldsymbol{k}^{\mathrm{T}}\boldsymbol{X} & \mathrm{tr}(\boldsymbol{G})\end{bmatrix}, \tag{8}$$

where $\boldsymbol{V}=\mathrm{diag}\{v_1(\boldsymbol{\theta}),\ldots,v_n(\boldsymbol{\theta})\}$, with

$$v_i(\boldsymbol{\theta}) = \begin{cases}\frac{1}{4}\mathrm{sech}^2\left(\frac{y_i-\mu_i}{2}\right) - \frac{1}{\alpha^2}\ \cosh(y_i-\mu_i) & \text{if } i\in D,\\ -\frac{1}{4}\xi_{i2}h(\xi_{i2}) - \frac{1}{4}\xi_{i1}^2h'(\xi_{i2}) & \text{if } i\in C,\end{cases} \tag{9}$$

$\boldsymbol{k}=(k_1(\boldsymbol{\theta}),\ldots,k_n(\boldsymbol{\theta}))^{\mathrm{T}}$, with

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

and $\boldsymbol{G} = \mathrm{diag}\{g_1(\boldsymbol{\theta}), \ldots, g_n(\boldsymbol{\theta})\}$, with

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

where $h(\cdot)$ is given in (7), $h'(\cdot)$ is its derivative and $\xi_{i1}, \xi_{i2}$ are given in (5).

## 4. Diagnostic analysis

### 4.1. Local influence

Local influence calculation can be carried out in the model given in (4). Consider $\hat{\boldsymbol{\theta}}_\omega$ as the MLE under the perturbed model. In addition, let $\boldsymbol{\Delta} = (\Delta_{ij})$ be a $(p+1) \times n$ matrix that depends on the perturbation scheme, where $\Delta_{ij} = \partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\theta_i\partial\omega_j$, with $i = 1, \ldots, p+1$ and $j = 1, \ldots, n$, which is evaluated at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$, being $\boldsymbol{\omega}_0$ the no perturbation vector (see Cook, 1986). Then, if the likelihood displacement $LD(\omega) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_\omega)\}$ is used, the normal curvature for $\boldsymbol{\theta}$ at the direction $\boldsymbol{l}$ is given by $C_\ell(\boldsymbol{\theta}) = 2|\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\Delta}^{\mathrm{T}}\ddot{\boldsymbol{L}}_{\theta\theta}^{-1}\boldsymbol{\Delta}\boldsymbol{l}|$, where $\|\boldsymbol{l}\| = 1$. For the log-BS regression models the elements of $\ddot{\boldsymbol{L}}_{\hat{\theta}\hat{\theta}}$ are $\ddot{\boldsymbol{L}}_{\hat{\beta}\hat{\beta}} = \boldsymbol{X}^{\mathrm{T}}\hat{\boldsymbol{V}}\boldsymbol{X}$, $\ddot{\boldsymbol{L}}_{\hat{\beta}\hat{\alpha}} = \boldsymbol{X}^{\mathrm{T}}\hat{\boldsymbol{k}}$ and $\ddot{\boldsymbol{L}}_{\hat{\alpha}\hat{\alpha}} = \mathrm{tr}(\hat{\boldsymbol{G}})$.

We can also calculate the normal curvatures $C_\ell(\boldsymbol{\beta})$ and $C_\ell(\alpha)$ to perform various index plots. For instance, the index plot of $\boldsymbol{l}_{\max}$ the eigenvector corresponding to $C_{\ell_{\max}}$ the largest eigenvalue of the matrix $B = -\boldsymbol{\Delta}^{\mathrm{T}}\ddot{\boldsymbol{L}}_{\theta\theta}^{-1}\boldsymbol{\Delta}$ and the index plots of $C_{\ell_i}(\boldsymbol{\beta})$ and $C_{\ell_i}(\alpha)$, named total local influence (see, for example, Lesaffre and Verbeke, 1998), where $\boldsymbol{l}_i$ denotes an $n \times 1$ vector of zeros with one at the $i$th position. Thus, the curvature at the direction $\boldsymbol{l}_i$ assumes the form $C_i = 2|\boldsymbol{\Delta}_i^{\mathrm{T}}\ddot{\boldsymbol{L}}_{\theta\theta}^{-1}\boldsymbol{\Delta}_i|$, where $\boldsymbol{\Delta}_i^{\mathrm{T}}$ denotes the $i$th row of $\boldsymbol{\Delta}$. Verbeke and Molenberghs (2000, Section 11.3) propose considering as point out those cases such that $C_i \geqslant 2\bar{C}$, where $\bar{C} = (1/n)\sum_{i=1}^n C_i$.

#### 4.1.1. Curvature calculations

Next, we will calculate for three perturbation schemes, the $(p+1) \times n$ matrix

$$\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ij})_{(p+1)\times n} = \left(\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\theta_i\partial\omega_j}\right), \quad i = 1, \ldots, p+1 \quad \text{and} \quad j = 1, \ldots, n, \tag{10}$$

considering the model defined in (4) and its log-likelihood function given in (5).

*Case-weights perturbation*: Consider the vector of weights $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^{\mathrm{T}}$ and the perturbed log-likelihood function

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) \propto \sum_{i\in D}\omega_i\left[\log(\xi_{i1}) - \frac{1}{2}\{\log(8\pi) + \xi_{i2}^2\}\right] + \sum_{i\in C}\omega_i\ \log(1 - \Phi[\xi_{i2}]), \tag{11}$$

where $0 \leqslant \omega_i \leqslant 1$ and $\boldsymbol{\omega}_0 = (1, \ldots, 1)^{\mathrm{T}}$. The matrix $\boldsymbol{\Delta}$ is given by

$$\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_\beta \\ \boldsymbol{\Delta}_\alpha \end{pmatrix}, \tag{12}$$

where $\boldsymbol{\Delta}_\beta$ is a $p \times n$ matrix expressed as $\boldsymbol{\Delta}_\beta = \boldsymbol{X}^{\mathrm{T}}\,\mathrm{diag}\{\hat{b}_1, \ldots, \hat{b}_n\}$ with

$$\hat{b}_i = \begin{cases} \frac{1}{2}(\hat{\xi}_{i1}\hat{\xi}_{i2} - \hat{\xi}_{i2}/\hat{\xi}_{i1}) & \text{if } i \in D, \\ \frac{\hat{\xi}_{i1}}{2}\,h(\hat{\xi}_{i2}) & \text{if } i \in C, \end{cases} \tag{13}$$

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

$$\hat{a}_i = \begin{cases} -\frac{1}{\hat{\alpha}} + \frac{(\hat{\xi}_{i2})^2}{\hat{\alpha}} & \text{if } i \in D, \\ \frac{\hat{\xi}_{i2}}{\hat{\alpha}}\,h(\hat{\xi}_{i2}) & \text{if } i \in C, \end{cases} \tag{14}$$

where $h(\cdot)$ is given in (7) and $\hat{\xi}_{i1}, \hat{\xi}_{i2}$ are $\xi_{i1}, \xi_{i2}$ given in (5) evaluated at $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^{\mathrm{T}}, \hat{\alpha})^{\mathrm{T}}$.

*Response perturbation*: We will 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 the estimated standard deviation of $Y$ and $\omega_i \in \mathbb{R}$. The perturbed log-likelihood function takes the form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) \propto \sum_{i\in D}\left[\log(\xi_{i1\omega_1}) - \frac{1}{2}\{\log(8\pi) + \xi^2_{i2\omega_1}\}\right] + \sum_{i\in C}\log(1-\Phi[\xi_{i2\omega_1}]), \tag{15}$$

where $\xi_{i1\omega_1} = (2/\alpha)\cosh((y_{i\omega}-\mu_i)/2)$, $\xi_{i2\omega_1} = (2/\alpha)\sinh((y_{i\omega}-\mu_i)/2)$ and $\boldsymbol{\omega}_0 = (0,\ldots,0)^{\mathrm{T}}$. The matrix $\boldsymbol{\Delta}$ assumes the same form given in (12), where

$$\boldsymbol{\Delta}_\beta = \boldsymbol{X}^{\mathrm{T}}\mathrm{diag}\{\hat{d}_1,\ldots,\hat{d}_n\},$$

with

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

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

$$\hat{c}_i = \begin{cases} \frac{S_y}{\hat{\alpha}}\hat{\xi}_{i1}\hat{\xi}_{i2} & \text{if } i \in D,\\ \frac{S_y}{2\hat{\alpha}}[\hat{\xi}_{i1}h(\hat{\xi}_{i2}) + \hat{\xi}_{i1}\hat{\xi}_{i2}h'(\hat{\xi}_{i2})] & \text{if } i \in C,\end{cases}$$

where $h(\cdot)$, $h'(\cdot)$ are given in (7) and $\hat{\xi}_{i1\omega_1}$, $\hat{\xi}_{i2\omega_1}$ are $\xi_{i1\omega_1}$, $\xi_{i2\omega_1}$ given in (15) evaluated at $\hat{\boldsymbol{\theta}}$.

*Explanatory variable perturbation*: Consider now an additive perturbation on a particular continuous explanatory variable, namely $\boldsymbol{x}_t$, by making $x_{it\omega} = x_{it} + \omega_i S_x$, where $S_x$ is a scale factor that may be the estimated standard deviation of $\boldsymbol{x}_t$ and $\omega_i \in \mathbb{R}$. The perturbed log-likelihood function takes here the form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) \propto \sum_{i\in D}\left[\log(\xi_{i1\omega_2}) - \frac{1}{2}\{\log(8\pi) + \xi^2_{i2\omega_2}\}\right] + \sum_{i\in C}\log(1-\Phi[\xi_{i2\omega_2}]), \tag{16}$$

where $\xi_{i1\omega_2} = (2/\alpha)\cosh((y_i-\mu_i-\beta_t\omega_i S_x)/2)$, $\xi_{i2\omega_2} = (2/\alpha)\sinh((y_i-\mu_i-\beta_t\omega_i S_x)/2)$ and $\boldsymbol{\omega}_0 = (0,\ldots,0)^{\mathrm{T}}$. The matrix $\boldsymbol{\Delta}$ assumes the previous form given in (12), where $\boldsymbol{\Delta}_\beta$ is a $p \times n$ matrix with $\boldsymbol{\Delta}_{\beta_{ij}}$ elements which when $j \neq t$ assume the form

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

and when $j = t$ one has

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

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

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

where $h(\cdot)$, $h'(\cdot)$ are given in (7) and $\hat{\xi}_{i1}$, $\hat{\xi}_{i2}$ are $\xi_{i1}$, $\xi_{i2}$ given in (16) evaluated at $\hat{\boldsymbol{\theta}}$.

*Non-censoring case perturbation*: Suppose that only the non-censoring cases are perturbed, which means to assume $\omega_i = 1$ for $i \in C$ in expression (11). Thus, the quantities $\hat{b}_i$ and $\hat{a}_i$ given in (13) and (14) reduce, respectively, to

$$\hat{b}_i = \begin{cases} \frac{1}{2}(\hat{\xi}_{i1}\hat{\xi}_{i2} - \hat{\xi}_{i2}/\hat{\xi}_{i1}) & \text{if } i \in D,\\ 0 & \text{if } i \in C,\end{cases}$$

and

$$\hat{a}_i = \begin{cases} -\frac{1}{\hat{\alpha}} + \frac{(\hat{\xi}_{i2})^2}{\hat{\alpha}} & \text{if } i \in D, \\ 0 & \text{if } i \in C, \end{cases}$$

where $\hat{\xi}_{i1}$, $\hat{\xi}_{i2}$ are $\xi_{i1}$, $\xi_{i2}$ given in (5) evaluated at $\hat{\boldsymbol{\theta}}$. Note that the expressions above for $i \in D$ agree with ones obtained by Galea et al. (2004).

*Censoring case perturbation*: Suppose now that only the censoring cases are perturbed, that is, $\omega_i = 1$ for $i \in D$ in expression (11). In this case, the quantities $\hat{b}_i$ and $\hat{a}_i$ given in (13) and (14) reduce, respectively, to

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

and

$$\hat{a}_i = \begin{cases} 0 & \text{if } i \in D, \\ \frac{\hat{\xi}_{i2}}{\hat{\alpha}}\, h(\hat{\xi}_{i2}) & \text{if } i \in C, \end{cases}$$

where $h(\cdot)$ is given in (7) and $\hat{\xi}_{i1}$, $\hat{\xi}_{i2}$ are $\xi_{i1}$, $\xi_{i2}$ given in (5) evaluated at $\hat{\boldsymbol{\theta}}$.

### 4.2. Residual analysis

In order to study departures from the error assumptions as well as presence of outlying observations, we will consider two kinds of residuals: deviance component residual (see, for example, McCullagh and Nelder, 1989) and martingale-type residual (see, for example, Barlow and Prentice, 1988; Therneau et al., 1990). More details may be found in Ortega et al. (2003).

#### 4.2.1. Deviance component residual

The deviance component residual is defined by $r_{\mathrm{DC}_i} = \mathrm{sign}(y_i - \hat{\mu}_i)\sqrt{2}[l_i(\tilde{\boldsymbol{\theta}}) - l_i(\hat{\boldsymbol{\theta}})]^{1/2}$, where $\widetilde{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ under the saturated model (with $n$ parameters), $\hat{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ under the model of interest (with $p$ parameters) and $\mathrm{sign}(z)$ denotes the sign of $z$. Davison and Gigli (1989) defined the deviance component residual for censored data as

$$r_{\mathrm{DC}_i} = \mathrm{sign}(y_i - \hat{\mu}_i)[-2\,\log(S(y_i, \hat{\boldsymbol{\theta}}))]^{1/2}, \tag{17}$$

where $S(y_i, \hat{\boldsymbol{\theta}})$ is the MLE of the survival function (see, for example, Ortega, 2001). Therefore, considering fixed or known $\alpha$ and $\alpha < 2$, we have that the deviance component residual for log-BS regression models becomes

$$r_{\mathrm{DC}_i} = \begin{cases} \mathrm{sign}(y_i - \hat{\mu}_i)\sqrt{2}\Big[-\log\Big(\cosh\Big[\frac{y_i-\hat{\mu}_i}{2}\Big]\Big) + \frac{2}{\hat{\alpha}^2}\sinh^2\Big(\frac{y_i-\hat{\mu}_i}{2}\Big)\Big]^{1/2} & \text{if } i \in D, \\ \mathrm{sign}(y_i - \hat{\mu}_i)\Big[-2\,\log\Big(1 - \Phi\Big[\frac{2}{\hat{\alpha}}\sinh\Big(\frac{y_i-\hat{\mu}_i}{2}\Big)\Big]\Big)\Big]^{1/2} & \text{if } i \in C. \end{cases}$$

#### 4.2.2. Martingale-type residual

Therneau et al. (1990) introduced the deviance component residual in counting process by using basically martingale residuals. The martingale residuals are skewness, have maximum value +1 and minimum value $-\infty$. In parametric lifetime models, the martingale residual can be expressed as $r_{\mathrm{M}_i} = \delta_i + \log[S(y_i, \hat{\boldsymbol{\theta}})]$, with $\delta_i = 0$ indicating if the observation is censored and $\delta_i = 1$ indicating if the observation is uncensored (see, for example, Klein and Moeschberger, 1997; Ortega et al., 2003). Thus, the martingale residual for log-BS regression models assumes the form

$$r_{\mathrm{M}_i} = \delta_i + \log\left(1 - \Phi\left[\frac{2}{\hat{\alpha}}\sinh\left(\frac{y_i - \hat{\mu}_i}{2}\right)\right]\right).$$

The deviance component residual proposed by Therneau et al. (1990) is a transformation of the martingale residual to attenuate the skewness. This transformation was motivated by the deviance component residuals found in generalized linear models. In particular, the deviance component residual for the Cox's proportional hazard model with no time-dependent explanatory variables is described as

$$r_{\mathrm{MD}_i} = \mathrm{sign}(r_{\mathrm{M}_i})[-2\{r_{\mathrm{M}_i} + \delta_i \ \log(\delta_i - r_{\mathrm{M}_i})\}]^{1/2},$$

where $r_{\mathrm{M}_i}$ is the martingale residual.

Although the residuals $r_{\mathrm{MD}_i}$ are not deviance components of log-BS regression models, we will use them in the following as a transformation of the martingale residuals in order to have residuals symmetrically distributed around zero. We will name $r_{\mathrm{MD}_i}$ martingale-type residual.

#### 4.2.3. Standardized residuals

Ortega (2001) suggests to standardize the deviance component residual $r_{\mathrm{DC}_i}$ and the martingale-type residual $r_{\mathrm{MD}_i}$ for censored data as

$$r^*_{\mathrm{DC}_i} = \frac{r_{\mathrm{DC}_i}}{\sqrt{1 - GL_{ii}}}$$

and

$$r^*_{\mathrm{MD}_i} = \frac{r_{\mathrm{MD}_i}}{\sqrt{1 - GL_{ii}}},$$

$GL_{ii}$ being the $i$th principal diagonal element of the generalized leverage matrix (see Wei et al., 1998) defined by

$$\boldsymbol{GL}(\boldsymbol{\theta}) = \boldsymbol{D}_\theta(-\ddot{\boldsymbol{L}}_{\theta\theta})^{-1}\ddot{\boldsymbol{L}}_{\theta y},$$

where $\boldsymbol{D}_\theta = \partial\boldsymbol{\mu}/\partial\boldsymbol{\theta}^{\mathrm{T}} = (\boldsymbol{X},\ \boldsymbol{0})$, $\ddot{\boldsymbol{L}}_{\theta\theta}$ is given in (8) and

$$\ddot{\boldsymbol{L}}_{\theta y} = \begin{pmatrix} \ddot{\boldsymbol{L}}_{\beta y} \\ \ddot{\boldsymbol{L}}_{\alpha y} \end{pmatrix},$$

where $\ddot{\boldsymbol{L}}_{\beta y} = -\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}$, with $\boldsymbol{V} = \mathrm{diag}\{v_1(\boldsymbol{\theta}), \ldots, v_n(\boldsymbol{\theta})\}$, $v_i(\boldsymbol{\theta})$ is given in (9) and

$$\ddot{\boldsymbol{L}}_{\alpha y_i} = \begin{cases} \frac{1}{\alpha}\xi_{i1}\xi_{i2} & \text{if } i \in D, \\ \frac{1}{2\alpha}\xi_{i1}h(\xi_{i2}) + \frac{1}{2\alpha}\xi_{i1}\xi_{i2}h'(\xi_{i2}) & \text{if } i \in C, \end{cases}$$

with all the expressions evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Index plots of $GL_{ii}$ may reveal those observations with high influence on their own predicted values.

#### 4.2.4. Simulation study

In order to investigate the form of the empirical distributions of the residuals $r^*_{\mathrm{DC}_i}$ and $r^*_{\mathrm{MD}_i}$ for the values of $n = 30, 50, 100$, $\alpha = 0.5, 1.0, 1.5$ and censoring proportion $p = 0.10, 0.30, 0.50$, we performed a small simulation study described in the sequel. We considered a single explanatory variable $x$ which follows a uniform distribution (0, 1). The Birnbaum–Saunders observations were generated through normal observations using the expression

$$T_i = \eta_i\left(1 + \frac{\alpha^2 Z_i^2}{2} + \alpha Z_i\left(1 + \frac{\alpha^2 Z_i^2}{4}\right)^{1/2}\right),$$

where $Z_i \sim N(0, 1)$ and $\eta_i = \exp(\beta_0 + \beta_1 x_i)$, with $\beta_0 = 4$ and $\beta_1 = -2$, for $i = 1, \ldots, n$. Following the procedures presented by Wang et al. (2006), we generated the censoring random variable $C_i$ from a log-normal distribution with cdf $\Phi((\log u - \mu)/\sigma)$, where $\Phi(\cdot)$ is the standard normal cdf. We fixed $\sigma$ as 1 and let $\mu$ vary to obtain different censoring proportions. The observed lifetimes and the censoring indicators were generated from $Z_i = \min(T_i, C_i)$ and $\delta_i = I_{(T_i \leqslant C_i)}$, respectively, being $I_C$ the indicator function. For each setting of $n$, $\alpha$ and $p$, 1000 random samples were generated. For each sample, the log-BS regression model given in (4), where $Y_i = \log(Z_i)$, with $i = 1, \ldots, n$, was fitted. Then, we obtained the residuals $r^*_{\mathrm{DC}_i}$ and $r^*_{\mathrm{MD}_i}$ for each generated sample. We performed normal probability plots between the mean quantiles of the residuals and the expected quantiles of the standard normal distribution.

The main conclusion obtained from the generated graphs is that the empirical distributions of the two residuals present a good agreement with the standard normal distribution for each $n$, $\alpha$ and censoring proportion given. In particular, as the censoring proportion decreases, the empirical distributions of the residuals approach faster to the standard normal distribution. In addition, as $\alpha$ increases, the empirical distributions present a weak skewness. The extension of the results for more general situations is not straightforward. Thus, as suggested by Atkinson (1981) and Williams (1987), the use of normal probability plots for $r^*_{\mathrm{DC}_i}$ and $r^*_{\mathrm{MD}_i}$ with envelope is recommended. Graphs of such residuals against the fitted values and explanatory variable values may be useful even though the interpretation of them is not necessarily the same of that used in normal linear regression models.

## 5. Application

Next, as illustration, we will consider the data set given in Krall et al. (1975) and reported in Lawless (1982, pp. 332–333) connected with survival analysis. This data set has recently been analyzed by Jin et al. (2003) and Ghosh and Ghosal (2006) that uses a semiparametric accelerated failure time model. The aim of the study is to relate the survival time ($t$) for multiple myeloma with a number of prognostic variables with censored data. The data present the survival times, in months, for 65 patients who were treated with alkylating agents, of the which 48 died during the study and 17 survived and also include measurements on each patient for the following five predictors: logarithm of a blood urea nitrogen measurement at diagnosis ($x_1$); hemoglobin measurement at diagnosis ($x_2$); age at diagnosis ($x_3$); sex ($x_4$, 0 for male and 1 for female) and serum calcium measurement at diagnosis ($x_5$).

Firstly, we consider the following regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i, \quad i = 1, \ldots, 65, \tag{18}$$

where $\varepsilon_i$ are independent r.v. such that $\varepsilon_i \sim SN(\alpha, 0, 2)$ and the response $y_i$ denotes the lifetime logarithm. The MLE of $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\mathrm{T}}, \alpha)^{\mathrm{T}} = (\beta_0, \beta_1, \ldots, \beta_5, \alpha)^{\mathrm{T}}$ are (estimated standard errors in parenthesis):

$$\hat{\beta}_0 = 4.500(1.282), \ \hat{\beta}_1 = -1.596(0.435), \ \hat{\beta}_2 = 0.142(0.050), \ \hat{\beta}_3 = 0.010(0.013),$$

$$\hat{\beta}_4 = 0.209(0.284), \quad \hat{\beta}_5 = -0.141(0.069) \quad \text{and} \quad \hat{\alpha} = 1.082(0.112).$$

We note that predictors $x_3$ and $x_4$ seem to be not significant at 10% in model (18).

In order to detect possible departures from the assumptions of log-BS errors in the model (18) as well as outlying and high leverage observations, we present in Fig. 2 the normal probability plots for $r^*_{\mathrm{DC}_i}$ and $r^*_{\mathrm{MD}_i}$ with generated envelopes and the index plot of $GL_{ii}$, in Fig. 3. As we can see, the graphs in Fig. 2 do not present unusual features so that the assumption of log-BS errors does not seem to be unsuitable. Also, no observation appears as possible outlier. In Fig. 3, observation 2 appears as a high leverage point. This observation corresponds to the youngest patient, 38 years

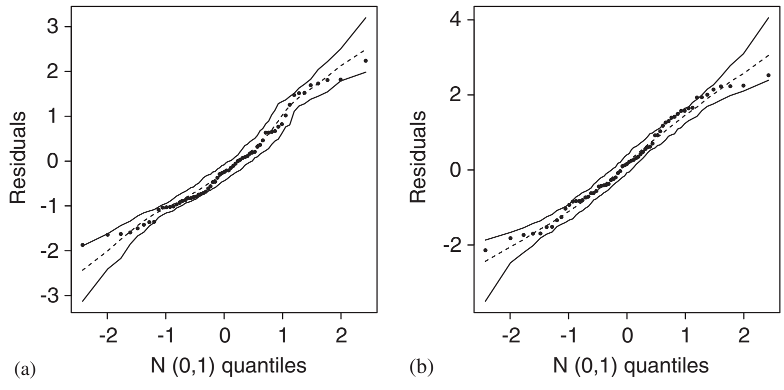


Fig. 2. Normal probability plots for: (a) $r^*_{\mathrm{DC}_i}$ and (b) $r^*_{\mathrm{MD}_i}$, with envelopes.

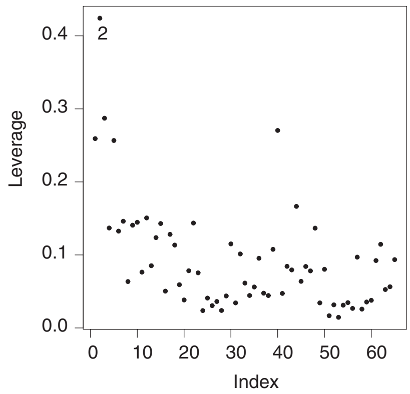


Fig. 3. Index plot of $GL_{ii}$.

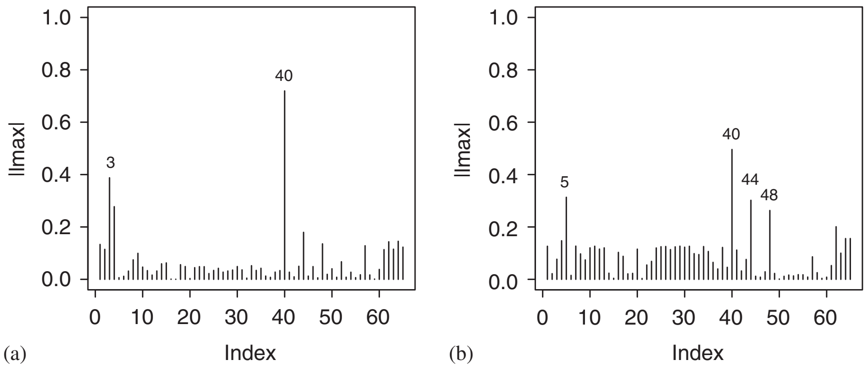


Fig. 4. Index plots of $|\boldsymbol{l}_{\max}|$ for: (a) $\boldsymbol{\theta}$ and (b) $\alpha$ under the case-weights perturbation.

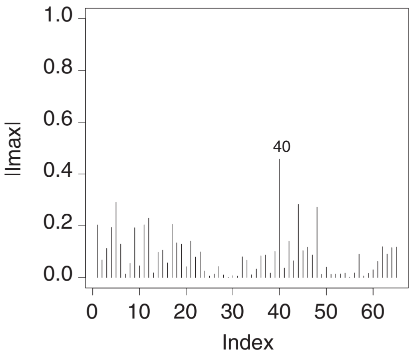


Fig. 5. Index plots of $|\boldsymbol{l}_{\max}|$ for $\boldsymbol{\theta}$ under the response perturbation.

old, male and just one month of survival since the treatment. His logarithm of blood urea nitrogen and hemoglobin measurements are larger than the average of the 65 patients and his serum calcium measurement is very high, the largest among the patients.

Figs. 4–7 show the index plots of $|\boldsymbol{l}_{\max}|$ for $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and $\alpha$, where $\boldsymbol{\theta} = (\boldsymbol{\theta}_{\mathbf{1}}^{\mathrm{T}}, \theta_2)^{\mathrm{T}}$, with $\boldsymbol{\theta}_{\mathbf{1}} = \boldsymbol{\beta} = (\beta_0, \ldots, \beta_5)^{\mathrm{T}}$ and $\theta_2 = \alpha$, under the perturbation scheme indicated. As we can see from these figures, observation 40 appears as the most

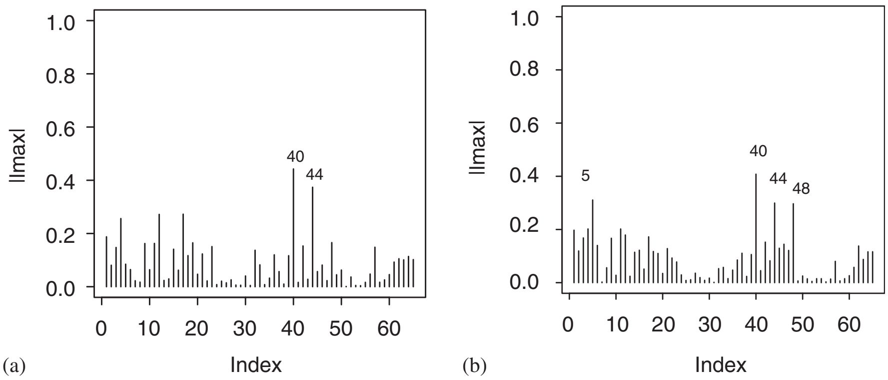


Fig. 6. Index plots of $|\boldsymbol{l}_{\max}|$ for: (a) $\boldsymbol{\theta}$ and (b) $\alpha$ under perturbation of $x_2$.

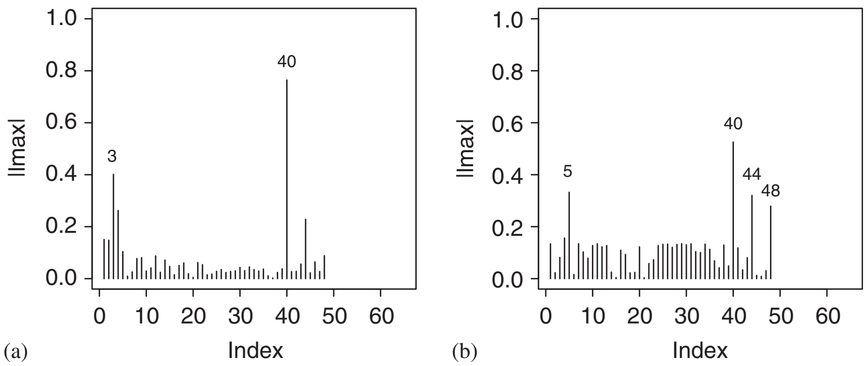


Fig. 7. Index plots of $|\boldsymbol{l}_{\max}|$ for: (a) $\boldsymbol{\theta}$ and (b) $\alpha$ under non-censoring perturbation.

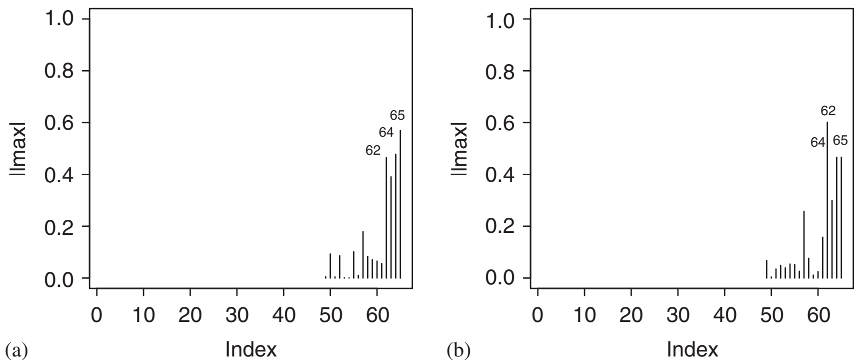


Fig. 8. Index plots of $|\boldsymbol{l}_{\max}|$ for (a) $\boldsymbol{\theta}$ and (b) $\alpha$ under censored case perturbation.

influential in all the graphs. This observation refers to a male patient, 74 years old and survival time since the treatment of 51 months. His logarithm of blood urea nitrogen and serum calcium measurements are larger than the average of the 65 patients and his hemoglobin measurement is below the average. Other observations appear with some outstanding influence on the parameter estimates. For instance, observations 3, 5, 44 and 48 appear as possible influential on $\hat{\boldsymbol{\beta}}$ and $\hat{\alpha}$ under the perturbation schemes used. When only the censoring cases are perturbed, we detected observations 62, 64 and 65 as possibly influential (see Fig. 8). Some index plots of $|\boldsymbol{l}_{\max}|$ for $\boldsymbol{\beta}$ and/or $\alpha$ were omitted, since these are similar to ones given in Figs. 4–8.

Table 1
Parameter estimates and the their $p$-values in parenthesis for the indicated sets

| Set | Parameter | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|---|---|---|---|---|---|---|---|
| A | Estimate | 4.500 | −1.596 | 0.142 | 0.010 | 0.209 | −0.141 |
| | | (0.000) | (0.000) | (0.004) | (0.455) | (0.461) | (0.042) |
| B | Estimate | 5.685 | −1.617 | 0.099 | −0.005 | 0.119 | −0.142 |
| | | (0.000) | (0.000) | (0.037) | (0.692) | (0.634) | (0.105) |
| E | Estimate | 4.598 | −1.580 | 0.129 | 0.009 | 0.243 | −0.135 |
| | | (0.001) | (0.001) | (0.020) | (0.521) | (0.435) | (0.067) |

Table 2
Relative changes (–RC– in %) and the corresponding $p$-values in parenthesis

| Set ($I$) | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ | $\hat{\beta}_5$ | $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|
| A−{case 2} | 4 | 6 | 1 | 54 | 3 | 27 | 0 |
| | (0.001) | (0.001) | (0.005) | (0.753) | (0.473) | (0.201) | |
| A−{case 3} | 15 | −3 | 3 | −62 | −4 | 30 | 1 |
| | (0.005) | (0.000) | (0.005) | (0.250) | (0.438) | (0.188) | |
| A−{case 5} | −20 | −3 | 25 | 42 | 68 | −9 | 4 |
| | (0.000) | (0.000) | (0.046) | (0.660) | (0.814) | (0.023) | |
| A−{case 40} | −14 | 0 | −14 | 71 | −12 | −34 | 6 |
| | (0.000) | (0.000) | (0.000) | (0.826) | (0.377) | (0.011) | |
| A−{case 44} | 12 | −1 | −14 | −22 | −46 | 12 | 3 |
| | (0.002) | (0.000) | (0.001) | (0.350) | (0.283) | (0.063) | |
| A−{case 48} | −5 | −2 | 5 | 4 | 52 | −7 | 3 |
| | (0.000) | (0.000) | (0.006) | (0.460) | (0.723) | (0.029) | |
| A−{case 62} | −5 | −6 | 9 | 23 | −7 | 11 | 2 |
| | (0.000) | (0.000) | (0.009) | (0.562) | (0.422) | (0.065) | |
| A−{case 64} | −2 | 3 | 5 | 14 | −9 | −1 | 1 |
| | (0.000) | (0.000) | (0.006) | (0.513) | (0.415) | (0.037) | |
| A−{case 65} | −2 | 6 | 8 | 5 | −9 | −6 | 1 |
| | (0.000) | (0.000) | (0.008) | (0.471) | (0.414) | (0.031) | |
| B | −26 | −1 | 30 | 148 | 43 | −1 | 24 |
| | (0.000) | (0.000) | (0.037) | (0.692) | (0.634) | (0.105) | |
| E | −2 | 1 | 9 | 7 | −16 | 4 | −3 |
| | (0.001) | (0.001) | (0.020) | (0.521) | (0.435) | (0.067) | |

The index plots of $C_i$ for $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and $\alpha$ under each perturbation scheme (omitted here) confirmed the outstanding influence of observations 3, 5, 40, 44, 48, 62, 64 and 65, which also were detected in the previous index plots of $|\boldsymbol{l}_{\max}|$.

Therefore, the diagnostic analysis detected as potentially influential the following nine cases: 2, 3, 5, 40, 44, 48, 62, 64 and 65. In order to reveal the impact of these nine observations on the parameter estimates, we refitted the model under some situations. First, we eliminated individually each one of these nine cases. Next, we removed from the "set $A$" (original data set) the totality of potentially influential observations (we called it "set $B$").

Table 1 presents the parameter estimates of the models under such situations. In Table 2, we have the relative changes (in percentage) of each parameter estimate, defined by $\mathrm{RC}_{\theta_j} = [(\hat{\theta}_j - \hat{\theta}_{j(I)})/\hat{\theta}_j] \times 100$, and the corresponding $p$-values, where $\hat{\theta}_{j(I)}$ denotes the MLE of $\theta_j$ after the "set $I$" of observations has been removed.

Following the same procedure suggested by Lee et al. (2006), we use the total relative changes: $\mathrm{TRC} = \sum_{j=1}^{7} |\mathrm{RC}_{\theta_j}|$, the maximum relative changes: $\mathrm{MRC} = \max_j |\mathrm{RC}_{\theta_j}|$ and the likelihood displacement: $\mathrm{LD}_I(\boldsymbol{\theta}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{(I)})\}$, where $\hat{\boldsymbol{\theta}}_{(I)}$ denotes the MLE of $\boldsymbol{\theta}$ after the "set $I$" of observations has been removed (see Cook et al., 1988). We find for the "set $B$" that $\mathrm{TRC} = 273.30$, $\mathrm{MRC} = 148.02$ and $\mathrm{LD}_{A-B}(\boldsymbol{\theta}) = 11.92$. In order to compare the impact of the influential observations, we repeat the analysis by removing the same number (nine observations) randomly selected from the non-influential observations (“set $E$”, formed by the observations of the “set $A$” less the non-influential observations randomly selected: 6, 14, 19, 30, 33, 41, 46, 54 and 63). In this case, we find that TRC = 41.74, MRC = 15.84 and $\mathrm{LD}_{A-E}(\boldsymbol{\theta}) = 0.19$. Hence, the results are more sensitive for the influential observation group.

As we can see from Table 1 the inferences do not change at the significance level of 10% when either set $A$ or set $B$ or set $E$ is considered. That is, the explanatory variables $x_3$ and $x_4$ should be removed from the model. However, looking at Table 2 we can notice that the elimination of observations 2 and 3 make the explanatory variable $x_5$ non-significant. The significance of this variable was masked by these observations, so it should be also removed from the model. Thus, the final model becomes given by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad i = 1, \ldots, 65. \tag{19}$$

The MLEs of the parameters (estimated standard errors in parenthesis) are

$$\hat{\beta}_0 = 4.409(0.772), \quad \hat{\beta}_1 = -1.869(0.423), \quad \hat{\beta}_2 = 0.109(0.049) \quad \text{and} \quad \hat{\alpha} = 1.140(0.118).$$

We may interpret the estimated coefficients of the final model as the following. The expected survival time should decrease approximately 85%$[(1-\mathrm{e}^{-1.869}) \times 100\%]$ as the logarithm of blood urea nitrogen measurement increases one unity, keeping the hemoglobin measurement fixed. Similarly, the expected survival time should increase approximately 12%$[\mathrm{e}^{0.109} \times 100\%]$ as the hemoglobin measurement increases one unity, keeping the logarithm of blood urea nitrogen measurement fixed.

## 6. Concluding remarks

The majority of the articles on log-BS regression models assume uncensored observations. In this work we present various results for the censored case. In particular, we derive the normal curvature of local influence under some perturbation schemes and we also present some ways to perform residual analysis. Two deviance-type residuals whose empirical distributions seem to be close to the standard normal distribution are proposed. A data set from the medical area is reanalyzed under log-BS regression models and the diagnostic analysis plays a very important role for selecting an appropriate model. The codes of the programs used for fitting log-BS regression models are available from the authors upon request.

## Acknowledgments

This study was supported partially by the research projects FONDECYT 1050862, FANDES C–13955/10 and DIPUV 42–2004, Chile; and CNPq, from Brasil. The authors are grateful for two anonymous referees for helpful comments and suggestions.

## References

Atkinson, A.C., 1981. Two graphical display for outlying and influential observations in regression. Biometrika 68, 13–20.
Barlow, W.E., Prentice, R.L., 1988. Residuals for relative risk regression. Biometrika 75, 65–74.
Birnbaum, Z.W., Saunders, S.C., 1969. A new family of life distributions. J. Appl. Probab. 6, 319–327.
Cook, R.D., 1986. Assessment of local influence (with discussion). J. Roy. Statist. Soc. B 48, 133–169.
Cook, R.D., Weisberg, S., 1982. Residuals and Influence in Regression. Chapman & Hall, London.
Cook, R.D., Peña, D., Weisberg, S., 1988. The likelihood displacement: a unifying principle for influence measures. Comm. Statist. Theory Methods 17, 623–640.
Davison, A.C., Gigli, A., 1989. Deviance residuals and normal scores plots. Biometrika 76, 211–221.
Desmond, A., 1985. Stochastic models of failure in random environments. Canad. J. Statist. 13, 171–183.
Díaz-García, J.A., Galea, M., Leiva, V., 2003. Influence diagnostics for elliptical multivariate linear regression models. Comm. Statist. Theory Methods 32, 625–641.
Escobar, L.A., Meeker, W.Q., 1992. Assessing influence in regression analysis with censored data. Biometrics 48, 507–528.
Fung, W.K., Kwan, C.W., 1997. A note on local influence based on the normal curvature. J. Roy. Statist. Soc. B 59, 839–843.
Galea, M., Paula, G.A., Bolfarine, H., 1997. Local influence in elliptical linear regression models. The Statistician 46, 71–79.
Galea, M., Riquelme, M., Paula, G.A., 2000. Diagnostics methods in elliptical linear regression models. Brazilian J. Probab. Statist. 14, 167–184.
Galea, M., Leiva, V., Paula, G.A., 2004. Influence diagnostics in log-Birnbaum–Saunders regression models. J. Appl. Statist. 31, 1049–1064.
Ghosh, S.K., Ghosal, S., 2006. Semiparametric accelerated failure time models for censored data. In: Upadhyay, S.K., Singh, U., Dey, D.K. (Eds.), Bayesian Statistics and its Applications. Anamaya Publishers, New Delhi, pp. 213–229.

Gradshteyn, I.S., Randzhik, I.M., 2000. Table of Integrals, Series, and Products. Academic Press, New York.
Jin, Z., Lin, D.Y., Wei, L.J., Ying, Z., 2003. Rank-based inference for the accelerated failure time model. Biometrika 90, 341–353.
Klein, J.P., Moeschberger, M.L., 1997. Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York.
Krall, J.V., Uthoff, V.A., Harley, J.B., 1975. A step-up procedure for selecting variables associated with survival. Biometrics 31, 49–57.
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., 1982. Statistical Models and Methods for Lifetime Data. Wiley, New York.
Lee, S.Y., Lu, B., Song, X.Y., 2006. Assessing local influence for nonlinear structural equation models with ignorable missing data. Comput. Stat. Data An. 50, 1356–1377.
Lesaffre, E., Verbeke, G., 1998. Local influence in linear mixed models. Biometrics 54, 570–582.
McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman & Hall, London.
Ortega, E,M., 2001. Influence analysis in generalized log-gamma regression models. Doctor Thesis, Department of Statistics, University of São Paulo, Brasil (in Portuguese).
Ortega, E.M., Bolfarine, H., Paula, G.A., 2003. Influence diagnostics in generalized log-gamma regression models. Comput. Stat. Data An. 42, 165–186.
Paula, G.A., 1993. Assessing local influence in restricted regression models. Comput. Stat. Data An. 16, 63–79.
Pettit, A.N., Bin Daud, I., 1989. Case-weight measures of influence for proportional hazards regression. Appl. Statist. 38, 51–67.
Press, W.H., Teulosky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in C: The Art of Scientific Computing. Prentice-Hall, London.
Rieck, J.R., 1989. Statistical analysis for the Birnbaum–Saunders fatigue life distribution, Unpublished Ph.D. Thesis, Clemson University, Department of Mathematical Sciences.
Rieck, J.R., Nedelman, J.R., 1991. A log-linear model for the Birnbaum–Saunders distribution. Technometrics 33, 51–60.
Tanaka, Y., Zhang, F., Mori, 2003. Local influence in principal component analysis: relationship between the local influence and influence function approaches revisited. Comput. Stat. Data An. 44, 143–160.
Therneau, T.M., Grambsch, P.M., Fleming, T.R., 1990. Martingale-based residuals for survival models. Biometrika 77, 147–160.
Verbeke, G., Molenberghs, G., 2000. Linear Mixed Models for Longitudinal Data. Springer, New York.
Wang, Z., Desmond, A.F., Lu, X., 2006. Modified censored moment estimation for the two-parameter Birnbaum–Saunders distribution. Comput. Stat. Data An. 50, 1033–1051.
Wei, B.C., Hu, Y.Q., Fung, W.K., 1998. Generalized leverage and its applications. Scand. J. Statist. 25, 25–37.
Williams, D.A., 1987. Residuals in generalized linear model diagnostic using the deviance and single case deletion. Appl. Statist. 36, 181–191.