Assessment of global and local influence in robust linear and nonlinear censored regression models

Isabel Cristina Gomes[a**], Lourdes C. Contreras Montenegro[a]
Victor Hugo Lachos[b], Marcos Oliveira Prates[a]

[a] Statistics Department,
Universidade Federal de Minas Gerais, Brazil,
[b] Statistics Department,
Universidade Estadual de Campinas, Brazil,

Abstract

The aim of this paper is to propose diagnostic methods to Normal/Independent (NI) nonlinear regression models for censored data. This class of models provides a useful generalization of normal nonlinear censored regression model since the error distribution cover heavy-tailed distributions such as Student-t, Slash, Contaminated Normal, among others. In this work, we develop global and local influence tools for nonlinear censored regression models with NI random errors to show that the non-normal distributions are more robust than the normal in presence of outliers. In order analyze the sensitivity of the maximum likelihood estimators of the parameters of model, we study the global and local influence methodology under the case-deleted and some perturbation schemes, such as case-weight, scale parameter, explanatory variable and coefficients of the model. We also present simulation studies that illustrate the behavior of diagnostic measures proposed. Finally, the results obtained from the analyzes of a real dataset is presented to illustrate the developed methodology.

Keywords: Censored regression model, EM algorithm, Local influence, Normal/Independent distributions.

1 Introduction

Nonlinear censored models with normally distributed random errors have received considerable contributions in different areas of surveys. In medical

**Corresponding author. Email: icgomes04@yahoo.com.br.

surveys the relationship between the survival time and age of some patient which has received certain treatment is often nonlinear. In this case the survival time is a right-censored variable because the patient can leave the study, die for other reason than the disease under study or for the end of the study. In engineering nonlinear censored models were applied in accelerated life tests studies which are considered right-censored for the end of the study (Heuchenne and Keilegom, 2007). The linear censored regression model can be seen as a particular case of the nonlinear model.

In the statistical literature, the censored models with normally distributed random errors are widely applied in several areas for modeling the symmetric data. However, it can be inappropriate when the data presents heavier tails than those expected under normality. An alternative to deal with outliers is the use of heavy-tailed distributions to represent the behavior of random errors. Some authors have adopted the Normal/Independent distributions (NI) (Lange and Sinsheimer, 1993) in robust regression models. From a Bayesian perspective and using this class of distributions, Liu (1996) studied the estimation of linear censored regression models, Rosa et al. (2003) estimated linear mixed models and Lachos et al. (2011) worked with linear and nonlinear censored mixed models. However, other authors have studied the nonlinear models with different distributions for the errors, Xie et al. (2009) estimated the parameter the model by EM algorithm considering skew-normal random errors and an AR(1) structure. Garay et al. (2011) and Labra et al. (2012) presented the EM estimation with skew-normal/independent random errors, where the last have considered heteroscedastic structure for the errors. Vanegas and Cysneiros (2010) used classical approach and NI distributions for the random errors. Wang (2007) and Meza et al. (2012) estimated a nonlinear mixed model, the first one through Monte Carlo EM methodology and second used EM algorithm and NI distributions for the random errors. Bayesian estimation was made by Cancho et al. (2011) with skew-normal/independent random errors and by Cancho et al. (2010), who also presented the classical approach and used skew-normal errors. The papers which has incorporated censored to estimation of nonlinear models was Heuchenne and Keilegom (2007), Matos et al. (2013) e Garay et al. (2015b).

A proper statistical anaylsis requires that when proposing a model to represent a particular event we should estimate it and present evidences to support the proposed model. A possible alternative is to use the goodness of fit techniques available in most statistical softwares, as parameter significance, coefficient determination and residual analysis (Montgomery and Peck, 1992). Moreover, the use of influence diagnostics methods allow to identify influential observations in the estimation process and evaluate the robustness of the estimates when the model or the data suffers small perturbations (Salgado, 2006; Cook and Weisberg, 1982). In the framework of influence diagnostics we can use global and/or local influence tools.

Global influence measures are methods which proposes the elimination of one or a set of observations and, next, analyze the estimates provided by the model with and without this data. Cook (1977) presented a measure to identify influential observations in linear regression models.

Recently, have been often discussed in the literature the local influence analysis, which goal is to evaluate the influence of small perturbations in the model or the data about the estimates provided by the model. Cook (1986) proposed a local influence analysis based on the change of the normal curvature of the neighborhood of a point where the postulated and perturbed models overlaps. Zhu and Lee (2001) improved the Cook method working with the conditional expectation of the log-likelihood function. Lachos (2002) studied diagnostic analysis of the Grubbs model. Ortega et al. (2003) worked with diagnostic analysis for censored regression models with Generalized Log-Gama distribution for the random errors. Zeller et al. (2010) proposed influence diagnostics for linear mixed models using distributions from the Skew-Normal/Independent family. Matos et al. (2013) analyzed influence diagnostics in linear and nonlinear censored models with Student-t distribution.

In this work we present the global and local influence analysis for nonlinear censored regression models with heavy-tailed distributed random errors. For this purpose we developed with the conditional expectation of the complete data log-likelihood function (Q function) to obtain the measures. To assess the global influence we use the Cook distance Cook (1977) and for the local influence we use the Zhu and Lee (2001) approach and analyze case-weight, scale, on one explanatory variable and in the coefficients perturbation schemes. The usefulness of our method is illustrated in simulation studies and in a real dataset analysis. We also present the influence measures and an application to the linear case.

The rest of this paper is organized as follows. A brief review of the NI distributions class is presented in Section 2. In Section 3 we show the nonlinear censored regression model with NI random errors and the EM algorithm was used for this estimation. The influence diagnostic analysis with global and local influence measures is presented in Section 4. The simulation study and real dataset application are given in Section 5 e 6. Finally, in Section 7 deals with some conclusions.

2 Normal/Independent Distributions

The Normal/Independent family of distributions includes symmetric distributions which extends the Normal case by the inclusion of kurtosis. In the last years this family has received increasing attention, because these distributions are more adequate to perform robust inference methods. Therefore, these distributions have been applied in different scientific areas to model data with atypical observations. A distribution of the NI class is defined as the distribution of a random variable (rv) Y given by

$$Y = \mu + \frac{Z}{\sqrt{U}}, \tag{1}$$

where U is positive rv independent of a rv Z with Normal distribution (mean zero and variance σ^2) and $\mu \in \mathbb{R}$ is a known constant. From now on this paper, this rv is denoted by $Y \sim NI\left(\mu, \sigma^2, \nu\right)$, where ν is the parameter indexing the distribution of U. The marginal probability density function (pdf) of Y is given by

$$f(y) = \int_0^\infty \frac{e^{-\frac{u}{2\sigma^2}(y-\mu)^2}\sqrt{u}}{\sqrt{2\pi\sigma^2}} dH(u), \quad y \in \mathbb{R},$$

where $H(u)$ is the cumulative distribution function (cdf) of U.

The Normal distribution is a particular case of this family and occurs when H is a degenerate probability mass function at one ($U = 1$ with probability 1). We can easily notice that for each properly choose of U we will have a different distribution. Some of the most used distributions of this family are presented are:

Pearson Type VII distribution: Occurs when $U \sim \text{Gamma}(\delta/2, \nu/2)$, with $\nu > 0$ and $\delta > 0$, where $\text{Gamma}(a, b)$ denote the Gamma distribution with a/b mean. In this particularly case, if $\delta = \nu$ we recover the Student-t distribution. More specifically, if $\delta = \nu = 1$ the Cauchy distribution is obtained.

Slash distribution: In this case $U \sim \text{Beta}(\nu, 1)$, with $\nu > 0$. When ν=1 we have the standard Slash distribution, and if $\nu \to \infty$ we have the Normal distribution Wang and Genton (2006).

Contaminated Normal distribution: For this distribution the mixture variable U is a discrete random variable assuming two values: $\lambda \in (0, 1)$ with probability ν and 1 with probability $1 - \nu$. In this case ν can be interpreted as the outliers proportion and λ interpreted as a scale factor.

3 Model specification

The Normal/Independent nonlinear regression model is represented by the following equation

$$Y_i \;=\; \eta(\mathbf{X}_i^\top, \boldsymbol{\beta}) + \epsilon_i, \quad i = 1, ..., n, \tag{2}$$

where $\epsilon_i \sim NI(0, \sigma^2, \nu)$, $\eta(\mathbf{X}_i^\top, \boldsymbol{\beta})$ is a nonlinear function of $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^\top$ that is twice continuously differentiable, Y_i is the ith response variable and $\mathbf{X}_i^\top = (x_{i1}, ..., x_{ip})$ is a vector of the values of p explanatory variables. By Eq. (1), we have that $Y_i \sim NI(\eta(\mathbf{X}_i^\top, \boldsymbol{\beta}), \sigma^2, \nu)$, for $i = 1, ..., n$. We call (2) the NI censored nonlinear regression (CNLNIR) model.

In the linear case, we have

$$Y_i = \mathbf{X}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad \epsilon_i \sim NI(0, \sigma^2, \nu), \quad i = 1, ..., n, \tag{3}$$

where Y_i is the ith response, $\boldsymbol{\beta}^\top = (\beta_1, ..., \beta_p)^\top$ is the parameter vector and $\mathbf{X}_i^\top = (X_{i1}, ..., X_{ip})$ is the covariable vector of the ith case. We call (3) the NI censored linear regression (CLNIR) model.

Without loss of gereneralilty, troughout the paper we present our diagnosis analysis from a right-censored regression framework, extension to left and interval cesoring are straight forward. In a right-cesoring setup, the observations can take the following values

$$Y_i^\star = \begin{cases} \kappa_i, & \text{if } Y_i \geq \kappa_i \\ Y_i, & \text{if } Y_i < \kappa_i \end{cases}$$

where $Y_i^\star$ is the observed response, Y_i is the true value and κ_i represent a known cut off point for $i = 1, ...n$.

3.1 Parameter estimation

The parameter vector of the NLNICR model is given by $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma^2)^T$. In this paper we assumed that the degrees of freedom ν is fixed and known as some authors have treated in their work, see, e.g., Lange et al. (1989); Meza et al. (2012). Meza et al. (2012) proposed fitting the model with several values of ν and choose the one which maximizes of log-likelihood function. Lange et al. (1989) concluded there is an increase in variance of the model when ν is estimated, comparing to a model in which ν is fixed and known. In our studies of simulation and in the real dataset application we consider the Meza et al. (2012) approach. It is important to notice that our main focus is to perform diagnosis analysis and only use the estimation methodology.

To estimate the parameters of the NLNICR model was used the methodology of the EM algorithm Dempster et al. (1977). Supposing that there are m censored observations of the response variable $\mathbf{Y}$, we can partition $\mathbf{Y}$ the following way, $\mathbf{Y} = (\kappa_1, ..., \kappa_m, Y_{m+1}, ..., Y_n)$, m censored values and $n - m$ uncensored values. Then the log-likelihood function for the parameters vector $\boldsymbol{\theta}$ can be expressed as

$$\begin{aligned} l(\boldsymbol{\theta}|\mathbf{Y}) &= \sum_{i=1}^{m} log\left[F\left(\frac{\eta(\mathbf{X}_i^\top, \boldsymbol{\beta}) - \kappa_i}{\sigma}\middle| 0, 1, \nu\right)\right] \\ &+ \sum_{j=m+1}^{n} log\left[f\left(y_j | \eta(\mathbf{X}_j^\top, \boldsymbol{\beta}), \sigma^2, \nu\right)\right], \end{aligned}$$

where $F(.|\mu, \sigma^2, \nu)$ denotes the cumulative distribution function (cdf) and $f(.|\mu, \sigma^2, \nu)$ probability density function (pdf) of the NI distribution with parameters (μ, σ^2, ν). We consider the censored observations Y_i as realizations of a latent random variable $\mathbf{Y}_L \sim NI(\eta(\mathbf{X}_i^\top, \boldsymbol{\beta}), \sigma^2, \nu)$. The augmented data vector is $\mathbf{Y}_c = (\mathbf{Y}, \mathbf{Y}_L, \mathbf{U})$. Then we can expressed the log-likelihood function as

$$\begin{aligned} l(\boldsymbol{\theta}|\mathbf{Y}_c) &= \frac{1}{2}\sum_{i=1}^{n} log(u_i) - \frac{n}{2}log(2\pi) - \frac{n}{2}log(\sigma^2) + \sum_{i=1}^{n} log\ h(u_i|\nu) \\ &- \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(u_i y_i^2 - 2u_i y_i \eta(\mathbf{X}_i^\top, \boldsymbol{\beta}) + u_i \eta(\mathbf{X}_i^\top, \boldsymbol{\beta})^\top \eta(\mathbf{X}_i^\top, \boldsymbol{\beta})\right), \end{aligned} \tag{4}$$

where $h(u|\nu)$ denote the pdf of the rv U. In the following, the superscript '(k)'

indicates the estimate of the parameter vector in the k stage of the algorithm. For the E step, we calculate the Q function expressed as

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = \mathbb{E}_{\boldsymbol{\theta}^{(k)}}\left[l(\boldsymbol{\theta}|\mathbf{Y}_c)|\mathbf{Y}\right],$$

where $\mathbb{E}_{\boldsymbol{\theta}^{(k)}}$ denotes the expectation using the $\boldsymbol{\theta}^{(k)}$ estimate to $\boldsymbol{\theta}$. The above expression is completely determined by the following expectations

$$\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)}) = \mathbb{E}_{\boldsymbol{\theta}^{(k)}}\left[U_i Y_i^s|Y_i\right], \quad s = 0, 1, 2, \quad i = 1, ..., n, \tag{5}$$

since $\mathbb{E}_{\boldsymbol{\theta}^{(k)}}\left[log\ U_i|Y_i\right]$ and $\mathbb{E}_{\boldsymbol{\theta}^{(k)}}\left[log\ h(u_i|\nu)|Y_i\right]$ only depends of ν, known, and s is the powers of Y_i in equation (4). Then we can rewrite the Q function as

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) &= \varsigma - \frac{n}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\Big[\mathcal{E}_{2i}(\boldsymbol{\theta}^{(k)}) - 2\eta(\mathbf{X}_i^\top, \boldsymbol{\beta})\mathcal{E}_{1i}(\boldsymbol{\theta}^{(k)}) \\ &\quad + \eta(\mathbf{X}_i^\top, \boldsymbol{\beta})^\top \eta(\mathbf{X}_i^\top \boldsymbol{\beta})\mathcal{E}_{0i}(\boldsymbol{\theta}^{(k)})\Big] \end{aligned}$$

where ς is a constant independent of $\boldsymbol{\theta}$.

For a uncensored observation we have $Y_i \sim NI(\eta(\mathbf{X}_i^\top, \boldsymbol{\beta}), \sigma^2, \nu)$ and, therefore, $\mathcal{E}_{si}(\boldsymbol{\theta}^{(k)}) = y_i^s \mathbb{E}_{\boldsymbol{\theta}^{(k)}}(U_i|Y_i)$, so we can find expression to $\mathbb{E}_{\boldsymbol{\theta}^{(k)}}(U_i|Y_i)$ for the distributions in the NI family. For the Student-t, Slash and Contaminated Normal distributions the expressions were obtained by Garay et al. (2015b).

For a censored observation $Y_i = \kappa_i$ if $Y_i \geq \kappa_i$, that is, $Y_i \in (\kappa_i, \infty)$, $i = 1, ..., n$. The expressions for the conditional expectations (5) were obtained by Garay et al. (2015b).

In the M step of the algorithm we need to maximize $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ over to the parameters $\boldsymbol{\beta}$ and σ^2. The leading expressions are:

$$\begin{aligned} \boldsymbol{\beta}^{(k+1)} &= \underset{\boldsymbol{\beta}}{\operatorname{argmin}}(\boldsymbol{\tau}^{(k)} - \boldsymbol{\mu})^\top \widehat{\boldsymbol{\mathcal{E}}}_0^{(k)}(\boldsymbol{\tau}^{(k)} - \boldsymbol{\mu}), \\ \sigma^{2^{(k+1)}} &= \frac{1}{n}\sum_{i=1}^{n}\Big[\mathcal{E}_{2i}(\boldsymbol{\theta}^{(k)}) - 2\mu_i\mathcal{E}_{1i}(\boldsymbol{\theta}^{(k)}) + \mu_i^\top \mu_i \mathcal{E}_{0i}(\boldsymbol{\theta}^{(k)})\Big], \end{aligned}$$

where $\widehat{\boldsymbol{\mathcal{E}}}_0^{(k)} = \operatorname{diag}(\widehat{\mathcal{E}}_{01}^{(k)}, ..., \widehat{\mathcal{E}}_{0n}^{(k)})^\top$, $\boldsymbol{\mu} = (\mu_1, ..., \mu_n)^\top$, $\mu_i = \eta(\mathbf{X}_i, \boldsymbol{\beta})$, and $\boldsymbol{\tau}^{(k)} = \left(\tau_1^{(k)}, ..., \tau_n^{(k)}\right)^\top$ represents the corrected observed response $\tau_i^{(k)} = \mathcal{E}_{1i}(\boldsymbol{\theta}^{(k)})/\mathcal{E}_{0i}(\boldsymbol{\theta}^{(k)})$. For more details and further calculations, see Garay et al. (2015b). The full development of the estimation process in the linear case can be seen in Garay et al. (2015a).

4 Diagnostic analysis

Diagnostic techniques have been proposed to detected observations that impact on model fitting. These techniques can be considered in two approaches: case-deletion measures Cook (1977) and local influence Cook (1986); Cook and Weisberg (1982).

4.1 Global influence

The case-deletion analysis is a the most used technique of global influence and study the effect of dropping the ith case from a dataset. Let $\mathbf{Y}_{c[i]}$ be the augmented data of the $l\left(\boldsymbol{\theta}|\mathbf{Y}_{c[i]}\right)$ with the complete-data log-likelihood function without the ith case and $\widehat{\boldsymbol{\theta}}_{[i]} = \left(\widehat{\boldsymbol{\beta}}^{\top}_{[i]}, \widehat{\sigma^2}_{[i]}\right)^{\top}$ the EM estimate for $\boldsymbol{\theta}$ in that case, assuming ν as a fixed and known constant.

To assess the influence of the ith observation over $\widehat{\boldsymbol{\theta}}$ we calculate the difference between $\widehat{\boldsymbol{\theta}}_{[i]}$ and $\widehat{\boldsymbol{\theta}}$. If the $\widehat{\boldsymbol{\theta}}_{[i]}$ is considered far from $\widehat{\boldsymbol{\theta}}$, then the ith case is considered influential.

In order to reduce the computational efforts in the $\widehat{\boldsymbol{\theta}}_{[i]}$ calculation, Zhu et al. (2001) proposed to use the Q function instead of the log-likelihood function. To obtain $\widehat{\boldsymbol{\theta}}_{[i]}$ we used

$$\widehat{\boldsymbol{\theta}}^{1}_{[i]} = \widehat{\boldsymbol{\theta}} + \left[-\ddot{Q}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})\right]^{-1} \dot{Q}_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}), \tag{6}$$

where

$$\ddot{Q}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\top}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} \quad \text{and} \quad \dot{Q}_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \frac{\partial Q_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}},$$

denotes the Hessian matrix and gradient vector of Q function, respectively. The quantity $\widehat{\boldsymbol{\theta}}^{1}_{[i]}$ is called one step approximation of the estimated parameters vector without the ith observation.

The adaptation of the Cook distance between $\widehat{\boldsymbol{\theta}}_{[i]}$ and $\widehat{\boldsymbol{\theta}}$ proposed by Cook and Weisberg (1982) is

$$GD_i = \left(\widehat{\boldsymbol{\theta}}_{[i]} - \widehat{\boldsymbol{\theta}}\right)^{\top} \left[-\ddot{Q}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})\right] \left(\widehat{\boldsymbol{\theta}}_{[i]} - \widehat{\boldsymbol{\theta}}\right), \; i = 1, ..., n. \tag{7}$$

Replacing the expression (6) in equation (7), we have

$$GD_i^1 = \dot{Q}_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})^\top \left[-\ddot{Q}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})\right]^{-1} \dot{Q}_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}).$$

An observation should be regarded as influential if GD_i is greater than $(p+1)/n$, where p is dimension of the $\boldsymbol{\beta}$ vector and n is the sample size (Massuia et al., 2014). Note that in this method which will vary from model to model is the gradient vector and the Hessian matrix.

For the nonlinear case the gradient vector is

$$\begin{aligned}
\frac{\partial Q_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} &= \frac{1}{\widehat{\sigma}^2}\sum_{i\neq j}\left[\mathcal{E}_{1j}(\widehat{\boldsymbol{\theta}})d\mu_j - \mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\mu_j d\mu_j\right], \text{ and} \\
\frac{\partial Q_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \sigma^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} &= -\frac{n}{2\widehat{\sigma}^2} + \frac{1}{2\widehat{\sigma}^4}\sum_{i\neq j}\left[\mathcal{E}_{2j}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{1j}(\widehat{\boldsymbol{\theta}})\mu_j + \mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\mu_j^2\right].
\end{aligned}$$

And the Hessian matrix is

$$\begin{aligned}
\ddot{Q}_{\boldsymbol{\beta}}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= \frac{1}{\widehat{\sigma}^2}\sum_{i=1}^{n}\left[\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}})D\mu_i - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\left(d\mu_i d\mu_i^\top + \mu_i \mathbf{1}_p^\top D\mu_i\right)\right], \\
\ddot{Q}_{\boldsymbol{\beta}\sigma^2}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{\widehat{\sigma}^4}\sum_{i=1}^{n}\left[\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}})d\mu_i - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\mu_i d\mu_i\right], \text{ and} \\
\ddot{Q}_{\sigma^2}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= \frac{n}{2\widehat{\sigma}^4} - \frac{1}{\widehat{\sigma}^6}\sum_{i=1}^{n}\left[\mathcal{E}_{2i}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}})\mu_i + \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\mu_i^2\right],
\end{aligned}$$

where $\mathbf{1}_p^\top$ is a vector $p \times 1$ with all entries equal to 1, $\mu_i = \eta(\mathbf{X}_i, \widehat{\boldsymbol{\beta}})$, $d\mu_i = \dfrac{\partial \eta(\mathbf{X}_i, \widehat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}}$, $D\mu_i = \dfrac{\partial^2 \eta(\mathbf{X}_i, \widehat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^\top}$.

And in the linear case, the gradient vector is

$$\begin{aligned}
\frac{\partial Q_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} &= \frac{1}{\widehat{\sigma}^2}\sum_{i\neq j}\left[\mathbf{X}_j\mathcal{E}_{1j}(\widehat{\boldsymbol{\theta}}) - \mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\mathbf{X}_j\mathbf{X}_j^\top\widehat{\boldsymbol{\beta}}\right], \text{ and} \\
\frac{\partial Q_{[i]}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \sigma^2}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} &= -\frac{1}{2\widehat{\sigma}^2}\sum_{i\neq j}\Bigg[1 - \frac{1}{\widehat{\sigma}^2}\Big(\mathcal{E}_{2j}(\widehat{\boldsymbol{\theta}}) - 2(\mathbf{X}_j^\top\widehat{\boldsymbol{\beta}})\mathcal{E}_{1j}(\widehat{\boldsymbol{\theta}}) \\
&\qquad + \mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\left(\mathbf{X}_j^T\widehat{\boldsymbol{\beta}}\right)^T\left(\mathbf{X}_j^T\widehat{\boldsymbol{\beta}}\right)\Big)\Bigg].
\end{aligned}$$

And the Hessian matrix is

$$
\begin{aligned}
\ddot{Q}_{\boldsymbol{\beta}}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{n}\mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\mathbf{X}_j^{\top}\mathbf{X}_j, \\
\ddot{Q}_{\sigma^2}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= \frac{1}{2\widehat{\sigma^4}}\sum_{i=1}^{n}\left[1-\frac{2}{\widehat{\sigma^2}}\left(\mathcal{E}_{2j}(\widehat{\boldsymbol{\theta}})-2(\mathbf{X}_j^{\top}\widehat{\boldsymbol{\beta}})\mathcal{E}_{1j}(\widehat{\boldsymbol{\theta}})\right.\right. \\
&\qquad \left.\left.+\mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\left(\mathbf{X}_j^{T}\widehat{\boldsymbol{\beta}}\right)^{\top}\left(\mathbf{X}_j^{\top}\widehat{\boldsymbol{\beta}}\right)\right)\right], \text{ and} \\
\ddot{Q}_{\boldsymbol{\beta}\sigma^2}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= -\frac{1}{\widehat{\sigma^4}}\sum_{i=1}^{n}\left[\mathbf{X}_j\mathcal{E}_{1j}(\widehat{\boldsymbol{\theta}})-\mathcal{E}_{0j}(\widehat{\boldsymbol{\theta}})\mathbf{X}_j\mathbf{X}_j^{\top}\widehat{\boldsymbol{\beta}}\right].
\end{aligned}
$$

4.2 Local influence

Let $\boldsymbol{\omega} = (\omega_1, ..., \omega_g)^{\top}$ be a perturbation vector varying in an open region $\boldsymbol{\Omega} \subset \mathbb{R}^g$ and let $l\left(\boldsymbol{\theta}, \boldsymbol{\omega}|\mathbf{Y}_c\right)$ the complete-data log-likelihood function of the perturbed model. There is a vector $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$ such that $l\left(\boldsymbol{\theta}, \boldsymbol{\omega}_0|\mathbf{Y}_c\right) = l\left(\boldsymbol{\theta}|\mathbf{Y}_c\right)\ \forall\ \boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^{p+1}$. The influence graph is defined as $\boldsymbol{\alpha}(\boldsymbol{\omega}) = \left(\boldsymbol{\omega}^{\top}, f_Q(\boldsymbol{\omega})\right)^{\top}$, where $f_Q(\boldsymbol{\omega})$ is the Q-displacement function:

$$
f_Q(\boldsymbol{\omega}) = 2\left[Q(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) - Q(\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})|\widehat{\boldsymbol{\theta}})\right],
$$

where $\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ is the estimate of $\boldsymbol{\theta}$ which maximizes the Q-function of the perturbed model, $Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})$. The $f_Q(\boldsymbol{\omega})$ function can be seen as a measure of difference between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})$. We will use the normal curvature $C_{f_Q,\mathbf{d}}$ of $\boldsymbol{\alpha}(\boldsymbol{\omega})$ at $\boldsymbol{\omega}_0$ in the direction of some unit vector $\mathbf{d}$ to describe the local behavior of the Q-displacement function.

Zhu and Lee (2001) showed that

$$
C_{f_Q,\mathbf{d}} = -2\mathbf{d}^{\top}\ddot{Q}_{\boldsymbol{\omega}_0}\mathbf{d} \text{ and } -\ddot{Q}_{\boldsymbol{\omega}_0} = \boldsymbol{\Delta}_{\boldsymbol{\omega}_0}^{\top}\left[-\ddot{Q}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})\right]^{-1}\boldsymbol{\Delta}_{\boldsymbol{\omega}_0},
$$

where

$$
\boldsymbol{\Delta}_{\boldsymbol{\omega}} = \left.\frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^{\top}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})} \quad \text{and} \quad \ddot{Q}_{\boldsymbol{\theta}}(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \left.\frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\top}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}.
$$

The information provided by the symmetric matrix $-2\ddot{Q}_{\boldsymbol{\omega}_0}$ is used to detect influential observations. For this, we consider the spectral decomposition

$$-2\ddot{Q}_{\boldsymbol{\omega}_0} = \sum_{k=1}^{g} \zeta_k \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T,$$

where $\{(\zeta_k, \boldsymbol{\epsilon}_k),\ k = 1, .., g\}$ are eigenvalue-eigenvector pairs of $-2\ddot{Q}_{\boldsymbol{\omega}_0}$, with $\zeta_1 \geq ... \geq \zeta_p > \zeta_{p+1} = ... = \zeta_g = 0$, and eigenvectors ϵ_k, $k = 1, ..., g$. Let

$$\tilde{\zeta}_k = \frac{\zeta_k}{\zeta_1 + ... + \zeta_g},\ \boldsymbol{\epsilon}_k^2 = (\epsilon_{k1}^2, ..., \epsilon_{kg}^2)^\top,\ \text{and}\ M(0) = \sum_{k=1}^{g} \tilde{\zeta}_k \boldsymbol{\epsilon}_k^2.$$

The lth component of $M(0)$ is $M(0)_l = \sum_{k=1}^{g} \tilde{\zeta}_k \epsilon_{kl}^2$. The evaluation of influential observations is based on visual inspection of $M(0)_l$, $l = 1, ..., g$, plotted against the index l. The lth case is said to be influential if $M(0)_l$ is greater than a benchmark value. Based on Poon and Poon (1999), Zhu and Lee (2001) proposed to use the conformal normal curvature, given by

$$B_{f_q,\mathbf{d}} = \frac{C_{f_Q,\mathbf{d}}}{\text{tr}\left[-2\ddot{Q}_{\boldsymbol{\omega}_0}\right]}. \tag{8}$$

The computation of the expression (8) is quite simple and it is such that $0 \leq B_{f_Q,\mathbf{d}} \leq 1$. Fix $\mathbf{d}_l$ to be a vector with the lth entry equals to 1 and all others entries equal to 0. Zhu and Lee (2001) showed that $\forall l$, $M(0)_l = B_{f_Q,\mathbf{d}_l}$. Therefore we are able to obtain $M(0)_l$ through $B_{f_Q,\mathbf{d}_l}$. The benchmark value proposed by Lee and Xu (2004) to $M(0)$ is $\overline{M(0)} + c^* SM(0)$, where $\overline{M(0)}$ and $SM(0)$ be the mean and standard error of $M(0)$, respectively, and c^* is a constant. In this paper we use $c^* = 3.5$, as made by Massuia et al. (2014).

We will evaluate the matrix $\boldsymbol{\Delta}_{\boldsymbol{\omega}_0}$ under the following perturbation schemes: of case weight, on the scale parameter, in an explanatory variable and in the coefficients of the model.

4.2.1 Perturbation schemes

In this Section we will detail the building of the matrix $\boldsymbol{\Delta}_{\boldsymbol{\omega}_0}$ under the case weight and on the scale parameter perturbations. For each perturbation scheme, we use partitioned form $\boldsymbol{\Delta}_{\boldsymbol{\omega}} = \left(\boldsymbol{\Delta}_{\beta}^{\top}, \boldsymbol{\Delta}_{\sigma^2}^{\top}\right)^{\top}$, where

$$\boldsymbol{\Delta}_{\beta} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\omega}^{\top}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})} \quad \text{and} \quad \boldsymbol{\Delta}_{\sigma^2} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial \sigma^2 \partial \boldsymbol{\omega}^{\top}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})}.$$

In the nonlinear case we have

Case weight perturbation: Consider the weights to the expected values of the complete-data log-likelihood function. In this case, let $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^{\top}$ and $\boldsymbol{\omega}_0 = (1, ..., 1)^{\top} = \mathbf{1}_n^{\top}$. So

$$\begin{aligned} \boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{1}{\widehat{\sigma^2}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}}) d\mu_i - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}}) \mu_i d\mu_i\right], \text{ and} \\ \boldsymbol{\Delta}_{\sigma^2} &= -\frac{n}{2\widehat{\sigma}^2} - \frac{1}{2\widehat{\sigma}^4} \sum_{i=1}^{n} \left[\mathcal{E}_{2i}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}}) \mu_i + \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}}) \mu_i^2\right]. \end{aligned}$$

Scale perturbation: To study the effect of perturbing the assumption over σ^2 we consider $\sigma^2(\omega_i) = \omega_i^{-1}\sigma^2$, $\omega_i > 0$, $i = 1, ..., n$. Under this scheme, the non-perturbed model occurs when $\boldsymbol{\omega}_0 = \mathbf{1}_n^{\top}$. Therefore, the matrix $\boldsymbol{\Delta}_{\boldsymbol{\omega}}$ is

$$\begin{aligned} \boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{1}{\widehat{\sigma^2}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}}) d\mu_i - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}}) \mu_i d\mu_i\right], \text{ and} \\ \boldsymbol{\Delta}_{\sigma^2} &= \frac{1}{2\widehat{\sigma}^4} \sum_{i=1}^{n} \left[\mathcal{E}_{2i}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}}) \mu_i + \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}}) \mu_i^2\right]. \end{aligned}$$

For the explanatory variable and coefficients perturbation schemes, the contamination is made inside of the nonlinear function $\eta(\mathbf{X}_i, \boldsymbol{\beta})$. Thus we have to calculate the entries of the $\boldsymbol{\Delta}_{\boldsymbol{\omega}}$ matrix for each nonlinear function. The expressions obtained for the application is presented in Appendix.

For the linear case we have

Case-weight perturbation: Consider the weights to the expected values of the complete-data log-likelihood. In this case, let $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^{\top}$ and $\boldsymbol{\omega}_0 = (1, ..., 1)^{\top} = \mathbf{1}_n^{\top}$. Such that

$$\begin{aligned} \boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{1}{\widehat{\sigma^2}} \left[\mathbf{X}^{\top} diag\left[\mathcal{E}_1(\widehat{\theta})\right] - \mathbf{A}\right], \text{ and} \\ \boldsymbol{\Delta}_{\sigma^2} &= -\frac{1}{2\widehat{\sigma^2}} \left[\mathbf{1}_n^{\top} - \frac{1}{\widehat{\sigma^2}} \mathbf{B}^{\top}\right], \end{aligned}$$

where $diag(\mathbf{W})$ denotes the diagonal of matrix $\mathbf{W}$, $\mathbf{A}$ is a matrix defined as $\mathbf{X}^\top\mathbf{X}\widehat{\boldsymbol{\beta}}\mathcal{E}_0(\widehat{\theta})^\top$, $\mathcal{E}_i(\widehat{\boldsymbol{\theta}}) = (\mathcal{E}_{i1}(\widehat{\boldsymbol{\theta}}), \ldots, \mathcal{E}_{in}(\widehat{\boldsymbol{\theta}}))^\top$, $i = 0, 1, 2$, $\mathbf{X}$ is the design matrix and $\mathbf{B}$ is an n-dimensional vector with coordinates $B_i = \mathcal{E}_{2i}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}})\mathbf{X}_i^\top\widehat{\boldsymbol{\beta}} + \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})(\mathbf{X}_i^\top\widehat{\boldsymbol{\beta}})^\top(\mathbf{X}_i^\top\widehat{\boldsymbol{\beta}})$, $i = 1, ..., n$.

Scale perturbation: To study the effect of perturbing over σ^2 we consider $\sigma^2(\omega_i) = \omega_i^{-1}\sigma^2$, $i = 1, ..., n$. Under this scheme, the non-perturbed model occurs when $\boldsymbol{\omega}_0 = \mathbf{1}_n^\top$. Therefore, the matrix $\boldsymbol{\Delta}_{\boldsymbol{\omega}}$ is

$$\begin{aligned}
\boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{1}{\widehat{\sigma^2}}\left[\mathbf{X}^\top diag\left[\mathcal{E}_1(\widehat{\theta})\right] - \mathbf{A}\right], \text{ and} \\
\boldsymbol{\Delta}_{\sigma^2} &= -\frac{1}{2\widehat{\sigma^4}}\mathbf{B}^\top.
\end{aligned}$$

Explanatory variable perturbation: In this case we will perturb a continue explanatory variable $x_{(,t)}(\boldsymbol{\omega}) = x_{(,t)} + \boldsymbol{\omega}^\top$, where $x_{(,t)} \in \mathbb{R}^n$ is the tth column of matrix $\mathbf{X}$ and $\boldsymbol{\omega} \in \mathbb{R}^n$. Then, each line of the design matrix will be of the form $\mathbf{X}_i(\boldsymbol{\omega})^\top = (x_{i1}, ..., x_{it} + \omega_i, ..., x_{ip}) = \mathbf{X}_i^\top + \omega_i\mathbf{c_t}^\top$, where $\mathbf{c_t}$ denotes a vector $p \times 1$ with the pth entry equals to 1 and all others entries equal to equal to 0. As a result, the additive perturbation have $\boldsymbol{\omega}_0 = 0$ when $t = 0$. To study the local influence we will replace $\mathbf{X}_i(\boldsymbol{\omega})^\top = \mathbf{X}_i^\top + \omega_i\mathbf{c_t}^\top$ in Q-function. This way we have

$$\begin{aligned}
\boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{1}{\widehat{\sigma^2}}\left[\mathbf{c_t}\mathcal{E}_1(\widehat{\boldsymbol{\theta}}) - 2\mathbf{c_t}\widehat{\boldsymbol{\beta}}^\top\mathbf{X}^\top diag\left(\mathcal{E}_0(\widehat{\boldsymbol{\theta}})\right) - 2\mathbf{c_t}\widehat{\boldsymbol{\beta}}^\top\mathbf{c_t}\boldsymbol{\omega}^\top\, diag\left(\mathcal{E}_0(\widehat{\boldsymbol{\theta}})\right)\right], \\
\boldsymbol{\Delta}_{\sigma^2} &= \frac{1}{\widehat{\sigma^4}}\left[\mathbf{c_t}^\top\widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\beta}}^\top\mathbf{X}^\top diag\left(\mathcal{E}_0(\widehat{\boldsymbol{\theta}})\right) + \mathbf{c_t}^\top\widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\beta}}^\top\mathbf{c_t}\boldsymbol{\omega}^\top diag\left(\mathcal{E}_0(\widehat{\boldsymbol{\theta}})\right)\right. \\
&\quad \left. -\mathbf{c_t}^\top\widehat{\boldsymbol{\beta}}\mathcal{E}_1(\widehat{\boldsymbol{\theta}})\right].
\end{aligned}$$

Coefficients perturbation: The perturbation in $\boldsymbol{\beta}$'s is introduced replacing $\boldsymbol{\beta}$ for $\boldsymbol{\beta}(\omega) = \boldsymbol{\beta}\omega_i$, with $i = 1, ..., n$ and $\boldsymbol{\omega} \in \mathbb{R}^n$ in the Q-function. Then,

$$\begin{aligned}
\boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{n}\left[\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}})\mathbf{X}_i^\top\omega_i - \mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\omega_i^2\mathbf{X}_i\mathbf{X}_i^\top\widehat{\boldsymbol{\beta}}\right], \text{ and} \\
\boldsymbol{\Delta}_{\sigma^2} &= \frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{n}\left[\mathbf{X}_i^\top\mathcal{E}_{1i}(\widehat{\boldsymbol{\theta}}) - 2\mathcal{E}_{0i}(\widehat{\boldsymbol{\theta}})\omega_i\mathbf{X}_i\mathbf{X}_i^\top\widehat{\boldsymbol{\beta}}\right].
\end{aligned}$$

5 Simulation studies

Some Monte Carlo simulations have been developed to compare the performance of the estimates on nonlinear censored models in the presence of outliers on the response variable. We consider the nonlinear growth-curve model

$$Y_i = \frac{\beta_1}{1 + exp(\beta_2 + \beta_3 x_i))} + \epsilon_i, \ i = 1, ..., n, \tag{9}$$

where $n = 100$, $\epsilon_i \sim N(0, \sigma^2)$ to $i = 2, 3, ...99$, $\epsilon_1 = -5$ and $\epsilon_{100} = 5$, $x_i \sim$ Unif(10; 20), $\beta_1 = 330$, $\beta_2 = 6, 5$, $\beta_3 = -0, 7$ and $\sigma^2 = 1$. We estimate the model (9) with 10%, 20% and 40% of censure. We observed that for the different censure scenarios, the diagnosis results obtained were very similar and the difference was observed in the estimated standard deviations of the censored regression. For this reason, from now on we focus and present the results for the simulation study with 20% of censure. The analyzed distributions were Normal, Student-t ($\nu = 3$), Slash ($\nu = 2$) and Contaminated Normal ($\nu = (0.2, 0.5)$). The choice of the ν values for the non-normal models were to maximize the log-likelihood function. The construction of errors determined the cases #1 and #100 as perturbed. To generate the censored observations we followed the propose made by Tsuyuguchi (2012). We generate the proposed model (9) and we define the censoring level κ_i as the rth value of the ordered vector $\mathbf{Y}$, that is, $\kappa_i = Y_{(r)}$, $\forall i = 1, ..., r$ where $r = n - n \times pc$ is the number of censored observations and pc is the percentage of censorship. Thus $Y_{(r)}$ becomes a censored observation and all values of $\mathbf{Y}$ greater or equal to $Y_{(r)}$ will have the same value of it. The initial values for the parameters were obtained by the *nls* function of the package *stats*, in software R 3.0.2 version (R Core Team, 2013).

We propose to use as the perturbation vector $\boldsymbol{\omega}$ the absolute value of the ordinary residuals (AOR)

$$r_i = \left| y_i - \frac{\hat{\beta}_1}{1 + exp(\hat{\beta}_2 + \hat{\beta}_3 x_i))} \right|, \ i = 1, ..., n, \tag{10}$$

where $\hat{\beta}_i$, $i = 1, 2, 3$ are the estimators of $\boldsymbol{\beta}$'s. We performed a Monte Carlo study with 1.000 replicates of the proposed models to assess the percentage of replications in which the contaminated observations was taken as influential, as well as the mean and standard deviation (SD) of the influence measures in these cases for the influence techniques we presented in this paper (see Table 1).

5.1 Global influence

The main interest of this study is evaluate the sensibility of the proposed model to atypical observations by the global influence techniques. We will present only the graphics and Monte Carlo study for the global estimation case, since the results obtained by analysis of $\boldsymbol{\beta}$'s and σ were quite similar.

We observed a considerable difference of the Cook distance between the Normal and non-normal models (see the first four lines in Table 1). We can see that the contaminated observations were influential for all the Monte Carlo replications in the Normal model. These cases, however, were not influential for any replications for Student-t and Slash models. For the Contaminated Normal model it was influential for some replications, nevertheless with measures substantially smaller than Normal model. The Figure 1 presents the generalized Cook distance for the four proposed models in one of the replications. We can see the Normal model was highly influenced by these observations, and the heavy-tailed models seems to be able to accommodate the effects of these cases.

5.2 Local influence

In this Subsection the main idea is evaluate the robustness of the proposed models to the outliers by local influence methods. In the Monte Carlo study (see Table 1) for the case-weight perturbation, we have noticed that the influence measures were smaller in the non-normal models and the perturbed observations were influential for almost all the replications to Normal model. In the Student-t and Slash models these cases were not influential in any replication and for the Contaminated Normal model it was influential for some replications, but with measure a somewhat greater than the benchmark value.

We observed a similar behavior in the scale perturbation scheme. In this case, however, the case #100 was not influential for any replications for the Contaminated Normal model. In the explanatory variable perturbation scheme these observations were not influential for any of the four models. For the coefficients perturbations only the case #100 were influential for all replications in the Normal models and for some replications in Contaminated Normal model, in this scenario with influence measure quite close between these two models.

Finally, we presented the graphical representation of the results of one replication. The both contaminated observations were influential to Normal model in the case-weight (Figure 2) and on scale (Figure 3) perturbation schemes. For the explanatory variable and coefficients perturbations (Figures 4 and 5) only the case #100 was influential to Normal model. In the coefficients perturbation, specifically, the case #24 (not contaminated by our purpose) appeared as influential for the non-normal models, but with less influence over the Student-t model.

Our findings suggests the heavy-tailed models were robust against outliers and the Student-t model was the best model, because it presented lower measures values.

6 Application

In this section we apply the proposed methodology in two real data for nonlinear an for linear censored regression models.

6.1 Nonlinear censored regression model

In this Subsection we will work with the low cycle fatigue data from a strain-controlled test presented in Nelson (2004). This study consists in count the number of cycles until the failure of 26 specimens of metals submitted to pseudo stress.

The following model was proposed to fit this data:

$$y_i = \beta_1 \exp(\beta_2 x_i) + \epsilon_i, \ i = 1, ..., 26,$$

where $y_i = log_{10}$Cycles, $x_i = 1/$(Pseudo-stress) and $\epsilon_i \sim NI(0, \sigma^2, \nu)$. The response is right censored because some metals could not fail at the end of the study. This dataset has 15,4% of censored observations. The Normal, Student-t ($\nu = 2.1$), Slash ($\nu = 1.2$) and Contaminated Normal ($\nu = (0.1, 0.3)$) models were tested. The first three lines in Table 3 shows the results. The Student-t is the model that provided the most accurate estimates (smaller standard error - SE).

The robustness of the models proposed was assessed by considering the influence of a single outlying observation on the EM estimate of $\boldsymbol{\theta}$. For this we replace y_{10} for $y_{10} + \tau$, τ between 0 and 5, and we calculate the relative change (in %) in the estimates. Figure 6 shows the heavy-tailed models were less influenced by outliers than the Normal model for the $\boldsymbol{\beta}$'s parameters. In the estimation of σ^2, the Contaminated Normal model was influenced by contamination. The Student-t is the model less affected by the contamination, providing stable estimates in all cases studied (Table 2).

To identify atypical observations or model misspecification we calculate the AOR for this model

$$r_i = \left| y_i - \hat{\beta}_1 \exp(\hat{\beta}_2 x_i) \right|, \; i = 1, ..., n, \tag{11}$$

where $\hat{\beta}_i$'s are the EM estimates for the β_i, $i = 1, 2$.. The observation #5 presented a value of residual far from the others observations. To the following analysis we consider the perturbation vector $\omega = AOR$ and we apply the global and local influence techniques proposed in Section 4. By Cook distance (Figure 7) we can see the observation is highly influential to Normal Model and quite less influential to Contaminated Normal model. The others heavy-tailed models were not influenced by this observation.

In the local influence diagnostics these observations were influential to Normal and Contaminated Normal models for the case-weight (Figure 8) and scale (Figure 9) perturbations schemes. In the explanatory variable (Figure 10) and coefficients (Figure 11) perturbations schemes there was no observation influential. The Student-t and Slash models were not influenced by this observations.

Analyzing the observation #5 of the dataset we notice that it was the metal which failed with smaller number of cycles between the metals submitted to pseudo stress less than 100. We can see the case #5 was influential in global model estimation and on scale parameter estimation for both Normal and Contaminated Normal models, although this influence is lower in Contaminated Normal model. To assess the impact of this observation we used the measures total relative changes (TRC) and maximum relative changes (MRC) suggested by Lee et al. (2006), defined by

$$\begin{aligned} TRC &= \sum_{i=1}^{n_p} \frac{|\hat{\theta}_i - \hat{\theta}_i^o|}{\hat{\theta}_i}, \text{ and} \\ MRC &= \max_i \frac{|\hat{\theta}_i - \hat{\theta}_i^o|}{\hat{\theta}_i}, \end{aligned}$$

where $\hat{\theta}_i$ and $\hat{\theta}_i^o$ are the EM estimates of the model obtained from the data with and without the influential observation. These measures for the four proposed models were presented in the last two lines of the Table 3. The impact of this observation was clearly smaller in the non-normal models.

Comparing the models in estimation and diagnostic analysis, the Student-t model is the better one because it provides more precise estimates and it is more robust than the others non-normal models analyzed (as shown in our simulation study) in presence of atypical observations.

6.2 Linear censored regression model

We will consider the motorette data reported in Schmee and Hahn (1979). This study consists in times to failure in hours of 40 motorettes tested at 4 different temperatures: 150°C, 170°C, 190°C and 220°C.

The following model was proposed to fit this data:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i,\ i = 1, ..., 40,$$

where $y_i = log_{10}(t_i)$, t_i is the ith failure time, $x_i = 1000/(\text{temp}_i + 273.2)$ e $\epsilon_i \sim NI(0, \sigma^2, \nu)$. The response is right censored because some motorettes may not fail at the end of the study. For this dataset we have censoring level of 57.5%. The Normal, Student-t ($\nu = 2.1$), Slash ($\nu = 1.2$) and Contaminated Normal ($\nu = (0.1, 0.3)$) models were tested. First three lines in Table 5 shows the results. The Student-t is the model provides more accurate estimates (smaller SE).

The robustness of the models proposed was assessed by considering the influence of a single outlying observation on the EM estimate of $\boldsymbol{\theta}$. For this we replace y_{10} for $y_{10} + \tau$, τ between 0 and 5, and we calculate the relative change (in %) in the estimates. Figure 12 and Table 4 show the heavy-tailed models were less influenced by outliers than the Normal model for the $\boldsymbol{\beta}$'s parameters. In the estimation of σ^2, the Contaminated Normal model has been so influenced as the Normal model. The Student-t is the model less affected by the contamination, providing stable estimates in all cases studied.

To identify atypical observations or model misspecification we calculate the AOR presented in (11). The observations #21 and #22 presented residuals lower than -2.5. By Cook distance (Figure 13) we can see the two observations are influential only for Normal Model, corroborating the analysis of this datased showed by Zhu et al. (2001). The case #11 presented a value of Cook distance closer to the benchmark value (0.6) for the Normal model.

In the local influence diagnostics these observations were influential to Normal model for all the four perturbations schemes analyzed. The heavy-tailed models were not influenced by this observations (Figures 14-17).

We can conclude that, for the Normal model, cases #21 and #22 were highly influential in global model estimation, for both $\boldsymbol{\beta}$'s and σ^2 estimation. In addition the values of the explanatory variable for these cases were influential in estimation process. In the dataset we can see that these two values presented the lower value of failure time and it were very different of the others failure times for the 190°C.

We would like to emphasize that the set of influence techniques do not have the purpose of exclude the influential cases, but recognize them and choose a better way to deal with it. To use the Normal model we may suggest the inspection of this observations to see if it are from measurement errors of typing errors, and if possible, try to corret it. If these cases were genuine observations we may propose a more adequate model to accomodate this effects or list this weakness of the estimation process. The fact the results obtained were highly dependent of this two observations. In the model estimated without this cases we observed reduction of 44% in the standard errors of the coefficients and 69% in the scale parameter (see Table 5). For the non-normal models, the results were robust against these cases.

Comparing the models in estimation and diagnostic analysis, the Student-t model is the better one because it provides more precise estimates and it is more robust than the others non-normal models analyzed (as shown in our simulation study) in presence of atypical observations.

7 Conclusions

In this paper we study nonlinear censored regression models with errors following distributions of the Normal/Independent family (NLNICR). The parameters estimation was obtained by EM algorithm through the method proposed by Garay et al. (2015a,b). We present extensions of some diagnostic methods to NLNICR models based on case-deletion and for the local influence analyze, we used four perturbations schemes: case weight, scale, covariable and coefficients. Simulations studies and an application were presented.

The simulation studies and application showed that the proposed methodology is able to correctly detect the influential observations and that the heavy-tailed models (Student-t, Slash and Contaminated Normal) were less influenced by outlying observations than the Normal model as expected. Therefore, we conclude that the influential diagnosis presented can provide a tool set to identify the goodness of fit of nonlinear censored regression models in the

NI family and that the non-Normal distributions can be required in modeling. In this case, we suggest the heavy tail distributions in the NI family as an alternative to model data with outliers without the necessity of the use of any transformations or deletion of data, which would mean loss of information.

Acknowledgments

The authors acknowledge financial support from CNPq and CAPES, Brazil. We would like to thank Aldo Garay for his precious help in the formulation of this article.

References

Cancho VG, Lachos VH, Ortega EMM, 2010. A nonlinear regression model with skew-normal errors. Statistical Papers 51, 547–558.

Cancho VG, Dey DK, Lachos VH, Andrade MG, 2011. Bayesian nonlinear regression models with scale mixtures of skew-normal distributions: Estimation and case influence diagnostics. Computational Statistics and Data Analysis 55, 588–602.

Cook RD, 1977. Detection of influential observation in linear regression. Technometrics 19, 15–18.

Cook RD, 1986. Assessment of local influence. Journal of the Royal Statistical Society, Series B 48, 133–169.

Cook RD, Weisberg S, 1982. Residual and influence in regression. London: Chapman & Hall.

Dempster AP, Laird NM, Rubin DB, 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 39, 1–38.

Garay AM, Lachos VH, Abanto-Valle CA, 2011. Nonlinear regression models based on scale mixtures of skew-normal distributions. Journal of the Korean Statistical Society 40, 115–124.

Garay AM, Lachos VH, Bolfarine H, Cabral CRB, 2015. Linear censored regression models with scale mixtures of normal distributions. Statistical Papers. Forthcoming.

Garay AM, Lachos VH, Lin T, 2015. Nonlinear censored regression models with scale mixtures of normal distributions. Statistics and its Interface, forthcoming.

Heuchenne C, Keilegom IV, 2007. Nonlinear regression with censored data. Technometrics 49, 34–44.

Labra FV, Garay AM, Lachos VH, Ortega EMM, 2012. Estimation and diagnostics for heteroscedastic nonlinear regression models based on scale mixtures of skew-normal distributions. Journal of Statistical Planning and Inference 142, 2149–2165.

Lachos VH, 2002. Infuência e diagnóstico em modelos de Grubbs. Master thesis. Universidade Estadual de Campinas, So Paulo.

Lachos VH, Bandyopadhyay D, Dey DK, 2011. Linear and nonlinear mixed-effects models for censored HIV viral loads using normal/independent distributions. Biometrics 67, 1594–1604.

Lange KL, Sinsheimer JS, 1993. Normal/independent distributions and their applications in robust regression. Journal of Computational and Graphical Statistics 2, 175–198.

Lange KL, Little RJA, Taylor JMG, 1989. Robust statistical modeling using t distribution. Journal of the American Statistical Association 84, 881–896.

Lee S, Xu L, 2004. Influence analysis of nonlinear mixed-effects models. Computational Statistics and Data Analysis. 45, 321–341.

Lee S, Lu B, Song X, 2006. Assessing local influence for nonlinear structural equation models with ignorable missing data. Computational Statistics and Data Analysis. 50, 1356–1377.

Liu C, 1996. Bayesian robust multivariate linear regression with incomplete data. Journal of the American Statistical Association 91, 1219–1227.

Massuia MB, Cabral CRB, Matos LA, Lachos VH. Influence diagnostics for student-t censored linear regression models. Statistics. 2014. DOI 10.1080/02331888.2014.958489

Matos LA, Lachos VH, Balakrishnan N, Labra FV, 2013. Influence diagnostics in linear and nonlinear mixed-effects models with censored data. Computational Statistics and Data Analysis 57, 450–464.

Meza C, Osorio F, De la Cruz R, 2012. Estimation in nonlinear mixed-effects models using heavy-tailed distributions. Statistics and Computing 22, 121–139.

Montgomery DC, Peck EA. Introduction to linear regression analysis. John Wiley; 1992.

Nelson WB, 2004. Accelerated testing: statistical models, test plans and data analysis. Wiley.

Ortega EMM, Bolfarine H, Paula GA, 2003. Influence diagnostics in generalized loggamma regression models. Computational Statistics and Data Analysis 42, 165–186.

Poon WY, Poon YS, 1999. Conformal normal curvature and assessment of local influence. Journal of the Royal Statistical Society, Series B 61, 51–61.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. (http://www.R-project.org/)

Rosa GJM, Padovani CR, Gianola D, 2003. Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. Biometrical Journal 45, 573–590.

Salgado FAO. Diagnóstico de infuência em modelos elípticos com efeitos mistos. Phd thesis. Universidade de São Paulo. So Paulo, 2006.

Schmee J, Hahn GJ. A simple method for regression analysis with censored data. Technometrics. 1979; 21:417–432.

Tsuyuguchi AB, 2012. Testes de bondade de ajuste para a distribuição Birnbaum-Saunders. Master thesis. Universidade Federal de Campina Grande.

Vanegas LH, Cysneiros FJA, 2010. Assessment of diagnostic procedures in symmetrical nonlinear regression models. Computational Statistics and Data Analysis 54, 1002–1016.

Xie F, Lin J, Wei B, 2009. Diagnostics for skew-normal nonlinear regression models with AR(1) errors. Computational Statistics and Data Analysis 53, 4403–4416.

Wang J, Genton MG, 2006. The multivariate skew-slash distribution. Journal of Statistical Planning and Inference 136, 209–220.

Wang J, 2007. EM algorithms for nonlinear mixed effects models. Computational Statistics and Data Analysis 51, 3244–3256.

Zeller CB, Labra FV, Lachos VH, Balakrishnan N, 2010. Influence analysis of skewnormal/independent linear mixed models. Computational Statistics and Data Analysis 54, 1266–1280.

Zhu H, Lee S, 2001. Local influence for incomplete-data models. Journal of the Royal Statistical Society, Series B 63, 111–126.

Zhu Z, Lee S, Wei B, Zhou J, 2001. Case-deletion measures for models with incomplete data. Biometrika, 88, 727–737.

Appendix

Expressions for the covariable and coefficients perturbations schemes to application. For the local influence diagnostic we calculate the matrix $\mathbf{\Delta}_{\boldsymbol{\omega}_0}$ to the covariable and coefficients perturbation schemes.

Covariable perturbation: We disturb the continue covariable $x_i(\omega_i) = x_i \times \omega_i$, where x_i is the ith entry of the vector $\mathbf{X}$. We have $\boldsymbol{\omega}_0 = 1$. We had to replace $x_i(\boldsymbol{\omega}) = x_i\omega_i$ in Q-perturbed function, inside of nonlinear function $\eta(\boldsymbol{\beta}, x_i) = \beta_1 \exp(\beta_2 x_i)$. This way we have

$$
\begin{aligned}
\mathbf{\Delta}_{\beta_1} &= \frac{1}{\widehat{\sigma^2}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\theta})\beta_2 x_i \exp(\beta_2 x_i \omega_i) - 2\mathcal{E}_{0i}(\widehat{\theta})\beta_1\beta_2 x_i \exp(2\beta_2 x_i \omega_i) \right], \\
\mathbf{\Delta}_{\beta_2} &= \frac{1}{\widehat{\sigma^2}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\theta})\beta_1 x_i \exp(\beta_2 x_i \omega_i) + \mathcal{E}_{1i}(\widehat{\theta})\beta_1\beta_2 x_i^2 \omega_i \exp(\beta_2 x_i \omega_i) \right. \\
&\quad \left. -\mathcal{E}_{0i}(\widehat{\theta})\beta_1^2 x_i \exp(2\beta_2 x_i \omega_i) - 2\mathcal{E}_{0i}(\widehat{\theta})\beta_1^2\beta_2 x_i^2 \omega_i \exp(2\beta_2 x_i \omega_i) \right], \text{ and} \\
\mathbf{\Delta}_{\sigma^2} &= -\frac{1}{\widehat{\sigma^4}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\theta})\beta_1\beta_2 x_i \exp(\beta_2 x_i \omega_i) - \mathcal{E}_{0i}(\widehat{\theta})\beta_1^2\beta_2 x_i \exp(2\beta_2 x_i \omega_i) \right].
\end{aligned}
$$

Coefficient perturbation: The perturbation on β's parameters is introduced replacing β by $\beta(\omega) = \beta\omega_i$, $i = 1, ..., n$ $\boldsymbol{\omega} \in \mathbb{R}^n$ in the Q-perturbed function. In this case

$$
\begin{aligned}
\mathbf{\Delta}_{\beta_1} &= \frac{1}{\widehat{\sigma^2}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\theta}) \exp(\beta_2 x_i \omega_i) + \mathcal{E}_{1i}(\widehat{\theta})\beta_2 \omega_i x_i \exp(\beta_2 x_i \omega_i) \right. \\
&\quad \left. -2\mathcal{E}_{0i}(\widehat{\theta})\beta_1 \omega_i \exp(2\beta_2 x_i \omega_i) - 2\mathcal{E}_{0i}(\widehat{\theta})\beta_1\beta_2 \omega_i^2 x_i \exp(2\beta_2 x_i \omega_i) \right], \\
\mathbf{\Delta}_{\beta_2} &= \frac{1}{\widehat{\sigma^2}} \sum_{i=1}^{n} \left[2\mathcal{E}_{1i}(\widehat{\theta})\beta_1 \omega_i x_i \exp(\beta_2 x_i \omega_i) + \mathcal{E}_{1i}(\widehat{\theta})\beta_1\beta_2 \omega_i^2 x_i^2 \exp(\beta_2 x_i \omega_i) \right. \\
&\quad \left. -3\mathcal{E}_{0i}(\widehat{\theta})\beta_1^2 \omega_i^2 x_i \exp(2\beta_2 x_i \omega_i) - 2\mathcal{E}_{0i}(\widehat{\theta})\beta_1^2\beta_2 \omega_i^3 x_i^2 \exp(2\beta_2 x_i \omega_i) \right], \text{ and} \\
\mathbf{\Delta}_{\sigma^2} &= -\frac{1}{\widehat{\sigma^4}} \sum_{i=1}^{n} \left[\mathcal{E}_{1i}(\widehat{\theta})\beta_1 \exp(\beta_2 x_i \omega_i) + \mathcal{E}_{1i}(\widehat{\theta})\beta_1\beta_2 \omega_i x_i \exp(\beta_2 x_i \omega_i) \right. \\
&\quad \left. -\mathcal{E}_{0i}(\widehat{\theta})\beta_1^2 \omega_i \exp(2\beta_2 x_i \omega_i) - \mathcal{E}_{0i}(\widehat{\theta})\beta_1^2\beta_2 \omega_i^2 x_i \exp(2\beta_2 x_i \omega_i) \right].
\end{aligned}
$$

Tables

Table 1: Influence analysis for the cases #1 and #100 in the simulations for kind of diagnostic and model in Monte Carlo study.

Influence diagnostic	Statistic	Normal #1	Normal #100	Student-t #1	Student-t #100	Slash #1	Slash #100	Cont. Normal #1	Cont. Normal #100
Case deletion	% Influential[a]	100%	100%	0%	0%	0%	0%	54%	73%
	Mean measure	4.985	4.817	0.045	0.054	0.053	0.055	0.116	0.154
	SD[b] measure	(0.692)	(0.642)	(0.006)	(0.002)	(0.006)	(0.007)	(0.048)	(0.043)
	Benchmark	0.080		0.080		0.080		0.080	
Case-weight	% Influential	98%	95%	0%	0%	0%	0%	88%	69%
	Mean measure	0.404	0.446	0.029	0.030	0.035	0.033	0.117	0.111
	SD measure	(0.031)	(0.032)	(0.003)	(0.002)	(0.004)	(0.003)	(0.024)	(0.025)
	Mean (SD) Benchmark	0.219 (0.005)		0.042 (0.003)		0.049 (0.005)		0.094 (0.009)	
Scale	% Influential	100%	100%	0%	0%	0%	0%	94%	0%
	Mean measure	0.444	0.414	0.019	0.001	0.025	0.010	0.182	0.002
	SD measure	(0.022)	(0.030)	(0.002)	(0.000)	(0.004)	(0.003)	(0.043)	(0.001)
	Mean (SD) Benchmark	0.229 (0.005)		0.054 (0.006)		0.062 (0.010)		0.086 (0.011)	
Explanatory variable	% Influential	0%	100%	0%	0%	0%	0%	0%	0%
	Mean measure	0.001	0.866	0.001	0.005	0.001	0.004	0.000	0.002
	SD measure	(0.003)	(0.028)	(0.000)	(0.004)	(0.000)	(0.004)	(0.000)	(0.004)
	Mean (SD) Benchmark	0.313 (0.010)		0.109 (0.010)		0.113 (0.001)		0.128 (0.022)	
Coefficients	% Influential	0%	100%	0%	0%	0%	0%	0%	56%
	Mean measure	0.011	0.980	0.003	0.001	0.008	0.002	0.007	0.837
	SD measure	(0.002)	(0.025)	(0.003)	(0.002)	(0.010)	(0.001)	(0.004)	(0.068)
	Mean (SD) Benchmark	0.353 (0.002)		0.196 (0.045)		0.101 (0.005)		0.317 (0.022)	

[a] % Influential: this statistic presents the percentage of Monte Carlo replicates in which the observation was regarded as influential (exceeded the benchmark value).
[b] SD is the abbreviation of standard deviation.

Table 2: Parameters estimation (SE in brackets) for the model without the #10 case and for different contamination level τ.

Parameters	Estimates					
	Without #10	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$
Normal model						
β_1	2.427	2.366	2.271	2.173	2.080	1.991
	(0.317)	(0.356)	(0.509)	(0.701)	(0.907)	(1.119)
β_2	62.994	66.241	71.409	77.090	82.733	88.314
	(31.178)	(34.658)	(49.418)	(67.978)	(87.921)	(18.480)
σ^2	0.078	0.101	0.208	0.397	0.667	1.017
Student-t model						
β_1	2.382	2.370	2.374	2.377	2.377	2.378
	(0.185)	(0.215)	(0.221)	(0.222)	(0.221)	(0.222)
β_2	65.980	66.615	66.309	66.211	66.166	66.140
	(18.222)	(20.973)	(21.545)	(21.647)	(21.631)	(21.688)
σ^2	0.018	0.024	0.026	0.026	0.026	0.026
Slash model						
β_1	2.380	2.370	2.375	2.377	2.377	2.378
	(0.183)	(0.209)	(0.209)	(0.209)	(0.210)	(0.211)
β_2	66.096	66.631	66.314	66.237	66.203	66.183
	(18.467)	(20.872)	(20.888)	(20.858)	(10.930)	(21.025)
σ^2	0.010	0.013	0.013	0.013	0.013	0.013
Contaminated Normal model						
β_1	2.390	2.385	2.360	2.329	2.321	2.330
	(0.194)	(0.203)	(0.269)	(0.352)	(0.410)	(0.505)
β_2	65.671	66.000	67.028	68.303	68.569	68.102
	(19.350)	(20.030)	(26.041)	(34.455)	(39.929)	(49.189)
σ^2	0.015	0.017	0.037	0.064	0.107	0.161

The estimation of the models with $\tau = 0$ is in the first three lines of the Table 3.

Table 3: Parameters estimation by EM algorithm and impact measures of influential observation.

Description	Parameter	Normal Estimate (SE [a])	Student-t Estimate (SE)	Slash Estimate (SE)	Contaminated Normal Estimate (SE)
Full data	β_1	2.445 (0.312)	2.395 (0.203)	2.396 (0.209)	2.405 (0.201)
	β_2	62.049 (30.536)	65.232 (19.892)	65.168 (20.969)	64.987 (19.956)
	σ^2	0.077	0.022	0.013	0.016
Without #5	β_1	2.366 (0.224)	2.388 (0.175)	2.388 (0.184)	2.375 (0.186)
	β_2	65.959 (22.030)	65.649 (17.280)	65.635 (18.705)	66.121 (18.672)
	σ^2	0.039	0.016	0.009	0.012
Impact measures	TRC	0.589	0.282	0.318	0.279
	MRC	0.494	0.273	0.308	0.250

[a] SE is the standard error of the estimates.

Table 4: Parameters estimation (SE in brackets) for the model without the #10 case and for different contamination level τ.

Parameters	Estimates					
	Without #10	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$
Normal model						
β_0	-5.938	-7.026	-8.511	-10.090	-11.696	-13.313
	(0.709)	(0.976)	(1.633)	(2.412)	(3.230)	(4.062)
β_1	4.273	4.805	5.547	6.338	7.144	7.956
	(0.321)	(0.444)	(0.749)	(1.111)	(1.492)	(1.879)
σ^2	0.067	0.114	0.255	0.487	0.807	1.216
Student-t model						
β_0	-5.607	-5.682	-5.640	-5.631	-5.627	-5.624
	(0.419)	(0.434)	(0.417)	(0.416)	(0.415)	(0.415)
β_1	4.107	4.141	4.119	4.115	4.113	4.112
	(0.190)	(0.197)	(0.189)	(0.188)	(0.188)	(0.188)
σ^2	0.014	0.013	0.012	0.012	0.012	0.012
Slash model						
β_0	-5.593	-5.989	-5.815	-5.776	-5.758	-5.747
	(0.544)	(0.870)	(0.819)	(0.811)	(0.813)	(0.803)
β_1	4.106	4.301	4.217	4.198	4.190	4.185
	(0.247)	(0.399)	(0.376)	(0.372)	(0.373)	(0.368)
σ^2	0.016	0.028	0.026	0.026	0.025	0.025
Contaminated Normal model						
β_0	-5.589	-5.839	-6.190	-6.572	-6.920	-7.271
	(0.471)	(0.717)	(0.942)	(1.291)	(1.728)	(2.282)
β_1	4.099	4.224	4.403	4.593	4.770	4.950
	(0.214)	(0.330)	(0.431)	(0.594)	(0.800)	(1.062)
σ^2	0.013	0.019	0.071	0.113	0.160	0.217

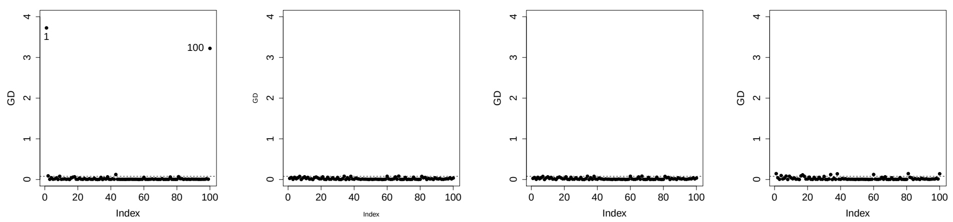
The estimation of the models with $\tau = 0$ is in the first three lines of the Table **??**.

Table 5: Parameters estimation by EM algorithm.

Description	Parameter	Normal Estimate (SE [a])	Student-t Estimate (SE)	Slash Estimate (SE)	Contaminated Normal Estimate (SE)
Full data	β_0	-6.019 (0.695)	-5.635 (0.407)	-5.600 (0.511)	-5.614 (0.462)
	β_1	4.311 (0.314)	4.120 (0.184)	4.107 (0.231)	4.111 (0.209)
	σ^2	0.067	0.014	0.010	0.013
Without #21 [b]	β_0	-5.775 (0.572)	-5.608 (0.383)	-5.563 (0.453)	-5.571 (0.389)
	β_1	4.189 (0.258)	4.109 (0.173)	4.089 (0.206)	4.092 (0.176)
	σ^2	0.046	0.013	0.009	0.012
Without #21 and #22	β_0	-5.422 (0.387)	-5.524 (0.358)	-5.538 (0.413)	-5.527 (0.360)
	β_1	4.028 (0.175)	4.072 (0.162)	4.077 (0.188)	4.073 (0.163)
	σ^2	0.021	0.014	0.007	0.011

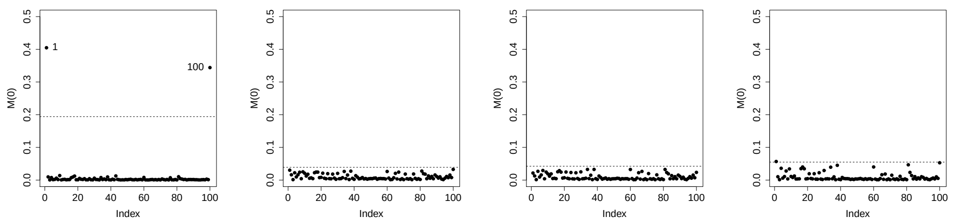
[a] SE is the standard error of the estimates.
[b] The observations #21 and #22 are equal (see Table **??**).

Figures



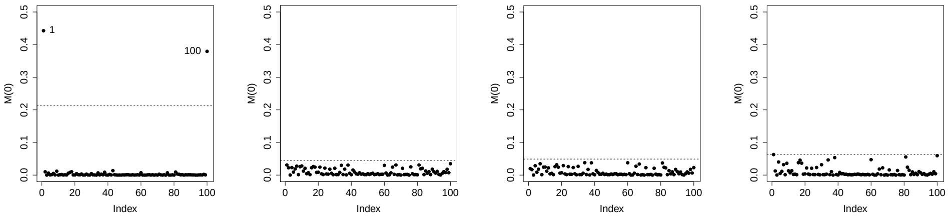
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 1: Global influence: generalized Cook distance in simulation study.



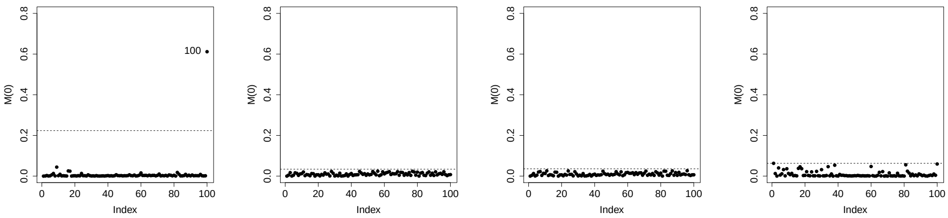
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 2: Local influence: case weight perturbation in simulation study.



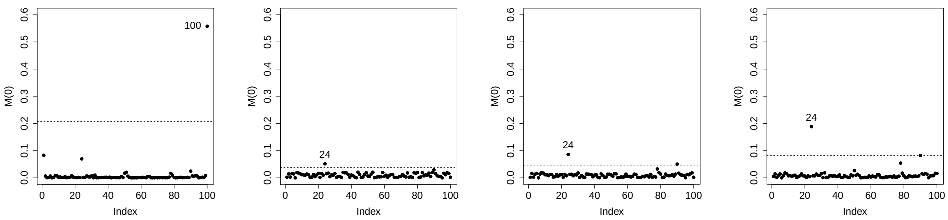
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 3: Local influence: scale perturbation in simulation study.



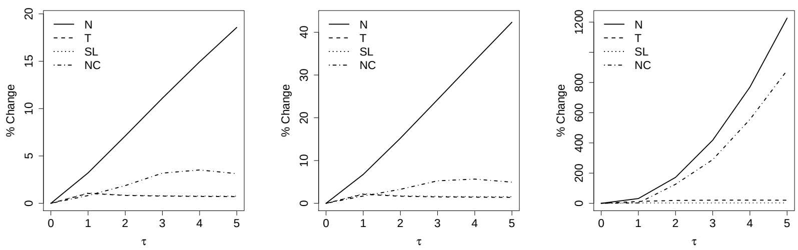
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 4: Local influence: on explanatory variable perturbation in simulation study.



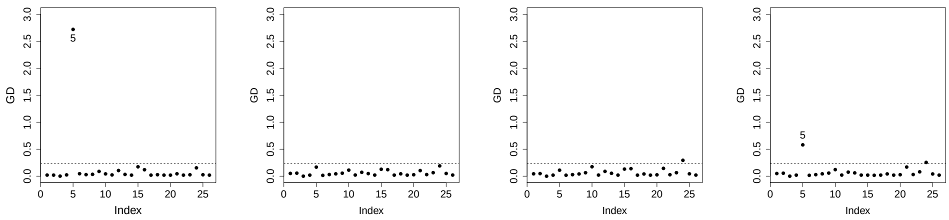
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 5: Local influence: coefficients perturbation in simulation study.



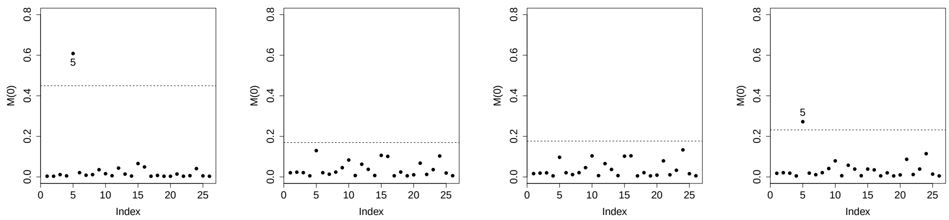
(a) β_0 **(b)** β_1 **(c)** σ^2

Figure 6: Relative changes in estimates for contamination level.



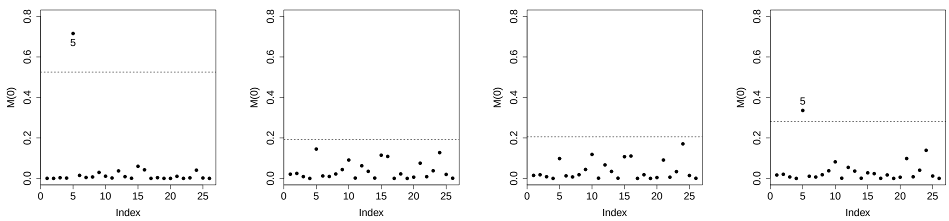
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 7: Global influence: generalized Cook distance in application.



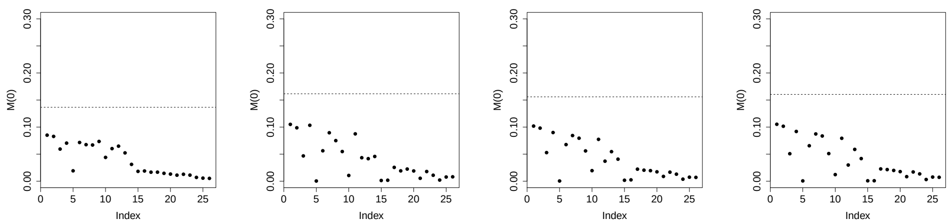
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 8: Local influence: case weight perturbation in application.



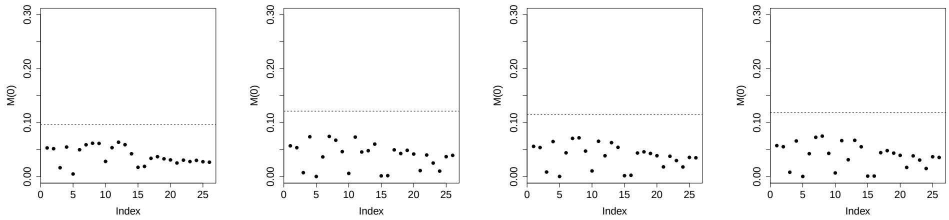
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 9: Local influence: scale perturbation in application.



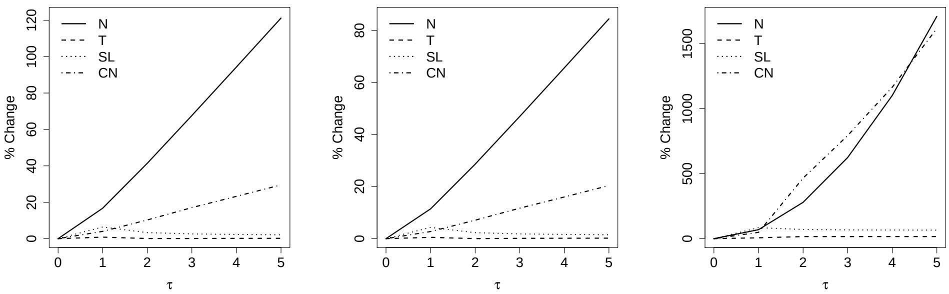
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 10: Local influence: on explanatory variable perturbation in application.



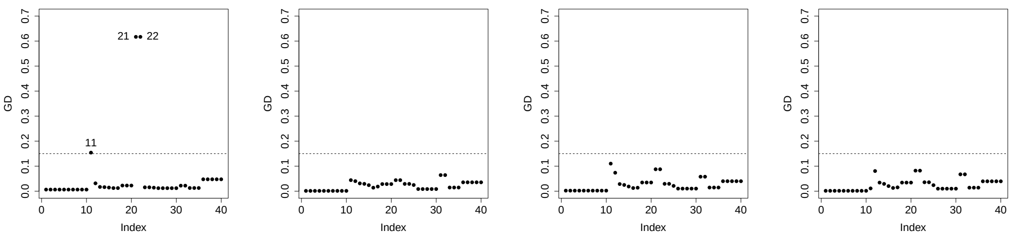
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 11: Local influence: coefficients perturbation in application.



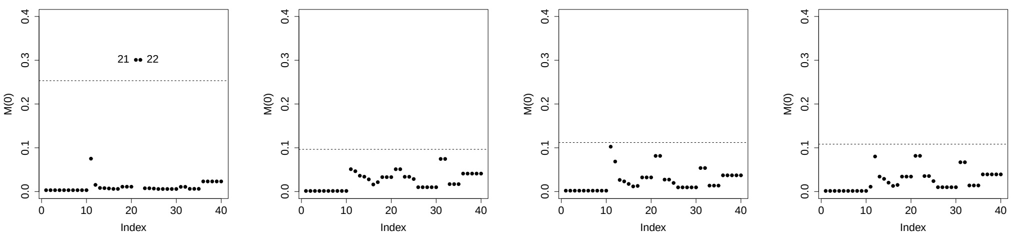
(a) β_0 **(b)** β_1 **(c)** σ^2

Figure 12: Relative changes in estimates for contamination level for censored linear model.



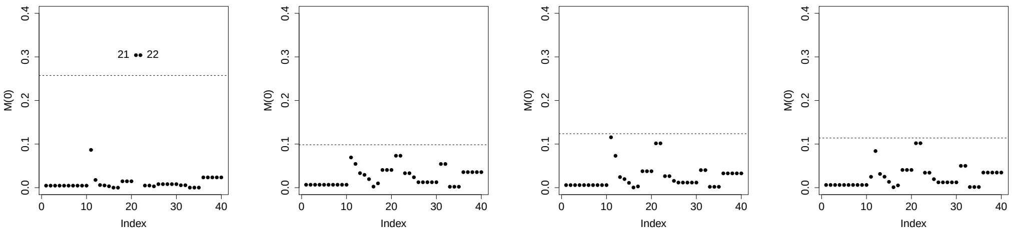
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 13: Global influence: generalized Cook distance for censored linear model.



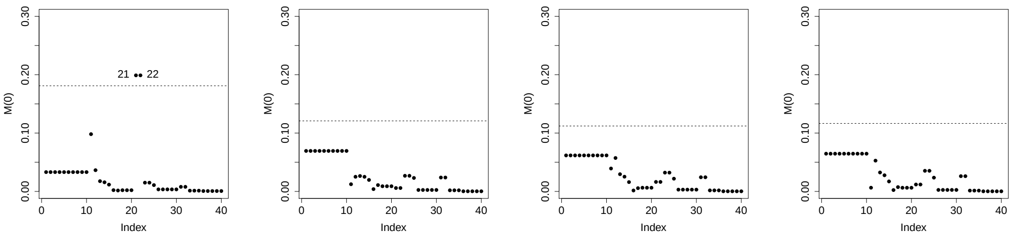
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 14: Local influence: case weight perturbation for censored linear model.



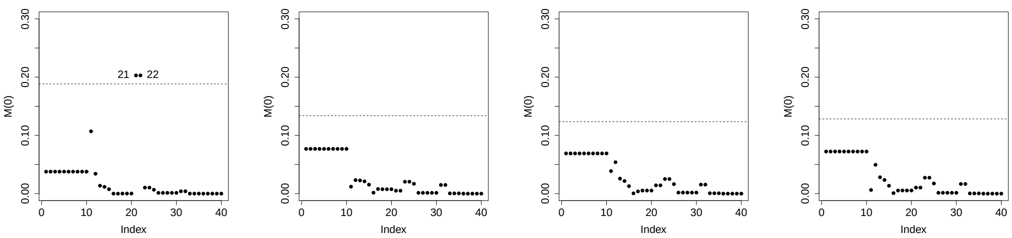
(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 15: Local influence: scale perturbation for censored linear model.



(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 16: Local influence: explanatory variable perturbation for censored linear model.



(a) Normal **(b)** Student-t **(c)** Slash **(d)** Cont. Normal

Figure 17: Local influence: coefficients perturbation for censored linear model.