Computational Statistics and Data Analysis 64 (2013) 153–164

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

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

Linear regression models with slash-elliptical errors

Izabel Cristina Alcantara[a,b], Francisco José A. Cysneiros[b,*]

[a] *Departamento de Estatística - UFPB, Cidade Universitária, Campus 1, 58051-900, João Pessoa - PB, Brazil*

[b] *Departamento de Estatística - UFPE, Cidade Universitária, 50740-540, Recife - PE, Brazil*

ARTICLE INFO

Article history:
Received 12 May 2012
Received in revised form 18 November 2012
Accepted 25 February 2013
Available online 15 March 2013

Keywords:
Heavy-tailed distribution
Leverage
Local influence
Residuals
Robust models
Slash-elliptical distribution

ABSTRACT

We propose a linear regression model with slash-elliptical errors. The slash-elliptical distribution with parameter q is defined as the ratio of two independent random variables Z and $U^{\frac{1}{q}}$, where Z has elliptical distribution and U has uniform distribution in $(0, 1)$. The main feature of the slash-elliptical distribution is to have greater flexibility in the degree of kurtosis when compared to the elliptical distributions. Other advantages of this distribution are the properties of symmetry, heavy tails and the inclusion of the elliptical family as a limit case when $q \to \infty$. We develop the methodology of estimation, hypothesis testing, generalized leverage and residuals for the proposed model. In the analysis of local influence, we also develop the diagnostic measures based on the likelihood displacement under the some perturbation schemes. Finally, we present a real example where slash-Student-t model is more stable than other considered models.

1. Introduction

Last decades, similar works have been developed based on symmetrical distribution. A review of some areas where symmetric distributions are applied is described in Chmielewski (1981). In many situations of statistical modeling there is a need of searching for less sensitive models of outlying observations. Galea et al. (2003) developed diagnostic methods for linear symmetrical models and Cysneiros and Paula (2005) developed restricted methods in symmetrical linear models. Paula et al. (2009) introduced the class of linear models with first-order autoregressive elliptical errors and diagnostic methods were derived.

Rogers and Tukey (1972) presented the slash distribution as the probability distribution of a standard normal variable divided by an independent standard uniform variable. In the general case, we say that a random variable S has standard slash distribution with parameter $q > 0$ if it can be expressed as the ratio of two independent random variables Z and $U^{\frac{1}{q}}$, where Z has standard normal distribution $N(0, 1)$ and U has uniform distribution in $(0, 1)$. The slash distribution has the properties of symmetry, heavy tails and converges to the normal distribution when $q \to \infty$.

Rogers and Tukey (1972) and Mosteller and Tukey (1977) discussed slash distribution and its properties. Kafadar (1982) proposed maximum likelihood estimators for location-scale parameters considering the slash distribution obtained through linear transformation $Y = \mu + \sqrt{\phi}S$. In the work of Andrews et al. (1972), Gross (1973) and Morgenthaler and Tukey (1991), the slash distribution is mainly used in simulation studies whose scenario involves extreme situations. Wang and Genton (2006) defining a skewed version of the slash distribution, assumed that the random variable Z has a multivariate skew normal distribution. Arslan (2008) introduced a new class of multivariate skew-slash distributions using the normal variance-mean mixture approach. Later, Arslan and Genç (2009) generalized the family of distributions proposed by Wang

* Corresponding author. Tel.: +55 81 21267422; fax: +55 81 21268422.
E-mail addresses: izabel@de.ufpb.br (I.C. Alcantara), cysneiros@de.ufpe.br, francisco.cysneiros@gmail.com (F.J.A. Cysneiros).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.02.029

and Genton (2006), constructed a family of multivariate distributions as a scale mixtures of the multivariate symmetric Kotz-type distribution and the uniform distribution. Lachos et al. (2008) derived diagnostic methods based on the local influence on scale-mixture models. Ferreira et al. (2011) developed EM algorithm for estimation of the parameters on scale-mixture models. Gómez et al. (2007) considered the standard slash distribution in the form:

$$Z = U^{1/q}S \sim N(0, 1)$$

and generalized slash distribution by replacing the distribution of Z by the family of univariate and multivariate elliptical distributions. This new family of distributions proposed by Gómez et al. (2007), known as slash-elliptical distribution, has the property of symmetry and greater flexibility in the degree of kurtosis when compared to the elliptical distribution. These properties were also observed by Genç (2007) to slash-power-exponential distribution. Another advantage of this family of distributions is to contain the elliptical family as a limiting case.

Given this new family of distributions, the linear regression model with error distribution in a univariate slash-elliptical family is developed, which will be called only by slash-elliptical distribution. The aim of this paper is to derive a methodology for estimating, hypothesis testing, generalized leverage, residual and diagnostic analysis based on the local influence approach. Section 2 introduces the slash-elliptical regression model and procedures for estimation are presented. Simulation studies of the proposed residual are presented. In addition, diagnostic measures based on the local influence approach are developed in Section 3. Section 4 is devoted to analysis of a real data set using slash-elliptical regression model and finally, some conclusions are presented in the final section.

2. Slash-elliptical regression model

Definition 1. A random variable Y has slash-elliptical distribution with location parameter $\mu \in \Re$ and scale $\phi > 0$ if Y can be expressed as

$$Y = \mu + \sqrt{\phi}\frac{V}{U^{1/q}},$$

where V and U are independent random variables, V has standard elliptical distribution, U has uniform distribution in $(0, 1)$ and q is the specific parameter of slash-elliptical distribution.

Definition 2. A random variable Y is called slash-elliptical variable and denoted by $\sim SEL(\mu, \phi, q, g(\cdot))$ if its density is given by:

$$f(y; \mu, \phi, q) = \frac{1}{\sqrt{\phi}} \begin{cases} \dfrac{q}{2|z|^{q+1}} H(z^2), & z \neq 0 \\ \dfrac{q}{q+1} g(0), & z = 0, \end{cases} \tag{1}$$

where $z = (y - \mu)/\sqrt{\phi}$ and $H(z^2) = \int_0^{z^2} t^{(q-1)/2} g(t)dt$, for some generating function of density $g(\cdot)$, with $g(t) > 0$ for $t > 0$ and $\int_0^{\infty} t^{-1/2} g(t)dt = 1$.

Definitions 1 and 2 can be found in Gómez et al. (2007) and Gómez and Venegas (2008). Whereas, the Definition 1 provides a method for generating slash-elliptical random variables based on elliptical and uniform distributions, the Definition 2 characterizes a slash-elliptical random variable by its density function.

Let $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ independent random variables, where $\epsilon_i \sim SEL(0, \phi, q, g(\cdot))$. We define the linear regression model with slash-elliptical errors and parameter q is fixed or known, which abbreviated to slash-elliptical model, by:

$$y_i = \mathbf{x}_i^t \beta + \epsilon_i, \quad i = 1, 2, \ldots, n, \tag{2}$$

where for each ith observation y_i is the response variable, $\mathbf{x}_i = (1, x_{i2}, \ldots, x_{ip})^t$ is the regressor vector $(p \times 1)$, and $\beta = (\beta_1, \ldots, \beta_p)^t$ is a vector of unknown parameters $(p \times 1)$.

2.1. Parameter estimation

To estimate the parameter vector $\theta = (\beta^t, \phi)^t$, we can use the maximum likelihood method, which is to maximize the log-likelihood function of the slash-elliptical model given by:

$$L(\theta) = -\frac{n}{2}\log(\phi) + \sum_{i \in A} a(z_i) + K, \tag{3}$$

with respect to θ, where $z_i = \frac{y_i - \mu_i}{\sqrt{\phi}}$, $a(z_i) = \log(H(z_i^2)) - (q + 1)\log|z_i|$ is a function that depends on θ through z_i, $A = \{i : z_i \neq 0\}$, $K = (n - n_A)\log\left(\frac{q}{q+1}g(0)\right) + n_A \log(\frac{q}{2})$ is constant with respect to θ and n_A is the number of elements in set A. The score function $U_\theta = (U_\beta^t, U_\phi)^t$ is given by

$$U_\beta = -\frac{1}{\phi}\mathbf{X}^t\mathbf{D}(b)(\mathbf{y} - \mathbf{X}\beta) \quad \text{and} \quad U_\phi = -\frac{1}{2\phi}\mathbf{z}^t\mathbf{v}(\ell),$$

where $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t$, $\mathbf{X} = (\mathbf{x}_1^t, \mathbf{x}_2^t, \ldots, \mathbf{x}_n^t)^t$, $a'(z_i) = \frac{2z_i^2 g(z_i^2)}{H(z_i^2)} - \frac{q+1}{z_i}$, $b(z_i) = \frac{a'(z_i)}{z_i}$, $\ell(z_i) = \frac{n}{n_A z_i} + z_i b(z_i)$, $\mathbf{D}(b) = \operatorname{diag}(b(z_1), \ldots, b(z_n))$ and $\mathbf{v}(\ell) = (\ell(z_1), \ldots, \ell(z_n))^t$. Thus, the score function can be expressed as

$$U_\theta = \widetilde{\mathbf{X}}\widetilde{\mathbf{M}}(\widetilde{\mathbf{y}} - \widetilde{\mathbf{X}}\theta), \tag{4}$$

where $\widetilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X}^t & 0 \\ 0 & 1 \end{bmatrix}$, $\widetilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}$ and $\widetilde{\mathbf{M}} = \begin{bmatrix} -\dfrac{\mathbf{D}(b)}{\phi} & 0 \\ 0 & -\dfrac{\mathbf{z}^t\mathbf{v}(\ell)}{2\phi^2} \end{bmatrix}$.

The Hessian matrix is given by

$$\ddot{L}_{\theta\theta} = \begin{bmatrix} \ddot{L}_{\beta\beta} & \ddot{L}_{\beta\phi} \\ \ddot{L}_{\phi\beta} & \ddot{L}_{\phi\phi} \end{bmatrix} = -(\widetilde{\mathbf{X}}^t\widetilde{\mathbf{V}}\widetilde{\mathbf{X}}), \tag{5}$$

where $g'(u_i) = \frac{\partial g(u_i)}{\partial u_i}$, $m(z_i) = \frac{2z_i^q g(z_i^2)}{H(z_i^2)}$, $c(z_i) = -\frac{q+1}{z_i^2} - \left(\frac{q}{z_i} + \frac{2z_i g'(z_i^2)}{g(z_i^2)}\right) m(z_i) + (m(z_i))^2$, $e(z_i) = z_i(b(z_i) - c(z_i))$, $d(z_i) = z_i^2(c(z_i) - 3b(z_i))$, $\mathbf{D}(c) = \operatorname{diag}(c(z_1), \ldots, c(z_n))$, $\mathbf{v}(e) = (e(z_1), \ldots, e(z_n))^t$, $d = \frac{1}{4\phi^2}\sum_{i\in A} d(z_i) - \frac{n}{2\phi^2}$, $\widetilde{\mathbf{V}} = \begin{bmatrix} \dfrac{\mathbf{D}(c)}{\phi} & -\dfrac{\mathbf{v}(e)}{2\phi^{3/2}} \\ -\dfrac{\mathbf{v}^t(e)}{2\phi^{3/2}} & d \end{bmatrix}$, $\ddot{L}_{\beta\beta} = -\frac{1}{\phi}\mathbf{X}^t\mathbf{D}(c)\mathbf{X}$, $\ddot{L}_{\phi\beta} = \ddot{L}_{\beta\phi}^t = \frac{1}{2\phi^{\frac{3}{2}}}\mathbf{v}^t(e)\mathbf{X}$, $\ddot{L}_{\phi\phi} = -d$, and the observation information matrix is given by $\ddot{J}_{\theta\theta} = (\widetilde{\mathbf{X}}^t\widetilde{\mathbf{V}}\widetilde{\mathbf{X}})$.

We develop the inferential methodology based on estimated asymptotic standard errors obtained by observed information matrix. Thus, by (4) and (5) we get the Newton–Raphson iterative process to estimate θ

$$\theta^{(m+1)} = (\widetilde{\mathbf{X}}^t\widetilde{\mathbf{V}}^{(m)}\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}\widetilde{\mathbf{V}}^{(m)}\widetilde{\mathbf{z}}^{(m)}, \tag{6}$$

where $\widetilde{\mathbf{z}}^{(m)} = \widetilde{\mathbf{X}}\theta^{(m)} + (\widetilde{\mathbf{V}}^{(m)})^{-1}\widetilde{\mathbf{M}}^{(m)}(\widetilde{\mathbf{y}} - \widetilde{\mathbf{X}}\theta^{(m)})$.

2.2. Simulation study

In order to evaluate the performance of the estimators of slash-elliptical model, we conducted a simulation study considering the model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \text{with } i = 1, 2, \ldots, n,$$

where $\beta_0 = 1$ and $\beta_1 = 1$ were kept fixed, and the distribution of the explanatory variable x_i was fixed in $U(0, 1)$ in the course of the simulations. Eight scenarios were considered for the distribution of ϵ_i with $q = 5$ and $q = 10$ for slash-normal, slash-contaminated normal (slash-CN) with $\lambda = 0.2$ and $\sigma = 1.5$, slash-Student-t with $\nu = 3$ and slash-slash with $\gamma = 5$. We generated 10,000 Monte Carlo replicates of the response variable for each of the eight scenario considered for sample sizes $n = 30$ and $n = 50$. We compute the empirical mean, standard error (SE), bias and mean square error (MSE) of the parameter estimator and the results are shown in Table 1.

All the results of this simulation study are developed upon Ox Console version 6.21 (Linux) Doornik (2007).

In all cases, the estimators slash-elliptical models had satisfactory performance. We also observed that, the mean squared errors (MSE) decreased with the increase in sample size and are smaller in models with $q = 10$ in comparison to the same models with $q = 5$ for all estimators.

Lange and Sinsheimer (1993) observed in their study that, slash and Student-t models had a similar performance. However, we observed that the estimators under slash-Student-t model have smaller bias and MSE than estimators under slash-slash model in all cases.

2.3. Goodness of fit and hypothesis testing

Once estimated the slash-elliptical model to a known or fixed q, we can verify the goodness of fit through information criterion such as Akaike Information Criterion—*AIC* (Akaike, 1974) and Bayesian Information Criterion—*BIC* (Schwarz, 1978). According to Kuha (2004), criterion *BIC* shows a better performance than *AIC*, where *BIC* is the most appropriate in case of discordance between these criterion. In the case of the parameter q is not known, one solution is: to fit model for a set of values of $q > 0$ and to calculate the criterion *AIC* or *BIC* of each fitted model. And, the model selected is the one that has the lowest value of *AIC* or *BIC*. In addition to measures of goodness of fit, test of significance of the parameters can also be used as a way to assess the inclusion or exclusion of explanatory variables. The inclusion of non-significant variables in the model

Table 1
Mean, standard error (SE), bias and mean square errors (MSE) for parameter estimators on the simulated models.

		($n = 30$)						($n = 50$)					
		($q = 5$)			($q = 10$)			($q = 5$)			($q = 10$)		
		β_0	β_1	ϕ	β_0	β_1	ϕ	β_0	β_1	ϕ	β_0	β_1	ϕ
Slash-normal	Mean	0.98	1.03	3.79	0.98	1.02	3.75	0.99	1.00	3.88	1.00	1.00	3.85
	SE	0.80	1.42	1.17	0.71	1.27	1.05	0.65	1.15	0.90	0.58	1.03	0.82
	Bias	−0.02	0.03	−0.21	−0.02	0.02	−0.25	−0.01	0.00	−0.12	−0.01	0.00	−0.15
	MSE	0.64	2.03	1.41	0.51	1.60	1.17	0.42	1.32	0.82	0.34	1.05	0.70
Slash-Student-t ($\nu = 3$)	Mean	1.00	1.00	3.87	1.00	1.00	3.86	1.01	0.99	3.93	1.01	0.99	3.92
	SE	0.97	1.74	1.56	0.88	1.58	1.52	0.79	1.38	1.17	0.72	1.25	1.14
	Bias	0.00	0.00	−0.13	0.00	0.00	−0.14	0.01	−0.01	−0.07	0.01	−0.01	−0.08
	MSE	0.95	3.03	2.45	0.78	2.50	2.32	0.62	1.91	1.38	0.51	1.57	1.31
Slash-CN ($\lambda = 0.2, \sigma = 1.5$)	Mean	1.00	0.99	3.80	1.00	0.99	3.77	0.99	1.00	3.89	0.99	1.01	3.87
	SE	0.87	1.56	1.22	0.78	1.40	1.13	0.72	1.26	0.93	0.64	1.13	0.87
	Bias	0.00	−0.02	−0.20	0.00	−0.01	−0.23	−0.01	0.00	−0.12	−0.01	0.01	−0.13
	MSE	0.76	2.45	1.52	0.61	1.97	1.33	0.51	1.60	0.89	0.41	1.28	0.77
Slash-slash ($\gamma = 5$)	Mean	0.99	1.00	4.82	0.99	1.00	4.86	0.99	1.02	4.90	1.01	0.99	3.92
	SE	0.99	1.78	1.58	0.89	1.60	1.53	0.80	1.40	1.21	0.72	1.25	1.14
	Bias	−0.01	0.00	0.82	−0.01	0.00	0.86	−0.01	0.02	0.90	0.01	−0.01	−0.08
	MSE	0.98	3.15	3.16	0.79	2.56	3.08	0.64	1.95	2.28	0.51	1.57	1.31

can reflect the goodness of fit of the model. We applied likelihood ratio, Wald and score tests as presented in Li (2001). The hypotheses considered were: $H_0 : \beta = \beta^0$ against $H_1 : \beta \neq \beta^0$ to a particular known vector β^0 of dimension $(s \times 1)$. The statistics of likelihood ratio (ξ_{RV}), Wald (ξ_W) and score (ξ_{RV}) tests are given by

$$\begin{aligned}
\xi_{RV} &= 2\{L(\hat{\beta}, \hat{\phi}) - L(\beta^0, \hat{\phi}^0)\},\\
\xi_W &= [\hat{\beta} - \beta^0]^t \widehat{\mathrm{Var}}(\hat{\beta})^{-1}[\hat{\beta} - \beta^0] \quad \text{and}\\
\xi_{SR} &= \mathbf{U}^t_{\hat{\beta}^0} \widehat{\mathrm{Var}}_0(\hat{\beta}^0)\mathbf{U}_{\hat{\beta}^0},
\end{aligned}$$

where $\widehat{\mathrm{Var}}(\hat{\beta}) = \hat{\phi}(\mathbf{X}^t\widehat{\mathbf{W}}\mathbf{X})^{-1}$ and $\widehat{\mathrm{Var}}_0(\hat{\beta}^0) = \hat{\phi}^0(\mathbf{X}^t\widehat{\mathbf{W}}^0\mathbf{X})^{-1}$ are the covariances matrices of β on the unrestricted and restricted models, respectively, $\hat{\beta}^0$ and $\hat{\phi}^0$ are the maximum likelihood estimator of β and ϕ on the restricted models. Under regular conditions, we have the statistics ξ_{RV}, ξ_W, and ξ_{SR}, under H_0, follow chi-square distribution with degrees of freedom s.

2.4. Residuals

We investigate in residuals analysis if the discrepancy between observed y_i and fitted $\hat{y}_i$ is statistically relevant. Based on Jørgensen (1984), a standardized residual is proposed. Then, considering the partition of $\widetilde{\mathbf{z}}$ in (6) for the iterative process of β, we obtain the relations:

$$\begin{aligned}
\mathbf{z} &= \mathbf{X}\beta + \mathbf{W}^{-1}\mathbf{S}(\mathbf{y} - \mathbf{X}\beta)\\
\mathbf{W}^{\frac{1}{2}}(\mathbf{z} - \mathbf{X}\beta) &= \mathbf{W}^{-\frac{1}{2}}\mathbf{S}(\mathbf{y} - \mathbf{X}\beta),
\end{aligned}$$

where $\mathbf{S} = -\mathbf{D}(b) - \frac{1}{4\phi^2 d}\mathbf{v}(e)\mathbf{v}^t(\ell)$. Therefore, we consider first the ordinary residual of the least-squares solution of the linear regression model of $\mathbf{z}$ on $\mathbf{X}$, which is defined by

$$\mathbf{r} = \widehat{\mathbf{W}}^{\frac{1}{2}}(\mathbf{z} - \mathbf{X}\hat{\beta}) = \hat{\mathbf{W}}^{-\frac{1}{2}}\hat{\mathbf{S}}(\mathbf{y} - \mathbf{X}\hat{\beta}), \tag{7}$$

then finding a residual for the slash-elliptical model. If we assume that $\mathrm{Var}(\mathbf{z}) \simeq \hat{\phi}\widehat{\mathbf{W}}^{-1}$, we have the variance of the ith component of $\mathbf{r}$ is approximately $\mathrm{Var}(r_i) \simeq \hat{\phi}(1 - h_{ii})$, where h_{ii} is the diagonal elements of the matrix $\widehat{H} = \widehat{\mathbf{W}}^{\frac{1}{2}}\mathbf{X}(\mathbf{X}^t\widehat{\mathbf{W}}\mathbf{X})^{-1}\mathbf{X}^t\widehat{\mathbf{W}}^{\frac{1}{2}}$. Thus the standard residual is defined by

$$r_i^* = \frac{r_i}{\sqrt{\hat{\phi}(1 - h_{ii})}}. \tag{8}$$

We conducted a simulation study to evaluate the performance of the proposed residuals considering the model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \text{with } i = 1, 2, \ldots, n,$$

where $\beta_0 = 1$ and $\beta_1 = 1$ were kept fixed, and the distribution of the explanatory variable x_i was fixed in $N(0, 1)$ in the course of the simulations. Eight scenarios were considered for the distribution of ϵ_i: slash-normal ($q = 3$), slash-Student-t ($q = 3, \nu = 3$), slash-Student-t ($q = 3, \nu = 4$), slash-Student-t ($q = 3, \nu = 6$), slash-normal ($q = 5$), slash-Student-t

Table 2
Mean ($\bar{r}$), standard deviation (S), skewness (b_1) and kurtosis (b_2) of the residuals for models with ($q = 3$).

Obs.	Slash-normal				Slash-Student-t ($\nu = 3$)				Slash-CN ($\lambda = 0.2, \sigma = 1.5$)				Slash-slash ($\gamma = 5$)			
	$\bar{r}$	S	b_1	b_2	$\bar{r}$	S	b_1	b_2	$\bar{r}$	S	b_1	b_2	$\bar{r}$	S	b_1	b_2
1	0.01	1.04	0.00	−1.09	−0.01	1.03	0.01	2.20	0.00	1.04	0.00	−1.19	0.02	1.04	−0.02	−1.21
2	0.01	1.04	0.00	−1.07	0.00	1.04	−0.04	2.34	−0.02	1.03	0.03	−1.16	0.02	1.04	−0.04	−1.20
3	0.01	1.04	0.00	−1.07	0.00	1.05	0.04	2.38	0.00	1.04	−0.02	−1.18	−0.01	1.04	0.01	−1.21
4	0.02	1.04	−0.03	−1.06	0.00	1.05	0.00	2.40	−0.01	1.04	0.01	−1.20	0.03	1.03	−0.03	−1.19
5	−0.02	1.05	0.02	−1.05	−0.01	1.09	−0.08	2.63	0.02	1.05	−0.01	−1.16	−0.02	1.05	0.01	−1.19
6	0.00	1.05	0.00	−1.05	0.02	1.09	0.10	2.49	0.01	1.04	0.00	−1.16	0.01	1.05	0.00	−1.20
7	0.00	1.05	0.00	−1.07	0.00	1.09	0.04	2.53	−0.01	1.04	0.01	−1.14	0.01	1.03	−0.01	−1.20
8	0.00	1.04	0.02	−1.04	0.00	1.06	−0.02	2.37	−0.01	1.03	0.01	−1.18	−0.01	1.04	0.01	−1.21
9	−0.01	1.05	0.01	−1.03	−0.01	1.09	−0.10	2.69	0.01	1.04	−0.01	−1.16	−0.01	1.05	0.00	−1.21
10	−0.01	1.03	0.02	−1.06	0.02	1.05	0.07	2.29	0.01	1.04	−0.01	−1.18	0.00	1.04	0.00	−1.21
11	−0.01	1.04	0.00	−1.06	−0.02	1.08	−0.08	2.32	0.00	1.04	0.00	−1.18	−0.01	1.04	0.02	−1.20
12	0.00	1.04	0.02	−1.06	0.00	1.07	−0.02	2.36	0.01	1.03	0.01	−1.18	0.00	1.04	0.00	−1.22
13	0.00	1.04	0.03	−1.05	0.00	1.06	0.00	2.48	0.00	1.04	−0.01	−1.18	−0.01	1.04	0.01	−1.21
14	0.00	1.04	0.00	−1.07	−0.01	1.06	−0.05	2.36	0.00	1.04	0.00	−1.17	0.00	1.03	−0.01	−1.20
15	0.01	1.04	−0.01	−1.04	0.02	1.06	0.05	2.36	−0.01	1.04	0.00	−1.19	0.01	1.04	−0.01	−1.23
16	0.00	1.05	−0.01	−1.08	0.02	1.05	0.02	2.40	−0.02	1.04	0.02	−1.20	0.01	1.04	−0.01	−1.21
17	0.01	1.04	−0.02	−1.05	−0.01	1.08	−0.07	2.37	0.00	1.04	0.01	−1.17	−0.02	1.04	0.02	−1.21
18	0.00	1.03	0.00	−1.05	−0.01	1.05	−0.08	2.53	0.00	1.04	−0.02	−1.19	0.00	1.03	0.01	−1.18
19	−0.01	1.03	0.01	−1.03	−0.02	1.06	0.01	2.28	0.01	1.04	0.01	−1.20	−0.01	1.05	0.01	−1.20
20	−0.01	1.03	0.02	−1.06	−0.02	1.05	0.04	2.29	0.00	1.04	0.00	−1.18	−0.01	1.03	0.02	−1.19

Table 3
Mean ($\bar{r}$), standard deviation (S), skewness (b_1) and kurtosis (b_2) of the residuals for models with ($q = 5$).

Obs.	Slash-normal				Slash-Student-t ($\nu = 3$)				Slash-CN ($\lambda = 0.2, \sigma = 1.5$)				Slash-slash ($\gamma = 5$)			
	$\bar{r}$	S	b_1	b_2	$\bar{r}$	S	b_1	b_2	$\bar{r}$	S	b_1	b_2	$\bar{r}$	S	b_1	b_2
1	0.01	1.05	0.00	−0.87	−0.01	1.04	0.00	1.74	0.00	1.04	0.00	−1.07	0.02	1.04	−0.02	−1.08
2	0.01	1.05	0.00	−0.85	0.00	1.06	−0.06	1.91	−0.03	1.03	0.03	−1.04	0.01	1.04	−0.04	−1.07
3	0.01	1.05	0.01	−0.84	0.00	1.06	0.06	1.91	0.00	1.04	−0.01	−1.05	−0.01	1.04	0.00	−1.07
4	0.02	1.04	−0.03	−0.84	0.00	1.06	−0.02	1.91	−0.01	1.04	0.01	−1.09	0.03	1.03	−0.03	−1.05
5	−0.02	1.05	0.01	−0.84	−0.01	1.10	−0.06	2.25	0.02	1.05	−0.01	−1.04	−0.02	1.05	0.00	−1.05
6	0.00	1.05	0.00	−0.83	0.01	1.09	0.07	2.04	0.00	1.05	0.00	−1.04	0.01	1.05	0.00	−1.07
7	0.00	1.05	0.00	−0.86	0.00	1.10	0.07	2.15	0.00	1.04	0.01	−1.02	0.01	1.03	−0.02	−1.06
8	0.00	1.04	0.02	−0.80	0.00	1.06	−0.04	1.90	−0.01	1.04	0.00	−1.05	−0.01	1.04	0.01	−1.07
9	−0.01	1.05	0.01	−0.82	0.00	1.10	−0.08	2.28	0.01	1.04	−0.01	−1.03	−0.01	1.05	0.00	−1.08
10	−0.01	1.04	0.02	−0.82	0.02	1.06	0.04	1.83	0.01	1.04	−0.01	−1.06	0.00	1.05	0.00	−1.06
11	−0.01	1.04	−0.01	−0.85	−0.02	1.09	−0.05	1.90	0.01	1.04	0.00	−1.07	−0.01	1.04	0.01	−1.06
12	0.00	1.04	0.03	−0.84	0.00	1.08	−0.03	1.92	0.01	1.03	0.00	−1.05	0.00	1.04	0.00	−1.09
13	0.01	1.05	0.03	−0.84	0.00	1.07	−0.01	2.05	0.00	1.04	−0.01	−1.05	−0.01	1.04	0.01	−1.07
14	0.00	1.05	0.00	−0.84	−0.01	1.07	−0.04	1.90	0.00	1.05	0.00	−1.05	0.00	1.03	−0.01	−1.06
15	0.01	1.04	0.00	−0.82	0.02	1.06	0.04	1.89	−0.01	1.04	−0.01	−1.06	0.01	1.04	−0.01	−1.10
16	0.00	1.05	−0.01	−0.85	0.02	1.06	0.02	1.90	−0.02	1.04	0.03	−1.08	0.01	1.04	−0.01	−1.07
17	0.01	1.05	−0.02	−0.82	0.00	1.08	−0.04	1.93	0.00	1.04	0.01	−1.05	−0.01	1.04	0.02	−1.07
18	0.00	1.04	−0.01	−0.84	−0.01	1.06	−0.07	2.06	0.00	1.05	−0.03	−1.06	0.00	1.04	0.01	−1.04
19	−0.01	1.04	0.01	−0.81	−0.02	1.07	0.01	1.89	0.01	1.04	0.00	−1.07	−0.01	1.05	0.01	−1.06
20	−0.01	1.04	0.02	−0.83	−0.02	1.05	0.04	1.77	0.00	1.04	0.00	−1.05	−0.01	1.03	0.02	−1.05

($q = 5, \nu = 3$), slash-Student-t ($q = 5, \nu = 4$) and slash-Student-t ($q = 5, \nu = 6$). We generated 10,000 Monte Carlo replicates of the response variable for each of the eight scenario considered and the sample size was fixed at $n = 20$. Statistical measures of simulated residuals were calculated and the results are shown in Tables 2 and 3. In all scenarios the proposed residuals had a empirical mean and skewness approximately zero and standard deviation approximately one. For scenarios where ϵ_i had slash-normal distribution the proposed residual presented a negative excess of kurtosis, indicating a platykurtic distribution, whereas in scenarios where ϵ_i had slash-Student-t distribution, the proposed residual shown a positive excess of kurtosis, indicating a leptokurtic distribution.

3. Diagnostic analysis

The methodology for diagnostic analysis through local influence for the family of linear models with elliptical errors can be found in the works by Galea et al. (1997, 2000, 2002, 2003). Galea et al. (2005) presented diagnostic methods for the class of symmetric nonlinear models. We develop the methodology of diagnostic analysis based on the approach of local influence for the class of slash-elliptical regression models and a measure of leverage is presented.

Assuming that the fitted model is correct, the diagnostic analysis allows to check if there are observations that play a strong influence on the fit of the model or observations with outliers in the regressors. An observation of the data set is said to be influential observation, if under small perturbation occur disproportionate variations in the estimates of the fitted model.

3.1. Generalized leverage

The measure of generalized leverage defined by $GL_{ij} = \frac{\partial \hat{y}_i}{\partial y_j}$ reflects the rate of instantaneous change in the ith predicted value $\hat{y}_i$, when the jth response variable y_j is increased by an infinitesimal. The generalized leverage GL_{ii} is proposed by Wei et al. (1998) as the measure of greatest influence of y_i on its own value set. High leverage suggests that the ith observation can have outliers in the regressors.

Denote the maximum likelihood estimator of θ by $\hat{\theta}(\mathbf{y})$, and μ by $\hat{\mu} = \mathbf{X}\hat{\beta}$, then Wei et al. (1998) showed that the matrix of generalized leverage can be expressed as

$$GL(\hat{\theta}) = D_\theta \ddot{L}_{\theta\theta}^{-1} \ddot{J}_{\theta\mathbf{y}}, \tag{9}$$

where $D_\theta = \frac{\partial \mu}{\partial \theta^t} = \frac{\partial \mathbf{X}\beta}{\partial \theta^t} = [\mathbf{X}, 0]$ is a $(n \times (p+1))$ matrix and $\ddot{J}_{\theta\mathbf{y}} = \frac{\partial^2 L(\theta)}{\partial \theta \partial \mathbf{y}^t}$ is a $((p+1) \times n)$ matrix evaluated at $\theta = \hat{\theta}$. Applying this methodology to the class of linear model with slash-elliptical errors, we have that the matrix $\ddot{J}_{\hat{\theta}\mathbf{y}} = (\ddot{J}^t_{\hat{\beta}\mathbf{y}}, \ddot{J}^t_{\hat{\phi}\mathbf{y}})^t$, where $\ddot{J}_{\hat{\beta}\mathbf{y}} = \frac{1}{\hat{\phi}} \mathbf{X}^t \mathbf{D}(\hat{c})$ and $\ddot{J}_{\hat{\phi}\mathbf{y}} = -\frac{1}{2\hat{\phi}^{3/2}} \mathbf{v}(\hat{e})$. Thus, the generalized leverage matrix is given by

$$GL(\hat{\theta}) = \mathbf{X}(\mathbf{X}^t \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^t \widehat{\mathbf{W}}. \tag{10}$$

The matrix $GL(\hat{\theta})$ is not symmetric, but has the property of idempotence and the following relationship with the matrix $\widehat{H}$:

$$\widehat{H} = \widehat{\mathbf{W}}^{\frac{1}{2}} GL(\hat{\theta}) \widehat{\mathbf{W}}^{-\frac{1}{2}} = \widehat{\mathbf{W}}^{\frac{1}{2}} \mathbf{X}(\mathbf{X}^t \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^t \widehat{\mathbf{W}}^{\frac{1}{2}}.$$

A graph of the diagonal elements of $GL(\hat{\theta})$ versus index can determine points with high influence on their own predicted value. Because $GL(\hat{\theta})$ is idempotent, we have that $\mathrm{rank}(GL(\hat{\theta})) = \mathrm{tr}(GL(\hat{\theta})) = p$ and we can consider the criterion $GL_{ii} \geq \frac{2p}{n}$ to identify whether the ith observation is a point of leverage (Cook and Weisberg, 1982).

3.2. Local influence on likelihood displacement

The analysis of local influence is a diagnostic method proposed by Cook (1986) that evaluates the effect of small perturbations in data or model through a measure of influence. Let $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ be $(n \times 1)$ perturbations vector, restricted to some open set Ω and ω_0 non-perturbation vector. In practice, we want to compare $\widehat{\theta}$ and $\widehat{\theta_\omega}$ by a measure of influence when ω varies on Ω. Considerable distances between these measures may indicate the presence of influential points. Among the measures of local influence, we emphasize the displacement of the likelihood proposed by Cook (1986) and a distance measure based on the residual of Pearson proposed by Thomas and Cook (1990).

Let $L(\theta)$ and $L(\theta_\omega)$ denote functions of log-likelihood of postulated and perturbed models, respectively, then the likelihood displacement is defined by

$$LD(\omega) = 2\{L(\widehat{\theta}) - L(\widehat{\theta_\omega})\}.$$

We note that if $\omega = \omega_0$, then $LD(\omega) = 0$, and if $\omega \neq \omega_0$, then $LD(\omega) \geq 0$, that is, ω_0 is a local minimum point of $LD(\omega)$. The analysis of local influence based on likelihood displacement is to determine how the surface $\alpha(\omega) = (\omega^t, LD(\omega))^t$, $\omega \in \Omega$, deviates from its tangent plane around ω_0 by analyzing the normal curvature of the surface $\alpha(\omega)$ in some arbitrary direction ℓ, with $\|\ell\| = 1$ (Cook and Weisberg, 1982). We can show that the normal curvature in direction ℓ may be expressed by

$$C_\ell = 2|\ell^t \ddot{F} \ell|,$$

where $\ddot{F} = \Delta^t \ddot{L}_{\theta\theta}^{-1} \Delta$ and $\Delta = \frac{\partial^2 L(\theta_\omega)}{\partial \theta \partial \omega^t}$ evaluated at $\omega = \omega_0$ and $\theta = \widehat{\theta}$.

If the interest is conducting an analysis of influence for the parameters β and ϕ separately, the $\ddot{F}$ matrix can be adapted to exclude from $\ddot{L}_{\theta\theta}$, the influence of other parameters. Thus, the normal curvature in the direction d for the parameters β and ϕ are:

$$C_\ell(\beta) = 2|\ell^t \Delta^t (\ddot{L}_{\theta\theta}^{-1} - L_1) \Delta \ell|$$

and

$$C_\ell(\phi) = 2|\ell^t \Delta^t (\ddot{L}_{\theta\theta}^{-1} - L_2) \Delta \ell|,$$

respectively, where $L_1 = \begin{bmatrix} 0 & 0 \\ 0 & \ddot{L}_{\phi\phi}^{-1} \end{bmatrix}$ and $L_2 = \begin{bmatrix} \ddot{L}_{\beta\beta}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$. The index graph of $d_{\max}$, the eigenvector corresponding to higher absolute eigenvalue of $\ddot{F}$, can reveal the most influential observations in $\hat{\theta}$. Another possibility for analysis was proposed by Lesaffre and Verbeke (1998), which consists of the construction of the graph of C_i's obtained from the calculation of C_{d_i} for the vector d_i formed by zeros with one in the ith position for each $i = 1, 2, \ldots, n$. Lesaffre and Verbeke (1998) also suggested that the observations such that $C_i > 2\overline{C}$, where $\overline{C} = \frac{\sum_{i=1}^n C_i}{n}$, can be influential points.

3.2.1. Scale perturbation

Assume that $y_i \sim SEL(\mathbf{x}_i^t\beta, \phi/\omega_i, q; g(\cdot))$, namely that the linear model $y_i = \mathbf{x}_i^t\beta + \epsilon_i$ is heteroscedastic with perturbation of the scale parameter $\phi_i = \phi/\omega_i$, for $\omega_i > 0$ and $i = 1, 2, \ldots, n$. For $0 < \omega_i < 1$, there is an inflating ϕ and when $\omega_i > 1$, there is a reduction of ϕ. When $\omega_i = 1$ then $\phi_i = \phi$ and the perturbed model reduces to the postulated. The log-likelihood of the perturbed model is given by

$$L(\theta, \omega) = -\frac{1}{2}\sum_{i=1}^{n} \log\left(\frac{\phi}{\omega_i}\right) + \sum_{i \in A} a(z_{i\omega}) + K,$$

where $z_{i\omega} = \sqrt{\omega_i}\frac{(y_i - \mathbf{x}_i^t\beta)}{\sqrt{\phi}} = \sqrt{\omega_i}z_i$. The perturbation matrix Δ with perturbation in scale is expressed by:

$$\hat{\Delta} = \begin{pmatrix} -\dfrac{1}{2\sqrt{\hat{\phi}}}\mathbf{X}^t\mathbf{D}(\hat{e}) \\ -\dfrac{1}{4\hat{\phi}}\hat{\mathbf{z}}^t\mathbf{D}(\hat{e}) \end{pmatrix},$$

where $\mathbf{D}(\hat{e}) = \mathrm{diag}(e(\hat{z}_1), \ldots, e(\hat{z}_n))$.

3.2.2. Cases-weight perturbation

Consider a scheme with perturbation in the ith case in which the log-likelihood function of θ, the perturbed model is expressed as

$$L(\theta, \omega) = \sum_{i=1}^{n} \omega_i \log(f_{y_i}(y_i)),$$

where $0 \le \omega_i \le 1$. For this perturbation scheme, the matrix Δ is given by

$$\hat{\Delta} = \begin{pmatrix} -\dfrac{1}{\sqrt{\hat{\phi}}}\mathbf{X}^t\mathbf{D}(\hat{a}') \\ -\dfrac{1}{2\hat{\phi}}\mathbf{v}^t(\hat{o}) \end{pmatrix},$$

where $a'(\hat{z}_i) = \frac{2\hat{z}_i^2 g(\hat{z}_i^2)}{H(\hat{z}_i^2)} - \frac{q+1}{\hat{z}_i}$, $\mathbf{D}(\hat{a}') = \mathrm{diag}(a'(\hat{z}_1), \ldots, a'(\hat{z}_n))$, $o(\hat{z}_i) = 1 + \hat{z}_i a'(\hat{z}_i)$ and $\mathbf{v}(\hat{o}) = (o(\hat{z}_1), \ldots, o(\hat{z}_n))^t$.

3.3. Local influence on the prediction

Consider the analysis of the effect of small perturbations on the prediction of a particular $(p \times 1)$ vector $\mathbf{v}$ of values of the regressors. Let $\widehat{\mu}(\mathbf{v}) = \mathbf{v}^t\widehat{\beta}$ and $\widehat{\mu}(\mathbf{v}, \omega) = \mathbf{v}^t\widehat{\beta}_\omega$ denote the predictions in the postulated and perturbed models, respectively, where $\widehat{\beta}_\omega$ is the vector of maximum likelihood estimates of θ in the perturbed model. A measure of local influence in the prediction proposed by Thomas and Cook (1990) based on the Pearson residual is defined as

$$f(\mathbf{v}, \omega) = \{\widehat{\mu}(\mathbf{v}) - \widehat{\mu}(\mathbf{v}, \omega)\}^2. \tag{11}$$

From the study of the behavior of the surface $\delta(\omega) = (\omega^t, f(\mathbf{v}, \omega))^t$ at around ω_0, one can show that the normal curvature of this surface in the direction d is $C_d(\mathbf{v}) = 2|d^t\ddot{F}d|$, where

$$\ddot{F} = \left.\frac{\partial^2 f(\mathbf{v}, \omega)}{\partial\omega\partial\omega^t}\right|_{\substack{\beta=\widehat{\beta} \\ \omega=\omega_0}} = \Delta^t(\ddot{L}_{\beta\beta}^{-1}\mathbf{v}\mathbf{v}^t\ddot{L}_{\beta\beta}^{-1})\Delta.$$

In this case, the vector $\ell_{\max}(\mathbf{v})$ for a given vector of values of the regressors, $\mathbf{v}$, is summarized in the expression

$$\ell_{\max}(\mathbf{v}) \propto \Delta^t\ddot{L}_{\beta\beta}^{-1}\mathbf{v}. \tag{12}$$

A proposed analysis is to consider $\mathbf{v} = \mathbf{x}_i$ for each observation i with $i = 1, 2, \ldots, n$, and calculate the largest value of the ith vector $\ell_{\max}(\mathbf{x}_i)$, which will be denoted by $\ell_{\max_i}(\mathbf{x}_i)$ (Galea et al., 2003). From the index graph of $(\ell_{\max_1}(\mathbf{x}_1), \ell_{\max_2}(\mathbf{x}_2), \ldots, \ell_{\max_n}(\mathbf{x}_n))^t$ we can identify there observations that have substantial influence on $\hat{y}$. Another proposal analysis is to build the index graph of the

$$\mathbf{C}_{\max}(\mathbf{x}_i) = \ell_{\max}(\mathbf{x}_i)^t\ell_{\max}(\mathbf{x}_i) = \mathbf{x}_i^t\ddot{L}_{\beta\beta}^{-1}\Delta\Delta^t\ddot{L}_{\beta\beta}^{-1}\mathbf{x}_i,$$

for each $\mathbf{x}_i$ with $i = 1, 2, \ldots, n$.

3.3.1. Response perturbation

Suppose that the ith response is perturbed additively, where $y_{i\omega} = y_i + s\omega_i$, and s is the estimated standard deviation of y, in such a way that $\omega_i = 0$, the perturbed model is equal to the postulated model. The log-likelihood function for the perturbed model is given by

$$L(\theta, \omega) = -\frac{n}{2}\log(\phi) + \sum_{i \in A} a(z_{i\omega}) + K,$$

where $z_{i\omega} = \frac{(y_i + s\omega_i - \mathbf{x}_i^t\beta)}{\sqrt{\phi}} = z_i + \frac{s\omega_i}{\sqrt{\phi}}$.

For the additive perturbation scheme on response, the obtained matrix Δ is given by

$$\hat{\Delta} = \frac{s}{\hat{\phi}}\mathbf{X}^t\mathbf{D}(\hat{c}).$$

3.3.2. Regressors perturbation

Now suppose that the ith observation of a particular explanatory variable t, $t = 2, 3, \ldots, p$, is perturbed additively by the scheme $x_{it\omega} = x_{it} + \omega_i$, which can also be expressed in vector form by $\mathbf{x}_{i\omega} = \mathbf{x}_i + \omega_i\mathbf{s}_t$, where $\mathbf{s}_t$ is a $(p \times 1)$ vector of zeros with one in the tth position. Also, in this case, when $\omega_i = 0$, the perturbed model is equal to the postulated model. The log-likelihood function for the perturbed model with additive perturbation in the tth explanatory variable is expressed by

$$L(\theta, \omega) = -\frac{n}{2}\log(\phi) + \sum_{i \in A} a(z_{i\omega}) + K,$$

where $z_{i\omega} = \frac{(y_i - \mathbf{x}_i^t\beta - \omega_i\beta_t)}{\sqrt{\phi}} = z_i - \frac{\omega_i\beta_t}{\sqrt{\phi}}$.

For the additive perturbation scheme in the tth explanatory variable, the Δ matrix obtained is given by

$$\hat{\Delta} = -\frac{\hat{\beta}_t}{\hat{\phi}}\mathbf{X}^t\mathbf{D}(\hat{c}) - \frac{1}{\sqrt{\hat{\phi}}}\mathbf{s}_t\mathbf{v}^t(\hat{a}'),$$

where $\mathbf{v}(\hat{a}') = (a'(\hat{z}_1), \ldots, a'(\hat{z}_n))^t$.

4. Application

The data set refers to the water salinity and river discharge taken in North Carolina's Pamlico Sound (Ruppert and Carroll, 1980). The values are biweekly averages of water salinity at time period i (y_i), salinity lagged two weeks (x_{1i}), trend one of the six biweekly periods in March–May (x_{2i}), and river discharge at time i (x_{3i}).

The data set of water salinity is frequently used to exemplify analysis techniques for symmetric and elliptic data. According to Aiktson (1985), the observations #16 and #5 are influential, under the assumption of normal errors. Davison and Tsai (1992) found #16, #5 and #3 points as influential observations in the Student-t $(\nu = 3)$ model. Galea et al. (1997), Liu (2000), and Galea et al. (2003) found #16 and #5 points as influential observations in elliptical models about the perturbation schemes in scale, response and regressors. We propose the initial model given by

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i, \quad i = 1, \ldots, 28,$$

where ϵ_i have distribution on elliptical class (normal and Student-t) and slash-elliptical class (slash-normal, slash-Student-t, slash-CN and slash-slash) for errors. The *BIC* criterion was used for selecting extra parameters. The selected models were: normal, Student-t $(\nu = 3)$, slash-normal $(q = 3)$, slash-Student-t $(q = 3, \nu = 2)$, slash-CN $(q = 5, \lambda = 0.5, \sigma = 4)$ and slash-slash $(q = 3, \gamma = 3)$ and results omitted here showed that the variable x_2 is not significant for all select models.

In Table 4 are presented the estimates and their asymptotic standard errors of the parameters(in parentheses) of the six models without the variable x_2. The asymptotic standard error values were calculated based on the observed information matrix and the p-values were based on the likelihood ratio test. Now, all variables are significant in the considered models and it was further observed that the standard errors of the slash-Student-t $(q = 3, \nu = 2)$ is smaller than in others. The lowest BIC is the model slash-CN $(q = 5, \lambda = 0.5, \sigma = 4)$ $(BIC = 100.11)$ followed slash-Student-t model $(BIC = 102.56)$.

A diagnostic analysis based on the proposed measures was performed. Index plot generalized leverage highlights #5 and #16 as leverage observation (see, Figs. 1(b), (c), (e), (f) and Figs. 1(a), (d), respectively). Index plot C_i under cases-weight scheme perturbation for six fitted models is shown seen in Fig. 2. The slash-Student-t model shows less sensitive to case #16 (see Table 5 and Fig. 2(d)) while the other models #16 is highlighted as influential point.

The Figs. 3 and 4 show the index $\ell_{\max}$ plot on response perturbation and the plot $|\mathbf{C}_{\max}|$ versus x_3 for explanatory variable x_3 perturbation scheme under local influence on predictors, respectively. We found that the observations #16 and #5 were also influential in the slash-elliptical models as well as in elliptical models. Thus, we can conclude that the observations #5 and #16 have substantial influence on the own predicted value and a small perturbation in the value of x_3 (for large values) leads to a substantial change in the prediction.

Table 4
Results from the six fitted models considered without x_2 to the water salinity data set.

		β_0	β_1	β_3	ϕ	BIC
Normal	Estimate	9.32	0.78	−0.29	1.52	104.49
	Std. error	2.44	0.08	0.09	0.41	
	p-value	<0.001	<0.001	0.001		
Student-t ($\nu = 3$)	Estimate	13.65	0.74	−0.45	0.67	103.01
	Std. error	1.99	0.06	0.07	0.25	
	p-value	<0.001	<0.001	<0.001		
Slash-normal ($q = 3$)	Estimate	13.13	0.75	−0.44	0.59	104.17
	Std. error	3.255	0.070	0.122	0.219	
	p-value	<0.001	<0.001	<0.001		
Slash-Student-t ($q = 3$, $\nu = 2$)	Estimate	14.27	0.73	−0.47	0.23	102.56
	Std. error	0.061	0.003	0.002	0.002	
	p-value	<0.001	<0.001	<0.001		
Slash-CN ($q = 5$, $\lambda = 0.5$, $\sigma = 4$)	Estimate	13.490	0.73	−0.43	0.11	100.11
	Std. error	1.516	0.040	0.056	0.044	
	p-value	<0.001	<0.001	<0.001		
Slash-slash ($q = 3$, $\gamma = 3$)	Estimate	13.25	0.75	−0.44	0.40	103.85
	Std. error	3.053	0.067	0.114	0.154	
	p-value	<0.001	<0.001	<0.001		

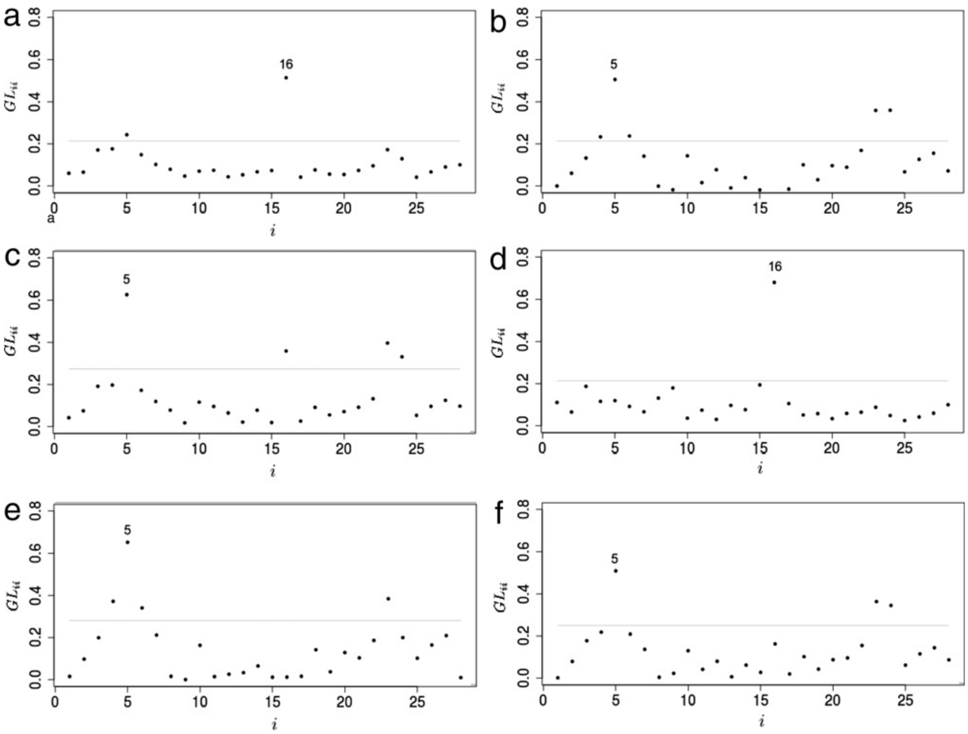

Fig. 1. Index plot generalized leverage: (a) normal, (b) Student-t ($\nu = 3$), (c) slash-normal ($q = 3$), (d) slash-Student-t ($q = 3$, $\nu = 2$), (e) slash-CN ($q = 5$, $\lambda = 0.5$, $\sigma = 4$) and (f) slash-slash ($q = 3$, $\gamma = 3$).

Table 5
Variation rate of the estimates of the parameters β_0, β_1, β_3, and ϕ, when dropout the influential points.

Model	Influence points	β_0	β_1	β_3	ϕ
Normal	5, 16	101.5	−7.3	134.8	−36.5
Student-t ($\nu = 3$)	5, 16	33.1	−3.6	40.9	−24.4
Slash-normal ($q = 3$)	5, 16	39.4	−3.6	49.1	−22.9
Slash-Student-t ($q = 3$, $\nu = 2$)	5, 16	24.2	−2.0	31.0	−17.5
Slash-CN ($q = 5$, $\lambda = 0.5$, $\sigma = 4$)	5, 16	31.1	−3.0	40.5	−21.3
Slash-slash ($q = 3$, $\gamma = 3$)	5, 16	37.5	−3.6	46.4	−23.8

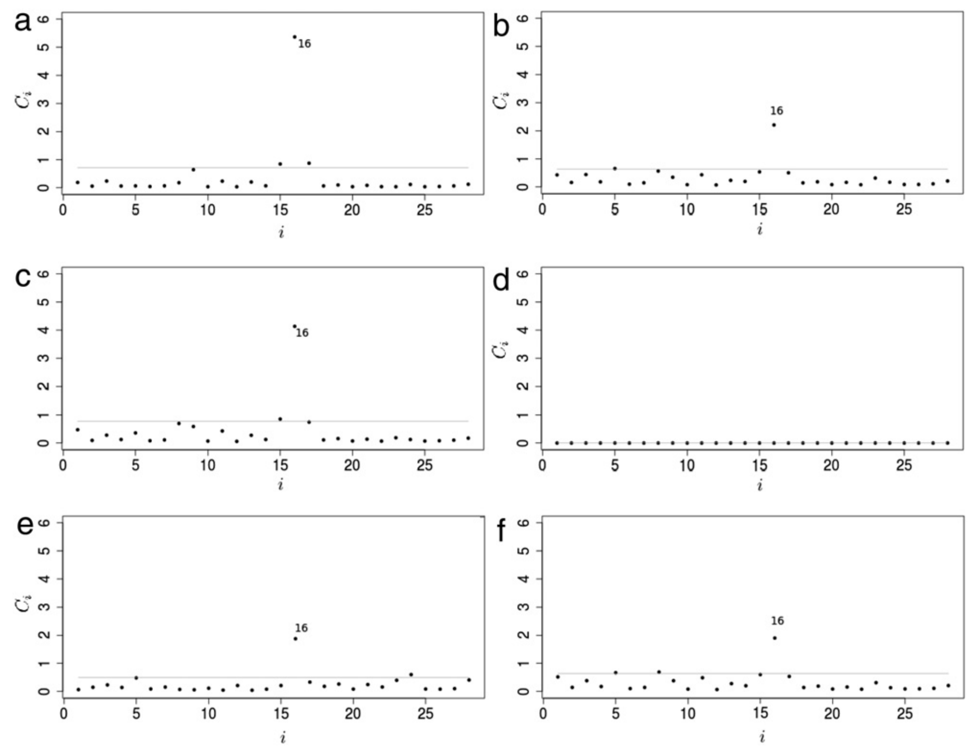

Fig. 2. Index plot C_i under case-perturbation scheme (local influence on likelihood displacement): (a) normal, (b) Student-t ($\nu = 3$), (c) slash-normal ($q = 3$), (d) slash-Student-t ($q = 3$, $\nu = 2$), (e) slash-CN ($q = 5$, $\lambda = 0.5$, $\sigma = 4$) and (f) slash-slash ($q = 3$, $\gamma = 3$).

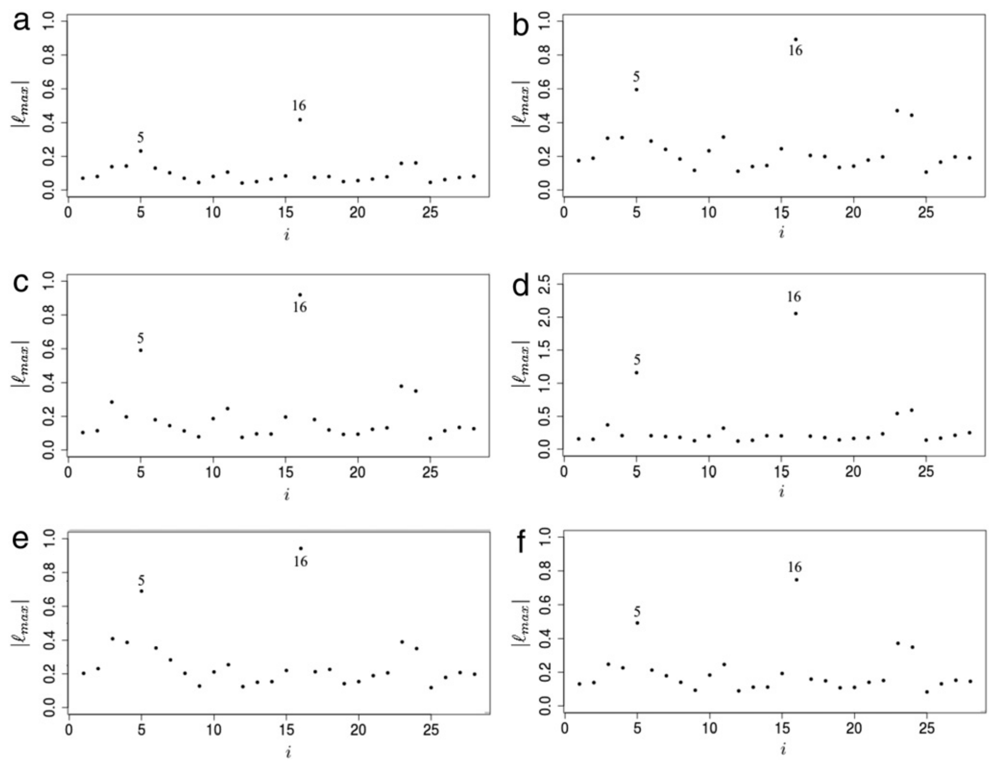

Fig. 3. Index plot $|\ell_{\max}|$ under response perturbation (local influence on the prediction): (a) normal, (b) Student-t ($\nu = 3$), (c) slash-normal ($q = 3$), (d) slash-Student-t ($q = 3$, $\nu = 2$), (e) slash-CN ($q = 5$, $\lambda = 0.5$, $\sigma = 4$) and (f) slash-slash ($q = 3$, $\gamma = 3$).

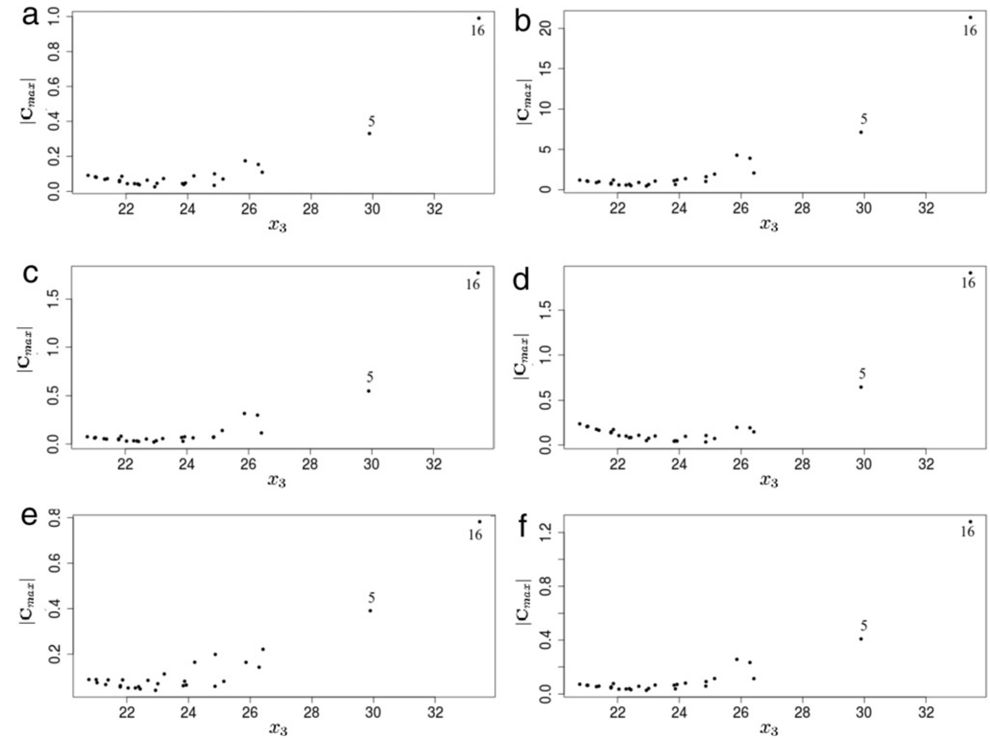

Fig. 4. Plot $|\mathbf{C}_{\max}|$ versus x_3 under explanatory variable x_3 perturbation scheme (local influence on the prediction): (a) normal, (b) Student-t $(\nu = 3)$, (c) slash-normal $(q = 3)$, (d) slash-Student-t $(q = 3, \nu = 2)$, (e) slash-CN $(q = 5, \lambda = 0.5, \sigma = 4)$ and (f) slash-slash $(q = 3, \gamma = 3)$.

In order to measure the variation of the parameter estimates when removing the influential points, we calculated the rates $r(\beta_i) = \{(\hat{\beta}_i^{(j)} - \hat{\beta}_i)/\hat{\beta}_i\} \times 100$, where $\hat{\beta}_i$ is the estimate of β_i model with all observations, and $\hat{\beta}_i^{(j)}$ in the model without the influential points. The variation rates of the estimated parameters can be found in Table 5. Note that, although the present results of slash-Student-t $(q = 3, \nu = 2)$ model, the diagnostics plots($|\ell_{\max}|$ and leverage) have better behavior than the other models. Also, we observed that the smallest variation on estimates of the parameters when removed from the influential points were in slash-Student-t $(q = 3, \nu = 2)$ model.

5. Concluding remarks

In this study, we proposed the linear regression model with slash-elliptical errors and parameter q known or fixed. We developed an inferential methodology based on estimated asymptotic standard errors obtained by observed information matrix. In the case that the parameter q is not known, we suggest using the *AIC* or *BIC* criteria to select the slash-elliptical model with the q corresponding to the best fit. We presented the asymptotic tests (likelihood ratio, Wald and score) that can be used to assess the inclusion or exclusion of explanatory variables.

We proposed a standardized residual and based on simulation studies, we concluded that the proposed residual has mean and standard error close to zero and one, respectively and is approximately symmetrical. However, the kurtosis is not three as in normal distribution and depend on extra parameters which leads us to believe that the distribution of this residual is not normal.

In the diagnostic analysis, we verified that generalized leverage matrix $GL(\hat{\theta})$ is idempotent and therefore we can consider the criterion $GL_{ii} \geq \frac{2p}{n}$ to identify leverage points. In the analysis of local influence, we considered the measures: likelihood displacement under the perturbation scheme in scale and cases-weight; and a distance measure based on the Pearson residual under the perturbation scheme in response variable and regressors.

We applied the proposed methodology to the data set of salinity of water, comparing the elliptical models (normal and Student-t $(\nu = 3)$) and slash-elliptical (slash-normal $(q = 3)$, slash-Student-t $(q = 3, \nu = 2)$, slash-CN $(q = 5, \lambda = 0.5, \sigma = 4)$ and slash-slash $(q = 3, \gamma = 3)$). We observed that the standard errors of the slash-Student-t $(q = 3, \nu = 2)$ model is smaller than in other models considered. Observations #16 and #5 are influential points in the slash-elliptical models, as well as in elliptical models. Although the #16 and #5 observation are also highlighted in slash-Student-t $(q = 3, \nu = 2)$ model, this model presents the smallest variation on estimates of the parameters when these points are removed. Thus, we can conclude that the slash-Student-t $(q = 3, \nu = 2)$ model is more stable when we dropout influential points, than other models.

Acknowledgments

This study was partially supported by CNPq and Capes, Brazil. The authors are grateful to referees and associate editor for helpful comments and suggestions.

References

Aiktson, A.C., 1985. Plots, Transformations, and Regression. Clarendon, Oxford.
Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control 19, 716–723.
Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., Tukey, J.W., 1972. Robust Estimates of Location: Survey and Advances. Princeton University Press, Princeton.
Arslan, O., 2008. An alternative multivariate skew-slash distribution. Statist. Probab. Lett. 78, 2756–2761.
Arslan, O., Genç, A.I., 2009. A generalization of the multivariate slash distribution. J. Statist. Plann. Inference 139, 1164–1170.
Chmielewski, M.A., 1981. Elliptically symmetric distributions: a review and bibliography. Int. Statist. Rev. 49, 67–74.
Cook, R.D., 1986. Assessment of local influence. J. R. Stat. Soc. Ser. B 48, 133–169 (with discussion).
Cook, R.D., Weisberg, S., 1982. Residuals and Influence in Regression. Chapman and Hall, London.
Cysneiros, F.J.A., Paula, G.A., 2005. Restricted methods in symmetrical linear regression models. Comput. Statist. Data Anal. 49, 689–708.
Davison, A.C., Tsai, C.L., 1992. Regression model diagnostics. Int. Statist. Rev. 60, 337–353.
Doornik, J.A., 2007. Object-Oriented Matrix Programming Using Ox. Timberlake Consultants Press, Oxford, London.
Ferreira, C.S., Bolfarine, H., Lachos, V.H., 2011. Skew scale mixtures of normal distributions: properties and estimation. Stat. Methodol. 8, 154–171.
Galea, M., Bolfarine, H., Vilca-Labra, F., 2002. Influence diagnostics of the structural error-in-variables model under the Student-t distribution. J. Appl. Stat. 29, 1191–1204.
Galea, M., Paula, G.A., Bolfarine, H., 1997. Local influence in elliptical linear regression models. The Statistician 46, 71–79.
Galea, M., Paula, G.A., Cysneiros, F.J.A., 2005. On diagnostic in symmetrical nonlinear models. Statist. Probab. Lett. 73, 459–467.
Galea, M., Paula, G.A., Uribe-Opazo, M., 2003. On influence diagnostic in the univariate elliptical linear regression models. Statist. Papers 44, 23–45.
Galea, M., Riquelme, M., Paula, G.A., 2000. Diagnostics methods in elliptical linear regression models. Braz. J. Probab. Stat. 14, 167–184.
Genç, A.I., 2007. A generalization of the univariate slash by a scale-mixtured exponential power distribution. Comm. Statist. Simulation Comput. 36, 937–947.
Gómez, H.W., Quintana, F.A., Torres, F.J., 2007. A new family of slash-distributions with elliptical contours. Statist. Probab. Lett. 77, 717–725.
Gómez, H.W., Venegas, O., 2007. Erratum to: "A new family of slash-distributions with elliptical contours" [Statist. Probab. Lett. 77 (2007) 717–725]. Statist. Probab. Lett. 78, 2273–2274.
Gross, A.M., 1973. A Monte Carlo swindle for estimators of location. J. Roy. Statist. Soc. Ser. C 22, 347–353.
Jørgensen, B., 1984. The delta algorithm and GLIM. Int. Statist. Rev. 52, 283–300.
Kafadar, K., 1982. A biweight approach to the one-sample problem. J. Amer. Statist. Assoc. 77, 416–424.
Kuha, J., 2004. AIC and BIC: comparisons of assumptions and performance. Sociol. Methods Res. 33, 188–229.
Lachos, V., Montenegro, L., Bolfarine, H., 2008. Inference and local influence assessment in skew-normal null measurement error models. J. Stat. Comput. Simul. 78, 395–415.
Lange, K., Sinsheimer, J.S., 1993. Normal/independent distributions and their applications in robust regression. J. Comput. Graph. Statist. 2, 175–198.
Lesaffre, E., Verbeke, G., 1998. Local influence in linear mixed models. Biometrics 54, 570–582.
Li, B., 2001. Sensitivity of Rao's score test, the Wald test and the likelihood ratio test to nuisance parameters. J. Statist. Plann. Inference 97, 57–66.
Liu, S., 2000. On local influence for elliptical linear models. Statist. Papers 41, 211–224.
Morgenthaler, S., Tukey, J.W., 1991. Configural Polysampling: A Route to Practical Robustness. Wiley, New York.
Mosteller, F., Tukey, J.W., 1977. Data Analysis and Regression. Addison-Wesley, Universidade de Michigan.
Paula, G.A., Medeiros, M.J., Vilca-Labra, F.E., 2009. Influence diagnostics for linear models with first autoregressive elliptical errors. Statist. Probab. Lett. 79, 339–346.
Rogers, W.H., Tukey, J.W., 1972. Understanding some long-tailed symmetrical distributions. Statist. Neerlandica 26, 211–226.
Ruppert, D., Carroll, R.J., 1980. Trimmed least squares estimation in the linear model. J. Amer. Statist. Assoc. 75, 828–838.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Statist. 6, 461–464.
Thomas, W., Cook, R.D., 1990. Assessing influence on predictions from generalized linear models. Technometrics 32, 59–65.
Wang, J., Genton, M.G., 2006. The multivariate skew-slash distribution. J. Statist. Plann. Inference 136, 209–220.
Wei, B.C., Hu, Y.Q., Fung, W.K., 1998. Generalized leverage and its applications. Scand. J. Statist. 25, 647–656.