RT-MAE 9928

DIAGNOSTIC METHODS IN ELLIPTICAL LINEAR REGRESSION MODELS

by

Manuel Galea, Marco Riquelme
and
Gilberto A. Paula

Palavras-Chave: Elliptical distributions; Influence diagnostic; Likelihood displacement; Local influence; Residuals.
Classificação AMS: 62J02, 62J20, 62F35.
(AMS Classification)

- Dezembro de 1999 -

Diagnostic Methods in Elliptical Linear Regression Models

Manuel Galea and Marco Riquelme
Universidad de Valparaíso, Chile
and
Gilberto A. Paula[1]
Universidade de São Paulo, Brazil

Summary

We discuss in this paper the development of various diagnostic methods in elliptical linear regression models. We show, in particular, that the distribution of some usual standardized residuals is invariant in the class of elliptical models. This invariance is also verified for some influence measures of dropping observations, such as the Cook distance. We also discuss the computation of the likelihood displacement as well as the normal curvature in the local influence method. An example with real data is given for illustration.

Key words: Elliptical distributions; Influence diagnostic; Likelihood displacement; Local influence; Residuals.

1 Introduction

Diagnostic methods for the normal linear regression model have been largely investigated in the Statistical literature (see, for instance, Belsley et al., 1980; Cook and Weisberg, 1982;

[1]Address for correspondence: Instituto de Matemática e Estatística, USP - Caixa Postal 66281 (Ag. Cidade de São Paulo), 05315-970 São Paulo - SP - Brazil.
e-mail: giapaula@ime.usp.br

Atkinson, 1985 and Chatterjce and Hadi, 1988). The majority of the works have given emphasis in studying the effect of eliminating observations on same key results from the fitted model, particularly the parameter estimates. Alternatively, Cook (1986) has proposed an interesting method, named local influence, to assess the effect of small perturbations in the model on the parameter estimates. Several authors have extended the local influence method to various regression models. Some references are Beckman et al.(1987), Lawrence (1988), Thomas and Cook (1990), Tsai and Wu (1992), Paula (1993), Kim (1995), Galea at al. (1997), Fung and Kwan (1997) among others.

The aim of this paper is to discuss the development of some traditional diagnostic methods, such as the deletion of individual observations, residual analysis and local influence in elliptical linear regression models. We show that if the interest is only on the coefficients the diagnostic graphics are in general invariant with the error distribution so that the well known graphics developed for the normal linear case can be applied. However, when the interest is also on the dispersion the diagnostic graphics depend on the error distribution. In the application we compare the behavior of some graphics for two particular models, Student-t and power exponential.

2 Elliptical linear regression model

The class of elliptical distributions has received an increasing attention in the Statistical literature (see, for instance, Fang et al, 1990; Fang and Zhang, 1990; Fang and Anderson, 1990; Gupta and Varga, 1993; Arellano, 1994 and Leiva, 1998). We say that an $(n \times 1)$ random vector $\mathbf{Y}$ has an elliptical distribution with an $(n \times 1)$ position parameter $\boldsymbol{\mu}$ and an $(n \times n)$ scale matrix $\boldsymbol{\Sigma}$ if its density function is expressed as

$$f_{\mathbf{Y}}(\mathbf{y}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}-\boldsymbol{\mu})], \tag{1}$$

$\mathbf{y} \in \mathcal{R}^n$, where the function $g : \mathcal{R} \rightarrow [0, \infty)$ is such that $\int_0^\infty u^{n-1} g(u^2) du < \infty$. The function $g(\cdot)$ is typically known as the density generator. For a vector $\mathbf{Y}$ distributed according to the density (1) we use the notation $\mathbf{Y} \sim El_n(\boldsymbol{\mu}, \Sigma, g)$, or, simply, $\mathbf{Y} \sim El_n(\boldsymbol{\mu}, \Sigma)$. In the case where $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma = \mathbf{I}$, we obtain the spherical family of densities. This class of symmetric distributions includes the normal, Student-t, contaminated normal and logistic (both, univariate and multivariate), among others, as considered, for example, by Fang et al. (1990). Table 1 below, taken from Fang et al. (1990), reports examples of distributions in the elliptical family. The notation c_1, c_2, c_3, c_4 and c_5 is used to denote normalizing constants.

Table 1. Multivariate elliptical distributions.

Distribution	Notation	Generating function
Normal	$N_n(\boldsymbol{\mu}, \Sigma)$	$g(u) = c_1 e^{-u/2}, \quad u \geq 0$
Student-t	$t_n(\boldsymbol{\mu}, \Sigma, \nu)$	$g(u) = c_2(1 + u/\nu)^{-(\nu+n)/2}, \quad u \geq 0$
Contaminated Normal	$CN_n(\boldsymbol{\mu}, \Sigma, \delta, \tau)$	$g(u) = c_1\{(1-\delta)e^{-u/2} + \delta\tau^{-n/2}e^{-u/(2\tau)}\}, \quad u \geq 0$
Cauchy	$C_n(\mu, \Sigma)$	$g(u) = c_3(1+u)^{-(n+1)/2}, \quad u \geq 0$
Logistic	$L_n(\mu, \Sigma)$	$g(u) = c_4 e^{-u}/(1+e^{-u})^2, \quad u \geq 0$
Power Exponential	$PE_n(\boldsymbol{\mu}, \Sigma, \alpha)$	$g(u) = c_5 e^{-u^\alpha/2}, \quad u \geq 0$

Some properties of the elliptical distributions (see, for instance, Fang et al.,1990) are given in the following:

(a) The characteristic function takes the form

$$E\{\exp(it^T\mathbf{Y})\} = \exp(it^T\boldsymbol{\mu})\psi(\mathrm{t}^T\Sigma\mathrm{t}), \mathrm{t} \in I\!R^n,$$

for some function $\psi(\cdot)$.

(b) There exist $E(\mathbf{Y}) = \boldsymbol{\mu}$ and $\mathrm{Var}(\mathbf{Y}) = -2\psi^{(1)}(0)\Sigma$, where $\psi^{(1)}(0) = \partial\psi(\mathrm{t}^T\Sigma\mathrm{t})/\partial\mathrm{t}\mid_{\mathrm{t}=0}$.

(c) If $\mathbf{T}(\mathbf{Y}) = \mathrm{a} + \mathbf{A}\mathbf{Y}$, where a is a $(p \times 1)$ vector of constants and $\mathbf{A}$ is a $(p \times n)$ matrix of constants, then

$$\mathbf{T}(\mathbf{Y}) = El_p(\mathrm{a} + \mathbf{A}\boldsymbol{\mu}, \mathbf{A}\Sigma\mathbf{A}^T).$$

(d) The distribution of the statistic $\mathbf{T}(\mathbf{Y})$ is invariant in the class of distributions $EL_n(\mathbf{0}, \mathbf{I}_n)$ with $P(\mathbf{Y} = \mathbf{0}) = 0$, if

$$\mathbf{T}(k\mathbf{Y}) \stackrel{d}{=} \mathbf{T}(\mathbf{Y}), \forall k > 0,$$

where the operator $\stackrel{d}{=}$ indicates the same distribution. In this case we have that $\mathbf{T}(\mathbf{Y}) \stackrel{d}{=} \mathbf{T}(\mathbf{Z})$, with $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$.

Consider now the linear regression model

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon,$$

where $\mathbf{Y}$ is a $n \times 1$ vector of responses, $\mathbf{X}$ is a known $n \times p$ matrix of rank p, β is a p-dimensional vector of parameters and ϵ is a p-dimensional error vector with distribution $El_n(\mathbf{0}, \phi\mathbf{I})$, where ϕ is the scale parameter. Thus, it follows that $\mathbf{Y} \sim El_n(\mathbf{X}\beta, \phi\mathbf{I})$. This is typically called the elliptical linear regression model. If $g(\cdot)$ is a continuous and decreasing function then the maximum likelihood estimators of β and ϕ are given by (see Fang and Anderson, 1990)

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \quad \text{and} \quad \hat{\phi} = Q(\hat{\beta})/u_0,$$

where $Q(\beta) = (\mathbf{Y} - \mathbf{X}\beta)^T(\mathbf{Y} - \mathbf{X}\beta)$ and u_0 maximizes the function

$$h_n(u) = u^{n/2}g(u), \quad u \geq 0. \tag{2}$$

Typically, if $g(\cdot)$ in (2) is continuous and decreasing then its maximum u_0 exists, and is finite and positive. Moreover, if $g(\cdot)$ is continuous and differentiable then u_0 is the solution to (Fang and Anderson, 1990)

$$g'(u) + \frac{n}{2u}g(u) = 0,$$

or, equivalently, the solution to the equation

$$\frac{n}{2u} + W_g(u) = 0, \tag{3}$$

where $W_g(u) = d\log g(u)/du = g'(u)/g(u)$. It is easy to see that for the normal and Student-t distributions $u_0 = n$, while for the power exponential $u_0 = (n/\alpha)^{1/\alpha}$. However, for the contaminated normal and logistic distributions, u_0 has to be obtained numerically. In the case of the logistic distribution, for example, equation (3) becomes

$$\frac{n}{2u} = \tanh\left(\frac{u}{2}\right),$$

where $\tanh(\cdot)$ denotes the hyperbolic tangent.

Using properties of the elliptical distributions we may show that

$$\hat{\beta} \sim El_n(\mathbf{X}\beta, \phi(\mathbf{X}^T\mathbf{X})^{-1})$$

and

$$F = \frac{(\hat{\beta}-\beta)^T(\mathbf{X}^T\mathbf{X})^{-1}(\hat{\beta}-\beta)}{ps^2} \sim F_{p,(n-p)},$$

where $F_{p,(n-p)}$ denotes the F distribution with p and $(n-p)$ degrees of freedom and $s^2 = Q(\hat{\beta})/(n-p) = u_0\hat{\phi}/(n-p)$. Then, an $100(1-\gamma)\%$ confidence region for $\boldsymbol{\beta}$, where $0 < \gamma < 1$, is given by

$$\mathbf{R} = \{\beta \in I\!R^p : (\hat{\beta}-\beta)^T(\mathbf{X}^T\mathbf{X})^{-1}(\hat{\beta}-\beta) \leq ps^2F_{p,(n-p)}(1-\gamma)\}, \tag{4}$$

where $F_{p,(n-p)}(1-\gamma)$ denotes the $100(1-\gamma)th$ quantile of the $F_{p,(n-p)}$ distribution.

In addition, the likelihood ratio statistic for testing $H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{C}$ against $H_1 : \mathbf{A}\boldsymbol{\beta} \neq \mathbf{C}$, where $\mathbf{A}$ is an $(q \times p)$ matrix of rank q and $\mathbf{C}$ is an $(q \times 1)$ vector of constants, is given by

$$\lambda = \frac{(\mathbf{A}\hat{\beta}-\mathbf{C})^T\{\mathbf{A}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\}^{-1}(\mathbf{A}\hat{\beta}-\mathbf{C})}{\mathbf{Y}^T(\mathbf{I}_n-\mathbf{P})\mathbf{Y}}, \tag{5}$$

where $\mathbf{P} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, $(n-p)\lambda/q \overset{H_0}{\sim} F_{q,(n-p)}$ and λ is independent of $g(\cdot)$.

3 Effects of individual observations

3.1 Residuals

In this section we discuss some properties of two standardized forms for the ordinary residual in elliptical linear regression model. The vector of the ordinary residual is defined by $\mathbf{e} =$

$\mathbf{Y} - \hat{\mathbf{Y}} = (\mathbf{I} - \mathbf{P})\mathbf{Y}$. Then, we have that

$$\mathbf{e} \sim El_n(\mathbf{0}, \phi(\mathbf{I} - \mathbf{P})),$$

and in particular $e_i \sim El(0, \phi(1 - p_{ii}))$, $i = 1, \ldots, n$. We may have two standardized versions for the residual e_i, namely

$$r_i = \frac{e_i}{s\sqrt{1 - p_{ii}}}$$

and

$$t_i = \frac{e_i}{s_{(i)}\sqrt{1 - p_{ii}}},$$

$i = 1, \ldots, n$, where $s = \sqrt{u_0\hat{\phi}/(n-p)}$ and $s_{(i)} = \sqrt{u_0^*\hat{\phi}_{(i)}/(n-p-1)}$, with u_0^* denoting the maximum of the function $h_{n-1}(u)$ and $\hat{\phi}_{(i)} = Q_{(i)}(\hat{\beta}_{(i)})/u_0^*$ denotes the maximum likelihood estimator of ϕ by dropping the ith observation. Furthermore, we may note that $r_i(k\epsilon) = r_i(\epsilon)$ and $t_i(k\epsilon) = t_i(\epsilon)$, $\forall k > 0$. Hence, if $\epsilon \sim El_n(\mathbf{0}, \phi\mathbf{I})$ it follows from the property (d) of Section 2 that

$$t_i \sim t_{(n-p-1)}$$

and

$$b_i = \frac{r_i^2}{(n-p)} \sim Beta(1/2, (n-p-1)/2),$$

$i = 1, \ldots, n$, where $t_{(n-p-1)}$ denotes the Student-t distribution with $n - p - 1$ degrees of freedom and t_i and b_i are independent of $g(\cdot)$. These results show that the distributions of t_i and b_i are invariant in the class of elliptical linear models.

Consider now the mean-shift perturbation elliptical linear model

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{d}_i\tau + \epsilon,$$

where $\mathbf{d}_i$ denotes an $(n \times 1)$ vector of zeros with one at the ith position. Then, from (5), the F statistic to assess if the ith observation is an outlier, that corresponds in testing $H_0 : \tau = 0$ against $H_1 : \tau \neq 0$, is given by

$$F_{(i)} = t_i^2 \sim F_{1,(n-p-1)} \text{ under } H_0.$$

The demonstration of this result is similar to the one of the normal linear case (see, for instance, Chatterjee and Hadi, 1988).

3.2 Cook distance

Based on the confidence region for β in the normal linear case, Cook (1977) proposed, in order to assess the influence of the ith observation on $\hat{\beta}$, a distance between $\hat{\beta}$ and $\hat{\beta}_{(i)}$, where $\hat{\beta}_{(i)}$ denotes the maximum likelihood estimator of β by dropping the ith observation. It may be shown from (4) that the Cook distance in the elliptical linear model is given by

$$\begin{aligned} D_i &= (\hat{\beta} - \hat{\beta}_{(i)})^T (\mathbf{X}^T\mathbf{X})(\hat{\beta} - \hat{\beta}_{(i)})/ps^2 \\ &= \left(\frac{p_{ii}}{1 - p_{ii}}\right)\frac{r_i^2}{p}, \end{aligned}$$

$i = 1, \ldots, n$. Note that D_i is invariant in the class of elliptical linear models.

3.3 Scale ratio

Similarly to the normal linear case (see, for instance, Belsley et al., 1980) we can assess the influence of the ith observation on the scale matrix $\mathbf{D}(\hat{\beta}) = \hat{\phi}(\mathbf{X}^T\mathbf{X})^{-1}$, by using the influence measure

$$SCR_i = \frac{\det\{\hat{\phi}_{(i)}(\mathbf{X}_{(i)}^T\mathbf{X}_{(i)})^{-1}\}}{\det\{\hat{\phi}(\mathbf{X}^T\mathbf{X})^{-1}\}}$$

$i = 1, \ldots, n$, where $\mathbf{X}_{(i)}$ denotes the matrix $\mathbf{X}$ without the ith row. Since $\det(\mathbf{X}_{(i)}^T\mathbf{X}_{(i)}) = (1 - p_{ii})\det(\mathbf{X}^T\mathbf{X})$ and

$$\frac{\hat{\phi}_{(i)}}{\hat{\phi}} = \left(\frac{u_0}{u_0^*}\right)\left\{1 - \frac{r_i^2}{(n-p)}\right\} = \left(\frac{u_0}{u_0^*}\right)(1 - b_i),$$

we have that

$$SCR_i = \left(\frac{u_0}{u_0^*}\right)^p \frac{(1 - b_i)^p}{(1 - p_{ii})},$$

$i = 1, \ldots, n$. This measures is not invariant in the class of elliptical linear models.

3.4 Andrews-Pregibon measure

Andrews and Pregibon (1978) proposed an influence measure to detect remote observations in the subspace explained by the response and explanatory variable vectors. This measure becomes, in the linear elliptical model, given by

$$\begin{aligned} AP_{(i)} &= \frac{Q_{(i)}(\hat{\beta}_{(i)})\det\{\mathbf{X}_{(i)}^T\mathbf{X}_{(i)}\}}{Q(\hat{\beta})\det(\mathbf{X}^T\mathbf{X})} \\ &= (1-p_{ii})(1-b_i), \end{aligned}$$

$i = 1, \ldots, n$. Note that $AP_{(i)}$ is also invariant in the class of elliptical linear models.

4 Likelihood displacement

Let $L(\boldsymbol{\theta})$ denote the log-likelihood function for the elliptical linear model, where $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \phi)^T$. The likelihood displacement (see, for instance, Cook and Weisberg, 1982 and Cook et al., 1988) is defined by

$$LD_i(\boldsymbol{\theta}) = 2\{L(\hat{\boldsymbol{\theta}}) - L(\hat{\boldsymbol{\theta}}_{(i)})\},$$

where $\hat{\boldsymbol{\theta}}_{(i)}$ denotes the maximum likelihood estimator of $\boldsymbol{\theta}$ by dropping the ith observation. In order to assess the influence of the ith observation on $\hat{\boldsymbol{\theta}}$ we may compare $LD_i(\boldsymbol{\theta})$ with some quantile of the chi-square distribution with $p+1$ degrees of freedom.

For the elliptical linear model we find

$$L(\boldsymbol{\theta}) = -\frac{n}{2}\log(\phi) + \log\left[g\left\{\frac{Q(\boldsymbol{\beta})}{\phi}\right\}\right], \tag{6}$$

that evaluated at $\hat{\boldsymbol{\theta}}$ leads to

$$L(\hat{\boldsymbol{\theta}}) = -\frac{n}{2}\log(\hat{\phi}) + \log\{g(u_0)\}.$$

Evaluating (6) at $\hat{\boldsymbol{\theta}}_{(i)} = (\hat{\boldsymbol{\beta}}_{(i)}^T, \hat{\phi}_{(i)})^T$ we obtain

$$L(\hat{\boldsymbol{\theta}}_{(i)}) = -\frac{n}{2}\log(\hat{\phi}_{(i)}) + \log\left[g\left\{\frac{Q(\hat{\boldsymbol{\beta}}_{(i)})}{\hat{\phi}_{(i)}}\right\}\right].$$

Since

$$\frac{Q(\hat{\beta}_{(i)})}{\hat{\phi}_{(i)}} = u_0^* \left\{1 + \frac{b_i}{(1-p_{ii})(1-b_i)}\right\},$$

we get

$$L(\hat{\theta}_{(i)}) = -\frac{n}{2}\log(\hat{\phi}_{(i)}) + \log\left[g\left\{u_0^*\left(1 + \frac{b_i}{(1-p_{ii})(1-b_i)}\right)\right\}\right],$$

and consequently the distance (6) may be expressed in the form

$$LD_i(\theta) = n\log\left\{\left(\frac{u_0}{u_0^*}\right)(1-b_i)\right\} + 2\log\left[g(u_0)/g\left\{u_0^*\left(1 + \frac{b_i}{(1-p_{ii})(1-b_i)}\right)\right\}\right],$$

$i = 1, \ldots, n$. Note that $LD_i(\theta)$ is not invariant in the class of elliptical linear models.

In particular for the Student-t distribution we have $g(u) = c_2(1 + u/\nu)^{-(n+\nu)/2}$, $u_0 = n$ and $u_0^* = n - 1$. Then,

$$LD_i(\theta) = n\log\left\{\left(\frac{n}{n-1}\right)(1-b_i)\right\} + (n+\nu)\log\left[\left\{\nu + \frac{(n-1)(p_{ii}b_i + 1 - p_{ii})}{(1-p_{ii})(1-b_i)}\right\}/(n+\nu)\right]. \tag{7}$$

When $\nu \to \infty$ expression (7) reduces to

$$LD_i(\theta) = n\log\left\{\left(\frac{n}{n-1}\right)(1-b_i)\right\} + \left(\frac{n-1}{1-p_{ii}}\right)\left(\frac{b_i}{1-b_i}\right) - 1,$$

which corresponds to the normal linear case, as expected (see, for instance, Cook et al., 1988).

It may be also shown that $LD_i(\theta)$ takes, for the power exponential model, the form

$$LD_i(\theta) = n\log\left\{\left(\frac{n}{n-1}\right)^{1/\alpha}(1-b_i)\right\} + \frac{1}{\alpha}\left\{(n-1)\left(1 + \frac{b_i}{(1-p_{ii})(1-b_i)}\right)^{\alpha} - n\right\}.$$

4.1 Parameter subsets

Suppose now we have interest on the parameter vector β with ϕ being considered as a nuisance parameter. The likelihood displacement is defined in this case (see, for instance, Cook et al., 1988) as

$$LD_i(\beta \mid \phi) = 2\{L(\hat{\theta}) - \max_{\phi} L(\hat{\beta}_{(i)}, \phi)\}. \tag{8}$$

We may show that the value of ϕ which maximizes $L(\hat{\beta}_{(i)}, \phi)$ is $\tilde{\phi} = Q(\hat{\beta}_{(i)})/u_0$. Then, expression (8) reduces to

$$LD_i(\beta \mid \phi) = n\log\left(1 + \frac{pD_i}{(n-p)}\right), \tag{9}$$

$i = 1, \ldots, n$. Expression (9) agrees with expression (4) given in Cook et al. (1988). It is interesting to note that (9) is invariant in the class of elliptical linear models.

Similarly, we may show that

$$LD_i(\phi \mid \beta) = n\log(\hat{\phi}_{(i)}/\hat{\phi}) + 2\log[g(u_0)/g\{u_0\hat{\phi}/\hat{\phi}_{(i)}\}], \tag{10}$$

$i = 1, \ldots, n$, but (10) depends on the elliptical distribution. In the normal case $g(u) = c_1\exp(-u/2)$, $u_0 = n$ and $u_0^* = n - 1$ so that (10) reduces to

$$LD_i(\phi \mid \beta) = n\log\left\{\left(\frac{n}{n-1}\right)(1 - b_i)\right\} + \left(\frac{nb_i - 1}{1 - b_i}\right), \tag{11}$$

$i = 1, \ldots, n$. As expected, expression (11) agrees with expression (11) given in Cook et al. (1988) for the normal linear model.

5 Local influence

Let $L(\boldsymbol{\theta})$ denote the log likelihood function from the postulated model (here $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \phi)^T$), and let $\boldsymbol{\omega}$ be a $q \times 1$ vector of perturbations restricted to some open subset $\Omega \in \mathbb{R}^q$. The perturbations are made on the likelihood function, such that it takes the form $L(\boldsymbol{\theta}|\boldsymbol{\omega})$. Denoting the vector of no perturbation by $\boldsymbol{\omega}_0$, we assume $L(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = L(\boldsymbol{\theta})$. To assess the influence of the perturbations on the maximum likelihood estimate $\hat{\boldsymbol{\theta}}$, one may consider the likelihood displacement

$$LD(\boldsymbol{\omega}) = 2\{L(\hat{\boldsymbol{\theta}}) - L(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\},$$

where $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denotes the maximum likelihood estimate under the model $L(\boldsymbol{\theta}|\boldsymbol{\omega})$.

In some situations, though, it may be of interest to assess the influence on a subset $\boldsymbol{\theta}_1$ of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$. For example, one may have interest on $\boldsymbol{\theta}_1^T = (\beta_1, \ldots, \beta_p)^T$ or $\boldsymbol{\theta}_1 = \phi$. In these cases, the likelihood displacement is defined as

$$LD_1(\omega) = 2[L(\hat{\boldsymbol{\theta}}) - L\{\hat{\boldsymbol{\theta}}_{1\omega}, \hat{\boldsymbol{\theta}}_2(\hat{\boldsymbol{\theta}}_{1\omega})\}],$$

where $\hat{\boldsymbol{\theta}}_{1\omega}$ is obtained from $\hat{\boldsymbol{\theta}}_\omega = (\hat{\boldsymbol{\theta}}_{1\omega}^T, \hat{\boldsymbol{\theta}}_{2\omega}^T)^T$ and $\hat{\boldsymbol{\theta}}_2(\hat{\boldsymbol{\theta}}_{1\omega})$ is the maximum likelihood estimate of $\boldsymbol{\theta}_2$ for $\hat{\boldsymbol{\theta}}_{1\omega}$ fixed in the perturbed model.

The idea of local influence (Cook, 1986) is concerned in characterizing the behavior of $LD(\omega)$ around ω_0. The procedure consists in selecting a unit direction $\boldsymbol{\ell}$, $\| \boldsymbol{\ell} \| = 1$, and then to consider the plot of $LD(\omega_0 + a\boldsymbol{\ell})$ against a, where $a \in \mathbb{R}$. This plot is called *lifted line*. Note that, since $LD(\omega_0) = 0$, $LD(\omega_0 + a\boldsymbol{\ell})$ has a local minimum at $a = 0$. Each lifted line can be characterized by considering the normal curvature $C_{\boldsymbol{\ell}}(\boldsymbol{\theta})$ around $a = 0$. This curvature is interpreted as the inverse radius of the best fitting circle at $a = 0$. The suggestion is to consider the direction $\boldsymbol{\ell}_{max}$ corresponding to the largest curvature $C_{\boldsymbol{\ell}_{max}}(\boldsymbol{\theta})$. The index plot of $\boldsymbol{\ell}_{max}$ may reveal those observations that under small perturbations exercise notable influence on $LD(\omega)$.

Cook(1986) showed that the normal curvature at the direction $\boldsymbol{\ell}$ takes the form

$$C_{\boldsymbol{\ell}}(\boldsymbol{\theta}) = 2|\boldsymbol{\ell}^T \boldsymbol{\Delta}^T (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta} \boldsymbol{\ell}|, \tag{12}$$

where $-\ddot{\mathbf{L}}$ is the observed Fisher information matrix for the postulated model ($\omega = \omega_0$) and $\boldsymbol{\Delta}$ is the $(p+1) \times q$ matrix with elements

$$\Delta_{ij} = \frac{\partial^2 L(\boldsymbol{\theta}|\omega)}{\partial \theta_i \partial \omega_j},$$

evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and $\omega = \omega_0$, $i = 1, \ldots, p+1$ and $j = 1, \ldots, q$.

Therefore, the maximization of (12) is equivalent to finding the largest eigenvalue $C_{\boldsymbol{\ell}_{max}}$ of the matrix $\mathbf{B} = \boldsymbol{\Delta}^T (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}$, and $\boldsymbol{\ell}_{max}$ is the corresponding eigenvector.

For the subset $\boldsymbol{\theta}_1$, the curvature at the direction $\boldsymbol{\ell}$ is given by

$$C_{\boldsymbol{\ell}}(\boldsymbol{\theta}_1) = 2|\boldsymbol{\ell}^T \boldsymbol{\Delta}^T (\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22}) \boldsymbol{\Delta} \boldsymbol{\ell}|,$$

where $\mathbf{B}_{22}$ is defined as

$$\mathbf{B}_{22} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix},$$

and $\ddot{\mathbf{L}}_{22}$ is obtained from the partition of $\ddot{\mathbf{L}}$ according to the partition of $\boldsymbol{\theta}$. The eigenvector $\boldsymbol{\ell}_{max}$ corresponds to the largest eigenvalue of the matrix $\mathbf{B} = \boldsymbol{\Delta}^T(\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22})\boldsymbol{\Delta}$.

Recently, Fung and Kwan (1997) presented an interesting discussion on the application of the local influence method for other influence measures than the likelihood displacement. They show that an influence measure, namely $\hat{T}_{\boldsymbol{\omega}}$, is scale invariant if $\dot{\Gamma} = \partial \hat{T}_{\omega} / \partial \omega |_{\omega = \omega_0} = \mathbf{0}$. When this derivative is non-zero the ordering among the components of $\boldsymbol{\ell}_{max}$ is not necessarily preserved under changes in the scale. In particular, for the likelihood displacement, we have $\dot{\Gamma} = \partial L(\hat{\beta}_{\boldsymbol{\omega}}) / \partial \omega |_{\omega = \omega_0} = \mathbf{0}$. This property also follows, for instance, for the influence measures proposed by Thomas and Cook (1990) and Paula (1993). But, it does not hold for other influence measures as pointed out by Fung and Kwan (1997).

5.1 Curvature derivation

We assume the perturbation scheme $\boldsymbol{\epsilon} \sim El_n(\mathbf{0}, \phi \mathbf{D}^{-1}(\omega))$, where $\mathbf{D}(\omega) = \text{diag}\{\omega_1, \ldots, \omega_n\}$ with ω_i denoting the weight corresponding to the ith case and $\boldsymbol{\omega}_0 = \mathbf{1}$. The normal curvature in the direction $\boldsymbol{\ell}$ for $\boldsymbol{\theta}$ (see Galea et al., 1997) is given by

$$C_{\boldsymbol{\ell}}(\boldsymbol{\theta}) = 2|\boldsymbol{\ell}^T[\mathbf{B}_1 + \mathbf{B}_2]\boldsymbol{\ell}|,$$

where

$$\mathbf{B}_1 = (2/\hat{\phi}) W_g(u_0) \mathbf{D}(\mathbf{e}) \mathbf{P} \mathbf{D}(\mathbf{e})$$

and

$$\mathbf{B}_2 = \frac{1}{\hat{\phi}^2} \frac{[W_g(u_0) + u_0 W_g'(u_0)]^2}{[\frac{n}{2} + u_0\{2W_g(u_0) + u_0 W_g'(u_0)\}]} \mathbf{D}(\mathbf{e}) \mathbf{e}\mathbf{e}^T \mathbf{D}(\mathbf{e}).$$

In particular, if we are interested in the vector $\boldsymbol{\beta}$, the normal curvature in the direction $\boldsymbol{\ell}$ yields

$$C_{\boldsymbol{\ell}}(\boldsymbol{\beta}) = \frac{4}{\hat{\phi}} |W_g(u_0)| |\boldsymbol{\ell}^T \mathbf{D}(\mathbf{e}) \mathbf{P} \mathbf{D}(\mathbf{e}) \boldsymbol{\ell}|,$$

where $\mathbf{D}(\mathbf{e}) = \text{diag}\{e_1, \ldots, e_n\}$. Then, the index plot for $\boldsymbol{\ell}_{max}$ obtained from the matrix $\mathbf{D}(\mathbf{e})\mathbf{P}\mathbf{D}(\mathbf{e})$ may show how to perturb $\mathbf{D}(\omega)$ to obtain larger changes in the regression coefficients.

Similarly, the normal curvature for ϕ in the direction $\boldsymbol{\ell}$ takes the form

$$C_{\boldsymbol{\ell}}(\phi) = \frac{2}{\hat{\phi}^2}|C_\omega||\boldsymbol{\ell}^T\mathbf{D}(\mathbf{e})\mathbf{e}\mathbf{e}^T\mathbf{D}(\mathbf{e})\boldsymbol{\ell}|,$$

where

$$C_\omega = [W_g(u_0) + u_0W_g'(u_0)]^2/[\frac{n}{2} + u_0\{2W_g(u_0) + u_0W_g'(u_0)\}].$$

In this case, for the largest curvature,

$$\boldsymbol{\ell}_{max} \propto \mathbf{D}(\mathbf{e})\mathbf{e},$$

which means that observations with large values for e_i^2 are most influential on $\hat{\phi}$.

Therefore, since $\mathbf{e}$ is invariant in the class of elliptical linear models the vector $\boldsymbol{\ell}_{max}$ is invariant when we are interested in the vector $\boldsymbol{\beta}$ or in the scale parameter ϕ. However, if we are interested in both, $\boldsymbol{\beta}$ and ϕ, the vector $\boldsymbol{\ell}_{max}$ depends on the elliptical distribution.

6 Application

As illustration consider the data set reported by Ruppert and Carroll (1980) on the salinity of water during the spring in Pamlico Sound, North Carolina. The response Y was biweekly salinity, and the explanatory variables were salinity lagged 2 weeks, x_1, a dummy variables x_2 for the time period and river discharge, x_3. The value of x_{1i} may differ from y_{i-1}, since the data are not a contiguous sequence. Several authors have been analyzed this data set, particularly under the diagnostic viewpoint. The linear model

$$Y_i = \beta_0 + \beta_1x_{1i} + \beta_2x_{2i} + \beta_3x_{3i} + \epsilon_i,$$

where ϵ_i follows an appropriate symmetric distribution, has been adopted.

Atkinson (1985), for instance, assumed a normal distribution for ϵ_i and using deletion diagnostic methods found cases 16 and 5 as the most influential on the parameter estimates. Davison and Tsai (1992) assumed a Student-t distribution with 3 degrees of freedom for ϵ_i and in their deletion diagnostic analysis cases 16, 5 and 3 appeared as the most influential. Galea et al. (1997) assumed that $\epsilon \sim El_n(0, \phi \mathrm{I})$ and applied the local influence method for assessing the influence of the observations on $\hat{\theta}$, $\hat{\beta}$ and $\hat{\phi}$ under some multivariate elliptical distributions. They found case 16 as the most influential on $\hat{\beta}$ and cases 9, 15, 16 and 17 most influential on $\hat{\phi}$.

Our analysis will be restricted on the particular distributions, Student-t and power exponential. Figures 1 and 2 present the index plot of $LD_i(\theta)$ under the power exponential and Student-t distributions for $\alpha = 0.1, 0.5$ and 1.2 and $\nu = 3, 30$ and 100 degrees of freedom, respectively. We can notice from Figure 1 that case 16 appears with more accentuated influence for $\alpha = 1.2$ rather than for $\alpha = 0.5$ and 0.1. It may due to the fact that the power exponential distribution has tails less weighted than the normal distribution as α becomes greater than one. Similar tendency is observed in Figure 2. The influence of observation 16 becomes less accentuated for small degrees of freedom. Figures 3 and 4 present the index plot of $LD_i(\phi \mid \beta)$. In these figures we can notice similar tendencies to the ones observed in Figures 1 and 2, respectively. The index plot of $LD_i(\beta \mid \phi)$, that is invariant under the elliptical distributions and is omitted here, points out observation 16 as the most influential.

Figures 5a and 5b present the index plot of $\mid \ell_{max} \mid$ for θ under the power exponential distribution with $\alpha = 0.1$ and 1.2. Note that case 16 appears most influential in both graphics. In Figures 5c and 5d one has the index plot of $\mid \ell_{max} \mid$ for θ under the Student-t distribution with $\nu = 3$ degrees of freedom and under the normal distribution. Similarly to the power exponential distribution case 16 is influential in both situations.

The main conclusion from these two analysis is that the robustness of the power exponential and Student-t distributions seems to be more evident under individual deletion of observations rather than under the local influence on all the observations. Finally, it is interesting to note that although the estimates $\hat{\beta}$ and $\hat{\phi}$ are the same for normal and Student-t models, the influence measures $LD_i(\theta)$ and $LD_i(\phi \mid \beta)$ change. The explanation for this result is that the distribution of $\hat{\beta}$ and $\hat{\phi}$ depend on the elliptical distribution and consequently the influence measures based on the likelihood displacement will also depend.

Acknowledgment: This work was supported by CNPq and FAPESP-Brazil and FONDECYT-1980451-Chile.

References

Andrews, D. F. and Pregibon, D. (1978). Finding the outliers that matter. *Journal of the Royal Statistical Society* B 40, 80-93.

Arellano, R. (1994). *Elliptical Distributions: Properties, Inference and Applications in Regression Models.* Unpublished Ph.D. Thesis, Department of Statistics, University of Sao Paulo, Brazil.

Atkinson, A. C. (1985). *Plots, Transformations and Regression.* Clarendon Press: Oxford.

Beckman, R. J., Nachtsheim, C. J. and Cook, R. D. (1987). Diagnostics for mixed-model analysis of variance. *Technometrics* 29, 413-426.

Belsley, D., Kuh, E. and Welsch, R. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: Wiley.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression.* New York: Wiley.

Cook, R. D. (1977). Detection of influential observations in linear regressions. *Technometrics* 19, 15-18.

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

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

Cook, R. D., Peña, D. and Weisberg S. (1988). The likelihood displacement: A unifying principle for influence measures. *Communications in Statistics, Theory and Methods* 17, 623-640.

Davison, A.C. and Tsai, C-L. (1992). Regression model diagnostics. *International Statistical Review* 60, 337-353.

Fang, K. T. and Anderson, T. W. (1990). *Statistical Inference in Elliptical Contoured and Related Distributions.* Allerton Press: New York.

Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis.* New York: Springer-Verlag.

Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Simmetric Multivariate and Related Distributions.* London: Chapman and Hall.

Fung, W. K. and Kwan, C. W. (1997). A note on local influence based on normal curvature.*Journal of the Royal Statistical Society* B 59, 839-843.

Galea, M.; Paula, G. A. and Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician* 46, 71-79.

Gupta, A. K. and Varga, T. (1993). *Elliptically Contoured Models in Statistics.* Kluwer Academic Publishers.

Kim, M. G. (1995). Local influence in multivariate regression. *Communications in Statistics, Theory and Methods* 24, 1271-1278.

Lawrance, A. J. (1988). Regression transformation diagnostics using local influence. *Journal of the American Statistical Association* 84, 125-141.

Leiva, V. (1998). Inference on the variation coefficient in elliptical distributions. Unpublished Ph.D. Thesis, University of Granada, Spain.

Paula, G. A. (1993). Assessing local influence in restricted regression models. *Computational Statistics and Data Analysis* 16, 63-79.

Ruppert, D. and Carroll, R. J. (1980). Trimmed least squares estimation in the linear model. *Journal of the American Statistical Association* 75, 828-838.

Thomas, W. and Cook, R. D. (1990). Assessing influence on predictions from generalized linear models. *Technometrics* 32, 59-65.

Tsai, C. H. and Wu, X. (1992). Assessing local,influence in linear regression models with first-order autoregressive or heteroscedastic error structure. *Statistics and Probability Letters* 14, 247-252.

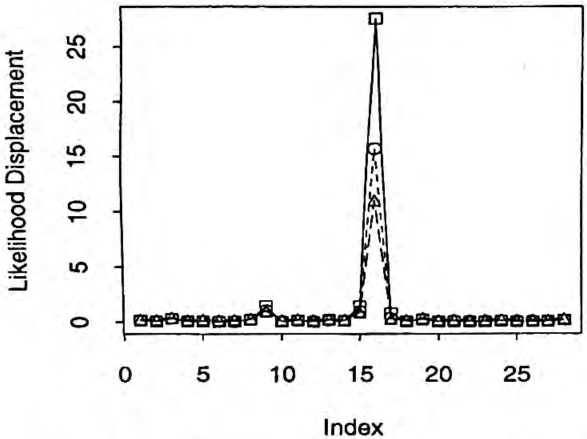

Figure 1: Index pot of $LD_i(\theta)$ for the power exponential distribution with $\alpha = 0.1$ ($-\triangle-$), $\alpha = 0.5$ ($-\circ-$) and $\alpha = 1.2$ ($-\square-$).

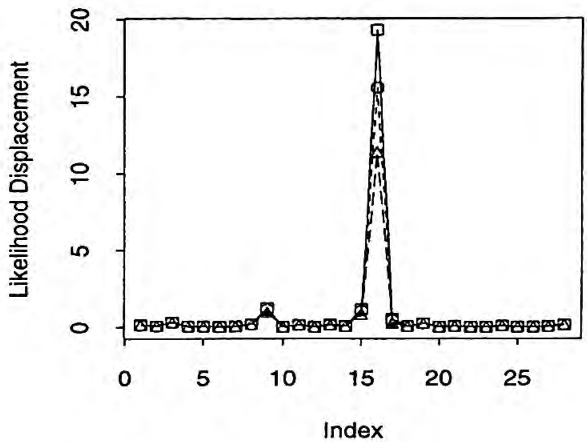

Figure 2: Index pot of $LD_i(\boldsymbol{\theta})$ for the Student-t distribution with $\nu = 3$ ($-\triangle-$), $\nu = 30$ ($-\circ-$) and $\nu = 100$ ($-\square-$) degrees of freedom.

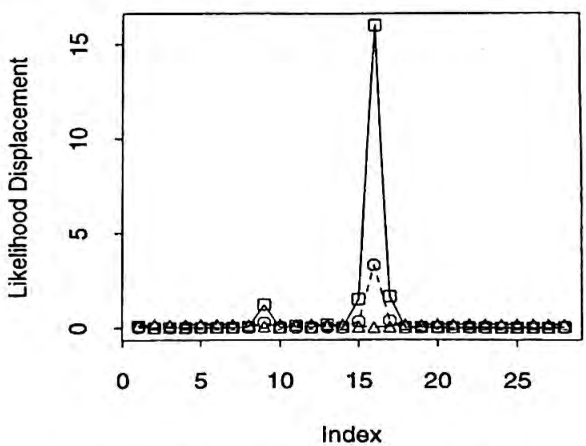

Figure 3: Index pot of $LD_i(\phi \mid \boldsymbol{\theta})$ for the power exponential distribution with $\alpha = 0.1$ ($-\triangle-$), $\alpha = 1.0$ ($-\circ-$) and $\alpha = 3.0$ ($-\square-$).

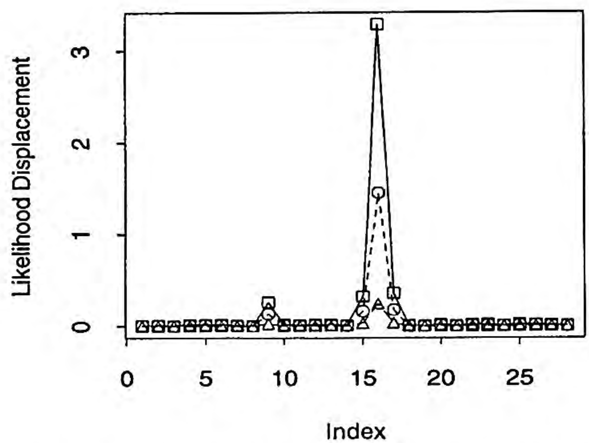

Figure 4: Index pot of $LD_i(\phi \mid \theta)$ for the Student-t distribution with $\nu = 3$ $(-\Delta-)$, $\nu = 30$ $(-\circ-)$ and $\nu = 100$ $(-\square-)$ degrees of freedom.

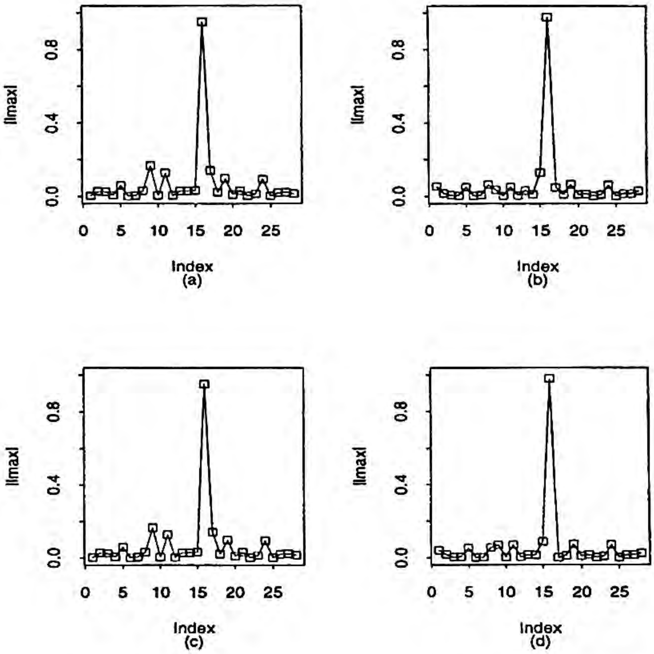

Figure 5: Index pot of $|\ell_{max}|$ for the power exponential distribution with $\alpha = 0.1$ (a) and $\alpha = 1.2$ (b), Student-t distribution with $\nu = 3$ degrees of freedom (c) and normal distribution (d).

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

9901 - **MONTOYA-DELGADO, L.E., IRONY, T.Z., WHITTLE, M.R., PEREIRA, C.A.B.** Unconditional Exact Test of Hardy-Weinberg Law. 1999. 19p.(RT-MAE-9901)

9902 - **MINKOVA, L.** Modelling of financial markets in high level of inflation. 1999. 8p. (RT-MAE-9902)

9903 - **GUERBEROFF, G.R.** Remark on the critical temperature of the hopfield model. 1999. 10p. (RT-MAE-9903)

9904 - **BAKEVA, V., KOLEV, N.** The optimal blocking time of the unreliable Geo/G_D/1 Queuing System. 1999. 12p. (RT-MAE-9904)

9905 - **KOLEV, N., MINKOVA, L.** Quotas on runs of successes and failures in a multi-state markov chain. 1999. 14p. (RT-MAE-9905)

9906 - **KOLEV, N., MINKOVA, L.** Run and frequency quotas in a multi-state Markov chain. 1999. 11p. (RT-MAE-9906)

9907 - **CHIANN, C., MORETTIN, P.A.** Estimation of time varying linear systems. 1999. 39p. (RT-MAE-9907)

9908 - **FONTES, L.R.G., NEVES, E.J., SIDORAVICIUS, V.** Limit velocity for a driven particle in a random medium with mass aggregation. 1999. 14p. (RT-MAE-9908)

9909 - **PEREIRA, C.A.B., STERN, J.M.** Evidence and credibility: Full bayesian significance test for precise hypotheses, 1999. 13p. (RT-MAE-9909)

9910 - **GALEA, M., PAULA, G.A., URIBE-OPAZO, M.** Assessing local influence in univariate elliptical linear regression models, 1999, 16p. (RT-MAE-9910)

9911 - **CARDOSO NETO, J., PAULA, G.A.** Wald one-sided test using generalized estimating equations approach, 1999, 18p. (RT-MAE-9911)

9912 - **CORGINI, M., SANKOVICH, D.P., TANAKA, N.I.** The imperfect boson gas: some mathematical approaches, 1999. 25p. (RT-MAE-9912)

9913 - **BOLFARINE, H., CASTRO, M.** Ancova models with measurement errors, 1999. 13p. (RT-MAE-9913)

9914 - **MACHADO, F.P., MENSHIKOV, M.V., POPOV, S.YU.** Recurrence and transience of multi type branching random walks, 1999. 20p. (RT-MAE-9914)

9915 - FERNÁNDEZ, R., GALVES, A. Identifying features in the presence of competing evidence. The case of first-language acquisition, 1999. 11p. (RT-MAE-9915)

9916 - AUBIN, E.C.Q., CORDEIRO, G.M. Bartlett-corrected tests for normal linear models when the error covariance matrix is nonscalar. 1999. 22p. (RT-MAE-9916)

9917 - CRIBARI-NETO, F., FERRARI, S.L.P., CORDEIRO, G.M. An improved heteroskedasticity-consistent covariance matrix estimator. 1999. 10p. (RT-MAE-9917)

9918 - MINKOVA, L. Further results on the family of inflated-parameter power series distributions. 1999. 13p. (RT-MAE-9918)

9919 - KOLEV, N. Run and frequency quotas under markovian fashion and their application in risk analysis. 1999. 15p. (RT-MAE-9919)

9920 - KOLEV, N. Two characterizations of the geometric distribution related to β-transformation. 1999. 11p. (RT-MAE-9920)

9921 - KOLEV, N. Correlated INAR(1) Process. 1999. 12p. (RT-MAE-9921)

9922 - KOLEV, N. Two extended partially correlated models. 1999. 13p. (RT-MAE-9922)

9923 - BARROSO, L.P., CORDEIRO, G.M., VASCONCELOS, K.L.P. Improved score tests for von mises regression models. 1999. 16p. (RT-MAE-9923)

9924 - VASCONCELLOS, K.L.P., BARROSO, L.P., CORDEIRO, G.M. Improved estimation for robust econometric regression models. 1999. 13p. (RT-MAE-9924)

9925 - GALEA-ROJAS, M., BOLFARINE, H., CASTRO, M. Local influence in comparative calibration models. 1999. 20p. (RT-MAE-9925)

9926 - MINKOVA, L. The pólya-aeppli process and ruin problems. 1999. 13p. (RT-MAE-9926)

9927 - CORDEIRO, M.G., FERRARI, S.L.P. Third-order asymptotic distributions of classic test statistics for one-parameter exponential family models. 1999. 18p.(RT-MAE-9927)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05315-970 - São Paulo, Brasil