One-Sided Tests in Univariate Elliptical Linear Regression Models

Francisco José de A. Cysneiros[1] and Gilberto A. Paula[2]

[1] Departamento de Estatística, Universidade Federal de Pernambuco - Caixa Postal 50749-540, Recife - PE - Brazil, e-mail: cysneiros@de.ufpe.br
[2] Instituto de Matemática e Estatística, USP - Caixa Postal 66281 (Ag. Cidade de São Paulo), 05311-970 São Paulo - SP - Brazil, e-mail: giapaula@ime.usp.br

Abstract: We discuss in this paper the problem of testing equality and inequality constraints in univariate elliptical linear regression models. First, the problem of testing the linear equality hypothesis $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$ against the linear inequality hypothesis $H_1 : \mathbf{C}\boldsymbol{\beta} \geq \mathbf{d}$, with at least one strict inequality in H_1 (case 1) and then, $H_1 : \mathbf{C}\boldsymbol{\beta} \geq \mathbf{d}$ against $H_2 : \boldsymbol{\beta} \in I\!R^p - H_1$ (case 2). This class of models includes all symmetric continuous distributions, such as normal, Student-t, Pearson VII, exponential power and logistic, among others. It is commonly used for the analysis of data containing influential or outlying observations with responses supposedly normal. Iterative processes for evaluating the parameters under equality and inequality constraints are presented. Under regular conditions the expressions of the statistics for three asymptotically equivalent statistical tests as well as their asymptotic null distribution are given. An illustrative example with presence of influential observations on the decisions from the statistical tests of different elliptical models is presented. The robustness aspects of such models are discussed.

Keywords: Hypothesis testing; Symmetric distributions; Multivariate one-sided tests; Restricted estimation; Robustness.

1 Univariate elliptical linear models

Let ϵ_i, $i = 1, \ldots, n$, be independent random variables with density function of the form

$$f_{\epsilon_i}(\epsilon) = \frac{1}{\sqrt{\phi}} g\{(\epsilon/\sqrt{\phi})^2\},\ \epsilon \in I\!R, \tag{1}$$

where $\phi > 0$ is the scale parameter, $g : I\!R \rightarrow [0, \infty]$ is such that $\int_0^\infty g(u^2)du < \infty$. We will denote $\epsilon_i \sim El(0, \phi)$. The function $g(.)$ is called density generator (see, for example, Fang, Kotz and Ng, 1990). Consider the linear regression model

$$y_i = \mu_i + \epsilon_i,\ i = 1, \ldots, n, \tag{2}$$

where $\mu_i = \mathbf{x}_i^T \boldsymbol{\beta}$, $\mathbf{x}_i^T = (x_{i1}, \ldots, x_{in})^T$ contains values of p explanatory variables, $y_1, \ldots, y_n$ are the observed response values, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$

is the parameter vector. The model defined by (1)-(2) is called univariate elliptical linear regression model. A joint iterative process to find the unrestricted estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\phi}$ is given by

$$\boldsymbol{\beta}^{(r+1)} = \{\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{X}\}^{-1}\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{y} \qquad \text{and} \tag{3}$$

$$\phi^{(r+1)} = \frac{1}{n}Q_v(\boldsymbol{\beta}^{(r)}), \text{ for } r = 0, 1, \ldots, \tag{4}$$

where $Q_v(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{D}(\mathbf{v})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$, $\mathbf{D}(\mathbf{v}) = \text{diag}\{v_1, \ldots, v_n\}$, $v_i = -2W_g(u_i)$, $u_i = (y_i - \mathbf{x}_i^T\boldsymbol{\beta})^2/\phi$ and $W_g(u) = g'(u)/g(u)$ with $g(u) = \partial g(u)/\partial u$. We should start the iterative process (3)-(4) with initial values $\boldsymbol{\beta}^{(0)}$ and $\phi^{(0)}$.

2 Restricted estimation

2.1 Equality constraints

Suppose first we are interested in estimating the parameter vector $\boldsymbol{\beta}$ under k linearly independent restrictions $\mathbf{C}_j^T\boldsymbol{\beta} - d_j = 0$, where $\mathbf{C}_j$, $j = 1, \ldots, k$, are $p \times 1$ vectors and d_j, $j = 1, \ldots, k$, are scalars, both known and fixed. The problem here is to maximize the log-likelihood function $\mathrm{L}(\boldsymbol{\theta})$ subject to the linear constraints $\mathbf{C}\boldsymbol{\beta} - \mathbf{d} = \mathbf{0}$, where $\mathbf{C} = (\mathbf{C}_1^T, \ldots, \mathbf{C}_k^T)^T$ and $\mathbf{d} = (d_1, \ldots, d_k)^T$. Similarly to Nyquist (1991), that investigated this kind of problem in generalized linear models, we will apply the methodology of penalty functions by considering a quadratic penalty function. The resulting iterative process is given by

$$\begin{aligned}\boldsymbol{\beta}^{0(r+1)} &= \{\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{X}\}^{-1}\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{y} + \{\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{X}\}^{-1}\mathbf{C}^T \times \\ &\quad \left[\mathbf{C}\{\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{X}\}^{-1}\mathbf{C}^T\right]^{-1} \times \\ &\quad \left[\mathbf{d} - \mathbf{C}\{\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{X}\}^{-1}\mathbf{X}^T\mathbf{D}(\mathbf{v}^{(r)})\mathbf{y}\right],\end{aligned} \tag{5}$$

for $r = 0, 1, \ldots$, where $\phi^{(r)}$ is obtained from (4). The authors have developed a library in S-Plus and R to fit univariate elliptical linear models based in some distributions and the iterative process (3-5) and more, some diagnostic graphics. This library is available in the web page `www.de.ufpe.br /~cysneiros/elliptical/elliptical.html`.

2.2 Inequality constraints

The problem of maximizing log-likelihood functions restricted to linear inequality parameter constraints $\mathbf{C}\boldsymbol{\beta} - \mathbf{d} \geq \mathbf{0}$ have been investigated by various authors (see, for instance, Robertson, Wright and Dykstra, 1988 and Fahrmeir and Klinger, 1994). Our primary interest is to obtain the maximum likelihood estimate of $\boldsymbol{\beta}$, denoted by $\tilde{\boldsymbol{\beta}}$, in model (1) subject to the constraints $\mathbf{C}\boldsymbol{\beta} - \mathbf{d} \geq \mathbf{0}$; that is, we want to solve the problem $\max_{\{C\beta - d \geq 0\}} \mathrm{L}(\boldsymbol{\beta}, \phi)$. We can apply the Kuhn-Tucker conditions to attain the restricted maximum. These conditions are equivalent to finding $\tilde{\boldsymbol{\beta}}$ from a searching procedure which consists in maximizing $\mathrm{L}(\boldsymbol{\beta}, \phi)$ subject to $\mathbf{C}_j^T\boldsymbol{\beta} - d_j = 0$, $j \in I$, for each $I \subseteq \{1, \ldots, k\}$. The inequality-restricted problem reduces to a equality-restricted problem that may be solved by the procedures given in Section 2.1.

3 One-sided tests

3.1 Case 1

We will consider in this section the problem of testing the hypotheses $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$ against $H_1 : \mathbf{C}\boldsymbol{\beta} \geq \mathbf{d}$, with at least one strict inequality in H_1. The usual statistics likelihood ratio, Wald and score take, in this case, the forms

$$
\begin{aligned}
\xi_{LR} &= 2\left[\frac{n}{2}\log\left(\frac{\hat{\phi}_0}{\tilde{\phi}}\right) + \sum_{i=1}^{n}\log\left\{\frac{g\{(y_i - \mathbf{x}_i^T\tilde{\boldsymbol{\beta}})^2/\tilde{\phi}\}}{g\{(y_i - \mathbf{x}_i^T\hat{\boldsymbol{\beta}}^0)^2/\hat{\phi}_0\}}\right\}\right], \\
\xi_W &= \frac{4d_g}{\tilde{\phi}}(\mathbf{C}\tilde{\boldsymbol{\beta}} - \mathbf{d})^T\{\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T\}^{-1}(\mathbf{C}\tilde{\boldsymbol{\beta}} - \mathbf{d}) \quad \text{and} \\
\xi_{SR} &= \frac{\hat{\phi}_0}{4d_g}\{\mathbf{U}_\beta(\hat{\boldsymbol{\beta}}^0, \hat{\phi}_0) - \mathbf{U}_\beta(\tilde{\boldsymbol{\beta}}, \tilde{\phi})\}^T(\mathbf{X}^T\mathbf{X})^{-1}\{\mathbf{U}_\beta(\hat{\boldsymbol{\beta}}^0, \hat{\phi}_0) - \mathbf{U}_\beta(\tilde{\boldsymbol{\beta}}, \tilde{\phi})\},
\end{aligned}
$$

respectively, where $d_g = E\{W_g^2(Z^2)Z^2\}$ with $Z \sim El(0,1)$ and $\mathbf{U}_\beta(\boldsymbol{\beta}, \phi) = \frac{1}{\phi}\mathbf{X}^T\mathbf{D}(\mathbf{v})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$. In addition, suppose the parameter space of $\boldsymbol{\beta}$ is open. Under the regular condition given in Gourieroux and Monford (1995, Section 21.3) it follows that the statistics ξ_{LR}, ξ_W and ξ_{SR} are asymptotically equivalent as a mixture of chi-square distributions, namely

$$
Pr\{\xi_{LR} \geq c\} = \sum_{\ell=0}^{k}\omega(k, \ell; \boldsymbol{\Delta})Pr\{\chi_\ell^2 \geq c\} + o(1), \tag{6}
$$

where $c \geq 0$, $\boldsymbol{\Delta} = \mathbf{C}\mathbf{K}_{\beta\beta}^{-1}\mathbf{C}^T$, $\mathbf{K}_{\beta\beta} = \frac{4d_g}{\phi}(\mathbf{X}^T\mathbf{X})$, χ_0^2 denotes the degenerate distribution at the origin and $\omega(k, \ell; \boldsymbol{\Delta})$'s are known as level probabilities which are expressed as functions of correlation coefficients associated with the matrix $\boldsymbol{\Delta}$. These correlation coefficients are the minimum information necessary to compute the asymptotic null distribution given in (6) because $\omega(k, \ell; \boldsymbol{\Delta})$ depends on $\boldsymbol{\Delta}$ only through its correlation matrix. Examining the expression of $\mathbf{K}_{\beta\beta}$ we can conclude that $\omega(k, \ell; \boldsymbol{\Delta})$ does not depend on $\boldsymbol{\beta}$. Then, the distribution given in (6) is unique and consequently invariant in the elliptical class. This property rarely occurs in other classes of regression models such as generalized linear models (see, for instance, Paula and Sen, 1995).

3.2 Case 2

Now, we will consider the hypotheses $H_1 : \mathbf{C}\boldsymbol{\beta} \geq \mathbf{d}$ against $H_2 : \boldsymbol{\beta} \in I\!R^p - H_1$. In this case, the usual statistics likelihood ratio, Wald and score take the forms

$$\begin{aligned}
\xi^c_{LR} &= 2\left[\frac{n}{2}\log\left(\frac{\tilde{\phi}}{\hat{\phi}}\right) + \sum_{i=1}^{n}\log\left\{\frac{g\{(y_i - \mathbf{x}_i^T\hat{\boldsymbol{\beta}})^2/\hat{\phi}\}}{g\{(y_i - \mathbf{x}_i^T\tilde{\boldsymbol{\beta}})^2/\tilde{\phi}\}}\right\}\right], \\
\xi^c_{W} &= \frac{4d_g}{\hat{\phi}}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{C}\tilde{\boldsymbol{\beta}})^T\{\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T\}^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{C}\tilde{\boldsymbol{\beta}}) \text{ and} \\
\xi^c_{SR} &= \frac{\tilde{\phi}}{4d_g}\mathbf{U}_\beta(\tilde{\boldsymbol{\beta}}, \tilde{\phi})^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{U}_\beta(\tilde{\boldsymbol{\beta}}, \tilde{\phi})^T.
\end{aligned}$$

An important asymptotic result observed in the last section is the lack of functional dependence of $\boldsymbol{\Delta} = \mathbf{C}\mathbf{K}^{-1}_{\beta\beta}\mathbf{C}^T$ on $\boldsymbol{\beta}$. The main consequence of this fact is that the asymptotic null distribution of ξ^c_{LR}, ξ^c_W and ξ^c_{SR} for the purpose of testing H_1 against H_2, is uniquely determined and given by

$$Pr\{\xi^c_{LR} \geq c\} = \sum_{\ell=0}^{k}\omega(k, k-\ell; \boldsymbol{\Delta})Pr\{\chi^2_\ell \geq c\} + o(1). \tag{7}$$

4 Example

We will reanalyze in this section the example discussed by Ramanathan (1993) on a study in which seven variables were observed in 40 metropolitan areas. The main interest is on regressing the number (in thousands) of subscribers with cable TV (Y) against the number (in thousands) of homes in the area (X_1), the per capita income for each television market with cable (X_2), the installation fee (X_3), the monthly service charge (X_4), the number of television signals carried by each cable system (X_5) and the number of television signals received with good quality without cable (X_6). Because Y corresponds to count data we will use a square root transformation in order to stabilize the variance of Y. Then, we will propose the model

$$\sqrt{y_i} = \beta_0 + \sum_{j=1}^{6}\beta_j x_{ji} + \epsilon_i, \qquad i = 1, \ldots, 40,$$

where $\epsilon_i \sim El(0, \phi)$ are mutually independent errors. In addition, it is reasonable to assume some constraints. For example, it is expected that the number of subscribers decreases as the monthly service charge increases, which leads to the restriction $\beta_4 \leq 0$. Following the same idea for the remaining variables one has the constraints $\beta_1 \geq 0, \beta_2 \geq 0, \beta_3 \leq 0, \beta_5 \geq 0$ and $\beta_6 \leq 0$. Applying one-sided t tests we can notice indications that the coefficients β_2, β_3 and β_4 seem to be individually equal to zero, at the significance level of 5%, while some doubt appears for the coefficient β_5 whose p-value is about 3%. The remaining coefficients β_1 and β_6 are highly significant in the direction of the constraints. Thus, in order to assess if the four coefficients β_2, β_3, β_4 and β_5 are jointly equal to zero, we apply the statistical tests defined in Sections 3.1 to assess, the hypotheses $H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ against $H_1 : \beta_2 \geq 0, \beta_3 \leq 0, \beta_4 \leq 0$ and $\beta_5 \geq 0$, with at least one strict inequality in H_1. Our main conclusion of this example based on diagnostic methods is that the transformation $\sqrt{Y}$ seems to stabilize the variance of the responses, but the Student-t with 6 degrees of freedom, exponential power and logistic-II models are less influenced by the outlying observation 14 than the normal model. The one-sided tests based on these three fitted models indicate for the rejection of the null hypothesis at the significance level of 5% while under the normal model the rejection of the null hypothesis becomes evident only after dropping the outlying observation 14. However, the Student-t model seems to be more robust against the influential observation 1 than the other three models. Continua ting the selection procedure the Student-t model appears as the best fitted model.

Acknowledgments: The first author received financial support from CAPES and the second author was supported by FAPESP and CNPq, Brazil.

References

Fahrmeir, L. and Klinger, J. (1994). Estimating and testing generalized linear models under inequality restrictions. *Statistical Papers* **35**, 211-229.

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

Gourieroux, G. and Monford, A. (1995). *Statistics and Econometric*. Vols. 1 and 2. Cambridge: Cambridge University Press.

Nyquist, H. (1991). Restricted estimation of generalized linear models. *Applied Statistics* **40**, 133-141

Paula, G. A. and Sen, P. K. (1995). One-sided tests in generalized linear models with parallel regression lines. *Biometrics* **51**, 1494-1501.

Ramanathan, R. (1993). *Statistical Methods in Econometrics*. New York : John Wiley.

Robertson, T.; Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*. New York : John Wiley.