RT-MAE-9310

Bartlett corrected tests for regression models with Student-t independent errors

by

Silvia L. de Paula Ferrari

and

Reinaldo B. Arellano

Palavras Chaves: Asymptotic expansion; Bartlett correction; chi-squared distribution; EM algorithm; likelihood ratio test; maximum likelihood; regression model; socre test ; Student-t distribution.
(Key words)

Classificação AMS: 62F05, 62J05
(AMS Classification)

- Maio de 1993 -

Bartlett corrected tests for regression models with Student-t independent errors

Silvia L. de Paula Ferrari and Reinaldo B. Arellano

Universidade de São Paulo,
Caixa Postal 20570, Agência Iguatemi,
01498-970 São Paulo, Brazil

Abstract

In this paper we develop Bartlett corrected likelihood ratio and score tests for regression linear models under the assumption that the error terms are independent and have Student-t distribution. We discuss the application of the EM algorithm for finding maximum likelihood estimates of the parameters. We also present simulation results comparing the performance of the original tests and their corrected versions.

Some key words: Asymptotic expansion; Bartlett correction; chi-squared distribution; EM algorithm; likelihood ratio test; maximum likelihood; regression model; score test; Student-t distribution.

1. Introduction

The likelihood ratio and the score statistics are widely used in testing problems involving large samples. It is well known that, under usual regularity conditions, both statistics have chi-squared distribution asymptotically. In order to improve the chi-squared approximation for the likelihood ratio statistic, one may multiply it by a suitable constant known as Bartlett correction. Under the null hypothesis, the modified statistic obtained this way has chi-squared distribution up to order n^{-1}, where n is the sample size. See Lawley (1956), Hayakawa (1977), Cordeiro (1987) and Barndorff-Nielsen and Cox (1984). Here and from now on, "up to order n^{-1}" means that terms of order less than n^{-1} are ignored. Recently, Cordeiro and Ferrari (1991) derived a Bartlett-type corrected score statistic having chi-squared distribution up to order n^{-1}. The corrections for both statistics are functions of joint cumulants of log-likelihood derivatives and, therefore, it can be very arduous to obtain these corrections in some particular cases. Matrix formulae for Bartlett corrections for the likelihood ratio statistic in generalized linear models and their extensions were obtained by Cordeiro (1983), Cordeiro and Paula (1989) and Cordeiro, Paula and Botter (1993) and for heteroskedastic normal regression models by Cordeiro (1993). Moreover, matrix formulae for Bartlett-type corrections for the score test in generalized linear models were recently obtained by Cordeiro, Ferrari and Paula (1993) and for the test of heteroskedasticity in normal regression models by Cribari-Neto and Ferrari (1992).

The main purpose of this paper is to develop corrections for the likelihood ratio and the score statistics in linear regression models when the errors are independent with Student-t distribution with ν degrees of freedom. The corrections are given in simple matrix form and can easily be implemented into computer programs. Moreover, they can be used analitically in order to obtain simple expressions for the corrections in special cases.

The t-model is a robust extension of the normal regression theory since it adds an extra parameter (ν) to model the kurtosis of the error distribution. The usefulness of the t-family for modeling data was demonstrated by Lange, Little and Taylor (1989). They illustrated the use of the t-family in a variety of settings including linear and non-linear regression, repeated measures and pedigree data. Moreover, they described three possible algorithms, including the EM algorithm, for finding maximum likelihood estimates of the parameters. The efficiency of the estimates and the coverage rates of confidence intervals of quantities of interest, when the extra parameter ν is estimated from the data and when it is treated as fixed, are studied by Taylor (1992). Properties and other recent applications of the t-family can be found in the papers by Mitchell (1989), Sutradhar and Ali (1986) and Albert, Delampady and Polasek (1991). However, the importance in statistical inference of the elliptic family, of which the t-family is a particular member, has been emphasized since the 70s (see Maronna, 1976). A comprehensive review about the t-family is given by Chmielewski (1981).

The plan of this paper is as follows. In Section 2 we describe the linear regression model with Student-t independent errors and present some results on estimation and tests for this model. We discuss the application of the EM algorithm for maximum likelihood estimation and develop likelihood ratio and score tests for the regression parameters. In Section 3 we develop corrections for these tests and in Section 4 we present applications for some special cases. In Section 5 simulation experiments are performed in order to compare the original tests and their corrected versions.

2. Estimation and tests in the t-model

2.1. The model

We consider the regression model

$$y = \mu + \sigma t,$$

where $y = (y_1, \ldots, y_n)'$, $\mu = X\beta$, X is an $n \times p$ nonstochastic matrix with rank $p < n$, $\beta = (\beta_1, \ldots, \beta_p)'$ is a $p \times 1$ vector of unknown regression parameters, $t = (t_1, \ldots, t_n)'$ is an $n \times 1$ vector of independent errors and $\sigma^2 = \phi^{-1}$ is a positive scale parameter. A usual assumption is that t_l, for $l = 1, \ldots, n$, have normal distribution with zero mean and unit variance. Instead, we replace it by the assumption that t_l, for $l = 1, \ldots, n$, have Student-t distribution with ν degrees of freedom. In other words, we assume that

$$y_l \overset{\text{ind}}{\sim} t(x_l'\beta, \phi, \nu), \quad l = 1, \ldots, n, \tag{1}$$

where $x_l' = (x_{l1}, \ldots, x_{lp})$ is the l-th row of X and $t(\mu, \phi, \nu)$ denotes the Student-t distribution with location parameter μ, scale parameter ϕ, ν degrees of freedom and with density

$$f(y; \mu, \phi, \nu) = \left(\frac{\phi}{\nu}\right)^{1/2} \frac{\Gamma\{(\nu+1)/2\}}{\Gamma(1/2)\Gamma(\nu/2)} \left(1 + \frac{\phi}{\nu}(y-\mu)^2\right)^{-(\nu+1)/2}, \quad -\infty < y < \infty.$$

The t-model includes the normal and Cauchy models as special cases ($\nu = \infty$ and $\nu = 1$, respectively). Hence, it is more appropriate when the error distribution has longer-than-normal tails. When $\nu < \infty$, maximum likelihood (ML) estimation of β and ϕ are robust in the sense that the observations with large squared distances $t_l^2 = \phi(y_l - \sum_{j=1}^{p} x_{lj}\beta_j)^2$ are downweighted. In particular, the ML estimates of β_j, for $j = 1, \ldots, p$, for the normal model satisfy the likelihood equations $\sum_{l=1}^{n}(y_l - \sum_{j=1}^{p} x_{lj}\beta_j)x_{lj} = 0$, for $j = 1, \ldots, p$, whereas the ML estimates of β_j, for $j = 1, \ldots, p$, based in the t-model (1) satisfy $\sum_{l=1}^{n} w_l(y_l - \sum_{j=1}^{p} x_{lj}\beta_j)x_{lj} = 0$, where $w_l = (\nu+1)/(\nu + t_l^2)$ is the weight assigned to the l-th observation. It is clear that w_l decreases with increasing t_l^2. Moreover, the degree of downweighting of outliers in w_l increases with decreasing ν. If ν is fixed at a reasonable value, it is a robustness tuning parameter and the ML estimation of β based on the t-model (1) is a kind of M estimation yielding robust estimates of location with a redescending influence function (Huber, 1981). When the sample is sufficiently large, ν may be estimated from the data by ML, yielding an adaptive procedure in the same sense used by Hogg (1974) in which, loosely speaking, the choice of the estimation method is based on the observed residual distribution (see Yu and Hogg, 1988, Lange, Little and Taylor, 1989 and Taylor, 1992).

The model (1) is different from that of Zellner (1976) who proposed a multivariate t distribution for the vector of errors (see also Singh, 1991). While (1) yields robust estimates of β, Zellner's model yields the standard least squares estimates of β with estimated standard errors which are inflated by the factor $\{\nu/(\nu-2)\}^{1/2}$, for $\nu > 2$. Moreover, in this case the likelihood ratio, the score and Wald's tests for the general hypothesis $H_0 : A\beta = \mathbf{0}$ are equivalent to the usual F-tests (see Ghosh and Sinha, 1980 and Ullah and Zinde-Walsh, 1981, 1985).

2.2. Estimation

The estimation method considered in this paper is ML. The log-likelihood function for the linear regression t-model (1) and a sample $y_1, \ldots, y_n$ is

$$L(\theta) = \sum_{l=1}^{n} \log f(y_l; \theta),$$

where

$$\log f(y_l;\theta) = \frac{\nu}{2}\log\nu + \log\Gamma\Big(\frac{\nu+1}{2}\Big) - \log\Gamma\Big(\frac{\nu}{2}\Big) - \log\Gamma\Big(\frac{1}{2}\Big) + \frac{1}{2}\log\phi - \frac{\nu+1}{2}\log\{\nu+\phi(y_l - x_l'\beta)^2\}$$

and $\theta = (\beta', \phi, \nu)'$. When ν is known $\theta = (\beta', \phi)'$ and the log-likelihood equations are

$$\frac{\partial L}{\partial \beta_j}\Big|_{\hat\theta} = \hat\phi \sum_{l=1}^{n} \hat w_l x_{lj}(y_l - x_l'\hat\beta) = 0, \quad \text{for } j = 1, \ldots, p,$$

$$\frac{\partial L}{\partial \phi}\Big|_{\hat\theta} = \frac{n}{2\hat\phi} - \frac{1}{2}\sum_{l=1}^{n} \hat w_l (y_l - x_l'\hat\beta)^2 = 0,$$

where $\hat\theta = (\hat\beta', \hat\phi)'$ and $\hat w_l = w_l(\hat\theta)$ are the ML estimates of $\theta = (\beta', \phi)'$ and $w_l(\theta) = (\nu+1)/\{\nu + \phi(y_l - x_l'\beta)^2\}$. These equations are solved iteratively using, for example, the scoring or the EM algorithm. The implementation of the scoring algorithm involves the information matrix $K = K(\theta)$. For known ν this matrix is given by $K = \mathrm{diag}\{(\nu+1)(\nu+3)^{-1}\phi X'X,\ n\nu\{2(\nu+3)\}^{-1}\phi^{-2}\}$. When ν is treated as an unknown parameter, it is easy to show that K is block diagonal between the regression parameters β and the remaining parameters (ϕ and ν). Hence, the ML estimation of β and that of (ϕ, ν) are asymptotically uncorrelated and the asymptotic standard errors of the estimates of β are unaffected by estimating the scale parameter ϕ or the degrees of freedom ν (see Appendix in Lange, Little and Taylor, 1989).

The EM algorithm (Dempster, Laird and Rubin, 1977, Little and Rubin, 1987; see also Little, 1988) augments the data $y = (y_1, \ldots, y_n)'$ by additional hypothetical data $q = (q_1, \ldots, q_n)'$ in such a way that the ML estimates of θ, given the complete data (y', q'), are easy to compute. Given the estimate $\theta^{(t)}$ at iteration t, the $(t+1)$-th iteration of the EM algorithm consists of an expectation (E) step and a maximization (M) step. The E step computes the expected value of the complete data log-likelihood with respect to the conditional distribution of q, given y and $\theta^{(t)}$. The M step maximizes the resulting function with respect to θ, yielding the new estimate $\theta^{(t+1)}$. The implementation of the EM algorithm is easy because the Student-t distribution may be obtained as a mixture of a normal distribution in a chi-squared distribution. In other words, if the vectors (y_l, q_l), for $l = 1, \ldots, n$, are independent and such that $y_l \mid q_l \sim N(x_l'\beta, (q_l\phi)^{-1})$ with $q_l \sim \nu^{-1}\chi^2_\nu$, then the marginal distribution of y_l is given by (1). Hence, the ML estimation for model (1) may be achieved by applying the EM algorithm with missing data q_l, for $l = 1, \ldots, n$. If ν is assumed to be known, the EM algorithm is an iteratively reweighted least squares procedure. The E step computes the weights

$$w_l^{(t)} = E\{q_l \mid y_l; \beta^{(t)}, \phi^{(t)}\} = \frac{\nu+1}{\nu + \phi^{(t)}(y_l - x_l'\beta^{(t)})^2}, \quad l = 1, \ldots, n.$$

The M step finds $\beta^{(t+1)}$ that minimizes the weighted sum of squares $\sum w_l^{(t)}(y_l - x_l'\beta)^2$, i.e. $\beta^{(t+1)} = (X'W^{(t)}X)^{-1}X'W^{(t)}y$, where $W^{(t)} = \mathrm{diag}\{w_1^{(t)}, \ldots, w_n^{(t)}\}$ and, if ϕ is unknown, computes also $\phi^{(t+1)} = \{n^{-1}\sum w_l^{(t)}(y_l - x_l'\beta^{(t+1)})^2\}^{-1}$. Since the M step is a weighted least squares estimation, it is noniterative if the regression is linear. Starting values for $\theta = (\beta', \phi)'$ are obtained by fitting a normal model by least squares, i.e. with $w_l^{(0)} = 1$, for $l = 1, \ldots, n$. If ν is treated as an unknown parameter, it may be estimated by repeating the aforementioned algorithm over a grid of values of ν. For an alternative method, see Appendix in Lange, Little and Taylor (1989).

2.3. Hypothesis testing

Let us assume that the parameters ν and ϕ are known and then $\theta = \beta$. In Section 3.2 we consider the case where ϕ is assumed to be unknown.

Consider the test of the hypothesis $H_0 : \beta_1 = \beta_1^{(0)}$ against the alternative $H : \beta_1 \neq \beta_1^{(0)}$, where $\beta_1 = (\beta_1, \ldots, \beta_r)'$ with $r \leq p$. Here the vector $\beta_2 = (\beta_{r+1}, \ldots, \beta_p)'$ denotes a nuisance parameter and $\beta_1^{(0)}$ denote an r-dimensional vector of known constants. Let $\hat{\beta}$ be the maximum likelihood estimate of β under the alternative hypothesis and let $\tilde{\beta}_2$ be the maximum likelihood estimate of β_2 under the null hypothesis. Functions evaluated at the point $\hat{\beta}$ will be written with a circumflex and functions evaluated at $\tilde{\beta} = (\beta_1^{(0)\prime}, \tilde{\beta}_2')'$ will be distinguished by the addition of a tilde.

The partition $\beta = (\beta_1', \beta_2')'$ leads to the corresponding partitioned design matrix $X = (X_1,\ X_2)$ and the partitioned information matrix and its inverse

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \text{ and } K^{-1} = \begin{pmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{pmatrix},$$

Here $K_{11} = (\nu+1)(\nu+3)^{-1}\phi X_1' X_1$, $K_{22} = (\nu+1)(\nu+3)^{-1}\phi X_2' X_2$ and $K_{12} = K_{21}' = (\nu+1)(\nu+3)^{-1}\phi X_1' X_2$. We also define $Z = \{z_{lm}\} = X(X'X)^{-1}X'$ and $Z_2 = \{z_{2lm}\} = X_2(X_2'X_2)^{-1}X_2'$, for $i, j = 1, \ldots, n$. It is noteworthy that $\phi^{-1}(\nu+3)(\nu+1)^{-1}Z$ and $\phi^{-1}(\nu+3)(\nu+1)^{-1}Z_2$ have simple interpretations as the asymptotic covariance matrices of $X\hat{\beta}$ and $X_2\tilde{\beta}_2$, respectively.

Let $U = (\partial L/\partial\beta_1, \ldots, \partial L/\partial\beta_p)'$ be the score vector. The likelihood ratio statistic $LR = 2(L(\hat{\beta}) - L(\tilde{\beta}))$ and the score statistic $S = \tilde{U}'\tilde{K}^{-1}\tilde{U}$ for testing H_0 against H may be written respectively as

$$LR = (\nu + 1)\sum_{l=1}^{n} \log\Big(\frac{\hat{w}_l}{\tilde{w}_l}\Big) \tag{2}$$

and

$$S = \tilde{s}' X_1 (R'R)^{-1} X_1' \tilde{s}, \tag{3}$$

where $s = \{(\nu + 3)(\nu + 1)^{-1}\phi\}^{1/2} W'(y - X\beta)$ is a vector of standardized residuals, $R = (I_n - Z_2)X_1$ with I_n representing the $n \times n$ identity matrix and $W = \text{diag}\{w_1, \ldots, w_n\}$ with

w_l given in Section 2.1. Notice that the score statistic does not involve estimation under the alternative hypothesis which represents a computational advantage over the likelihood ratio statistic.

If $n^{-1}X'X \rightarrow \Omega$, as $n \rightarrow \infty$, where Ω is a $p \times p$ non-singular matrix, both statistics have chi-squared distribution with r degrees of freedom asymptotically under the null hypothesis. However, in finite samples, the chi-squared approximation may be a poor one. In the next section, we derived modified likelihood and score statistics whose distributions are better approximated by the reference chi-squared distribution.

3. Bartlett corrected tests

The purpose of this section is to develop improved likelihood ratio and score tests for the hypothesis $H_0 : \beta_1 = \beta_1^{(0)}$ against the alternative $H : \beta_1 \neq \beta_1^{(0)}$ in the regression model defined in (1). Following Lawley (1956), in regular problems the likelihood ratio test may be improved by multiplying the statistic by a correction factor, the Bartlett correction. The modified statistic has the form

$$LR^* = LR(1-d), \tag{4}$$

where d equals r^{-1} multiplied by the term of order n^{-1} of the expansion of the first moment of LR under the null hypothesis. On the other hand, Cordeiro and Ferrari (1992) showed that, in regular problems, the score statistic may be improved by a Bartlett-type correction which is not exactly a Bartlett correction because it involves a polynomial in the original statistic. The modified score statistic is given by

$$S^* = S\{1-(c+bS+aS^2)\}, \tag{5}$$

where a, b and c come from the expansion of the distribution function of S under the null hypothesis and are of order n^{-1}. Both modified statistics have chi-squared distribution with r degrees of freedom up to order n^{-1} under the null hypothesis.

In order to present the coefficients a, b, c and d involved in the corrections, we now introduce some standard notation for cumulants of total log-likelihood derivatives (Lawley, 1956, Hayakawa, 1977, Harris, 1985). In what follows, all the suffices take the range $1, \ldots, p$.

Let $U_i = \partial L/\partial\beta_i$, $U_{ij} = \partial^2 L/\partial\beta_i\partial\beta_j$ and so on. We have $\kappa_{ij} = E(U_{ij})$, $\kappa_{ijk} = E(U_{ijk})$, $\kappa_{i,j} = E(U_iU_j)$, $\kappa_{ij,k} = E(U_{ij}U_k)$, $\kappa_{ij,kr} = E(U_{ij}U_{kr}) - \kappa_{ij}\kappa_{kr}$, $\kappa_{i,j,kr} = E(U_iU_jU_{kr}) - \kappa_{i,j}\kappa_{kr}$ and $\kappa_{i,j,k,r} = E(U_iU_jU_kU_r) - \kappa_{i,j}\kappa_{k,r} - \kappa_{i,k}\kappa_{j,r} - \kappa_{i,r}\kappa_{j,k}$. We define the derivatives of the cumulants by $\kappa_{i,j}^{(k)} = \partial\kappa_{i,j}/\partial\beta_k$, $\kappa_{i,j}^{(kr)} = \partial^2\kappa_{i,j}/\partial\beta_k\partial\beta_r$, etc. Under standard regularity conditions the cumulants satisfy certain relations which may facilitate their calculation, such as $\kappa_{i,j,k} + \kappa_{ijk} = -\kappa_{i,jk} - \kappa_{j,ik} - \kappa_{k,ij}$, $\kappa_{i,j,k}^{(r)} = \kappa_{ijkr} + \kappa_{ijk,r}$, $\kappa_{i,j,k} = -\kappa_{ijk} - \sum_{(3)}^{*}\kappa_{i,jk}$ and $\kappa_{i,j,k,r} = -3\kappa_{ijkr} + 2\sum_{(4)}^{*}\kappa_{ijk}^{(r)} - \sum_{(6)}^{*}\kappa_{ij}^{(kr)} + \sum_{(3)}^{*}\kappa_{ij,kr}$, where $\sum_{(k)}^{*}$ means that the summation is extended to all the k permutations of indices. The total information matrix K has elements $\kappa_{i,j} = -\kappa_{ij}$, and $\kappa^{i,j} = -\kappa^{ij}$ denotes the corresponding elements of its inverse.

Lawley (1956) derived an expansion for the expected value of $2(L(\hat{\beta}) - L(\beta))$. He showed that, up to order n^{-1}, $E\{2(L(\hat{\beta}) - L(\beta))\} = p + \epsilon_p$, where ϵ_p is a term of order n^{-1} given by

$$\epsilon_p = \sum{}' \{l_{ijkr} - l_{ijkrst}\}, \tag{6}$$

where $\sum'$ denotes the summation over the parameters,

$$l_{ijkr} = \kappa^{ij}\kappa^{kr}\Big(\frac{1}{4}\kappa_{ijkr} - \kappa_{ijk}^{(r)} + \kappa_{ij}^{(kr)}\Big) \tag{7}$$

and

$$l_{ijkrst} = \kappa^{ij}\kappa^{kr}\kappa^{st}\Big\{\kappa_{iks}\Big(\frac{1}{6}\kappa_{jrt} - \kappa_{jt}^{(r)}\Big) + \kappa_{ikr}\Big(\frac{1}{4}\kappa_{jst} - \kappa_{jt}^{(s)}\Big) + \kappa_{ik}^{(s)}\kappa_{jt}^{(r)} + \kappa_{ik}^{(r)}\kappa_{jt}^{(s)}\Big\}. \tag{8}$$

The notation here follows Cordeiro (1983).

For the test of the hypothesis $H_0 : \beta_1 = \beta_1^{(0)}$ against the alternative $H : \beta_1 \neq \beta_1^{(0)}$, the likelihood ratio statistic may be written as $LR = 2(L(\hat{\beta}) - L(\beta_1^{(0)}, \beta_2)) - 2(L(\tilde{\beta}) - L(\beta_1^{(0)}, \beta_2))$ and therefore its expected value up to order n^{-1} is $E(LR) = r + \epsilon_p - \epsilon_{p-r}$, where ϵ_{p-r} is obtained in the same way as ϵ_p with the suffices in the summation taking the range $r+1, \ldots, p$. Now it is easy to verify that the coefficient d in the Bartlett correction factor (see eq. (4)) for the likelihood ratio statistic is

$$d = \frac{\epsilon_p - \epsilon_{p-r}}{r}, \tag{9}$$

where ϵ_p and ϵ_{p-r} are evaluated at $(\beta_1^{(0)}, \beta_2)$.

From an expansion for the distribution function of S derived by Harris (1985), Cordeiro and Ferrari (1991) obtained the coefficients a, b and c which make the modified score statistic in (5) have chi-squared distribution up to order n^{-1} under H_0. They obtained

$$a = \frac{A_3}{12r(r+2)(r+4)}, \quad b = \frac{A_2 - 2A_3}{12r(r+2)}, \quad c = \frac{A_1 - A_2 + A_3}{12r}, \tag{10}$$

where

$$\begin{aligned} A_1 = {} & 3\sum{}'(\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t})a_{ij}a_{st}m_{kr} - 6\sum{}'(\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}a_{kr}m_{st} \\ & +6\sum{}'(\kappa_{i,jk} - \kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{rs,t})a_{js}a_{kt}m_{ir} - 6\sum{}'(\kappa_{i,j,k,r} + \kappa_{i,j,kr})a_{kr}m_{ij}, \end{aligned} \tag{11}$$

$$\begin{aligned} A_2 = {} & -3\sum{}'\kappa_{i,j,k}\kappa_{r,s,t}a_{kr}m_{ij}m_{st} + 6\sum{}'(\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}m_{kr}m_{st} \\ & -6\sum{}'\kappa_{i,j,k}\kappa_{r,s,t}a_{kt}m_{ir}m_{js} + 3\sum{}'\kappa_{i,j,k,r}m_{ij}m_{kr} \end{aligned} \tag{12}$$

and

$$A_3 = 3\sum{}'\kappa_{i,j,k}\kappa_{r,s,t}m_{ij}m_{kr}m_{st} + 2\sum{}'\kappa_{i,j,k}\kappa_{r,s,t}m_{ir}m_{js}m_{kt}. \tag{13}$$

Here a_{ij} and m_{ij}, for $i, j = 1, \ldots, p$, denote the elements of the $p \times p$ matrices

$$A = \begin{pmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{pmatrix} \quad \text{and} \quad M = K^{-1} - A,$$

where K_{22} comes from the partitioned matrix K (see Section 2.3). It should be noted that in Harris' expression for A_2 the term $K_{\cdot,\cdot,\cdot} * M * M * J * K_{\cdot,\cdot,\cdot,\cdot}$ should be replaced by $K_{\cdot,\cdot,\cdot,\cdot} * M * M * J * K_{\cdot,\cdot,\cdot}$. The coefficients a, b, c and d may be functions of unknown parameters. If this is the case, those parameters may be replaced by their corresponding estimates under the null hypothesis since the order of approximation for the distribution functions of the modified statistics by the chi-squared distribution remains the same (see Lawley, 1956 and Cordeiro and Ferrari, 1991).

We now aim to develop matrix formulae for a, b, c and d for the test of the hypothesis mentioned in the begining of this section for the linear regression model defined in (1). The formulae are easy to handle in computer programs because they involve only simple operations on matrices and vectors. Besides, they are useful to obtain simple expressions for the corrections in particular cases. First, we assume that the scale parameter ϕ is known and then we relax this assumption.

3.1. Known scale parameter

We shall introduce the notation: $\partial\mu_l/\partial\beta_i = x_{li} = (i)_l$, $\partial^2\mu_l/\partial\beta_i\partial\beta_j = x_{li}x_{lj} = (ij)_l$ and so on. The computation of the cumulants involved in the corrections for the likelihood ratio and score tests depend on expectations of the form $E\{t^j(\nu+t^2)^{-i}\}$, where t has Student-t distribution with ν degrees of freedom and $i = 1,\ldots,6$ and $j = 0,\ldots,8$. It is easy to show that for $\nu + 2(i-j) > 0$

$$E\{t^j(\nu+t^2)^{-i}\} = \begin{array}{ll} \frac{1}{\nu^{i-j}}\frac{B\left(\frac{j+1}{2},\,\frac{\nu+2(i-j)}{2}\right)}{B\left(\frac{1}{2},\,\frac{\nu}{2}\right)}, & \text{if } j \text{ is even} \\ 0, & \text{if } j \text{ is odd,} \end{array}$$

where $B(.,.)$ is the beta function. By using this expression and the relations among cumulants introduced in Section 3, we obtain

$$\begin{aligned} \kappa_{ij} &= -\kappa_{i,j} = -\phi\frac{\nu+1}{\nu+3}\sum(ij)_l, \quad \kappa_{ijk} = \kappa_{i,jk} = \kappa_{i,j,k} = \kappa_{ij}^{(k)} = 0, \\ \kappa_{ijkr} &= 6\phi^2\frac{(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}\sum(ijkr)_l, \quad \kappa_{ij,kr} = \frac{3(\nu+3)^2-8}{(\nu+2)(\nu+3)^2}\kappa_{ijkr}, \\ \kappa_{i,j,kr} &= \frac{(\nu+1)\{(\nu+2)^2-5\}}{(\nu+2)(\nu+3)^2}\kappa_{ijkr}, \quad \kappa_{i,j,k,r} = -3\kappa_{i,j,kr}, \quad \kappa_{ij}^{(kr)} = \kappa_{ijk}^{(r)} = 0, \end{aligned} \tag{14}$$

where $\sum$ denotes the summation over the sample.

Let $Z_d = \text{diag}\{z_{11},\ldots,z_{nn}\}$ and $Z_{2d} = \text{diag}\{z_{211},\ldots,z_{2nn}\}$ represent the matrices with the diagonal elements of Z and Z_2 respectively, $\rho_{ZZ} = n\text{tr}(Z_dZ_d)$, $\rho_{ZZ_2} = n\text{tr}(Z_dZ_{2d})$ and $\rho_{Z_2Z_2} = n\text{tr}(Z_{2d}Z_{2d})$. The matrices Z and Z_2 are defined in Section 2.3. From (9), (10) and (14) we get after some algebra (see Appendix)

$$d = \frac{3}{2nr}(\rho_{ZZ} - \rho_{Z_2Z_2})h_1, \tag{15}$$

$$c = \frac{3}{2nr}(3\rho_{ZZ} - 2\rho_{ZZ_2} - \rho_{Z_2Z_2})h_2, \tag{16}$$

$$b = -\frac{9}{2nr(r+2)}(\rho_{ZZ} - 2\rho_{ZZ_2} + \rho_{Z_2Z_2})h_2, \tag{17}$$

$$a = 0, \tag{18}$$

where $h_i = h_i(\nu)$ for $i = 1, 2$ and

$$h_1 = \frac{(\nu+2)(\nu+3)^2}{\nu(\nu+1)(\nu+5)(\nu+7)}, \quad h_2 = \frac{(\nu+2)^2-5}{\nu(\nu+5)(\nu+7)}. \tag{19}$$

Now from equations (4)-(5) and (15)-(19) one may easily compute the Bartlett corrected statistics.

Since the coefficient a equals zero, the quadratic term in the correction for the score statistic vanishes and therefore the correction is a polynomial of first degree in the unmodified statistic. The expressions for b, c, and d are functions of ν and the diagonal matrices Z_d and Z_{2d}. They depend neither on the scale parameter ϕ nor on the unknown parameters. Moreover, the expressions in (15)-(17) may be easily implemented in some computer programs since they involve only simple operations on matrices. It is important to emphasize that the corrections are valid for all positive ν and hence they include the Cauchy model as a special case. For the normal model ($\nu = \infty$), h_1 and h_2 vanish and no correction is obtained. This could be expected since the likelihood ratio and the score statistics have an exact chi-squared distribution under the null hypothesis.

3.2 Unknown scale parameter

In this section we assume that the scale parameter ϕ is unknown and develop matrix formulae for the corrections for the likelihood ratio and the score statistics which now also depend on cumulants involving derivatives of the log-likelihood function with respect to ϕ. Since the information matrix for (β', ϕ) is block diagonal (see Section 2.2), β and ϕ are globally orthogonal. The orthogonality between β and ϕ greatly simplify the computation of the corrections.

If ϕ is assumed to be unknown, the likelihood ratio statistic may be written as

$$LR = (\nu + 1) \sum_{l=1}^{n} \log\Big(\frac{\hat{w}_l}{\tilde{w}_l}\Big) + n\log\Big(\frac{\hat{\phi}}{\tilde{\phi}}\Big) \tag{20}$$

and the score statistic may be written as in equation (3) with ϕ replaced by $\tilde{\phi}$. Here $\hat{\phi} = \{n^{-1} \sum_{l=1}^{n} \hat{w}_l (y_l - x_l'\hat{\beta})^2\}^{-1}$ and $\tilde{\phi} = \{n^{-1} \sum_{l=1}^{n} \tilde{w}_l (y_l - x_l'\tilde{\beta})^2\}^{-1}$ are the maximum likelihood estimates of ϕ under the full model and the model restricted to the null hypothesis, respectively. Notice that the matrices $\tilde{W}$ and $\tilde{T}$ involved in S are evaluated at $(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\phi})$.

Let us introduce the following notation for cumulants involving the parameter ϕ:

$\kappa_{\phi\phi} = E(\partial^2 L/\partial\phi^2)$, $\kappa_{\phi i} = E(\partial^2 L/\partial\phi\partial\beta_i)$, $\kappa_{\phi ij} = E(\partial^3 L/\partial\phi\partial\beta_i\partial\beta_j)$, and so on. We get

$$\begin{aligned}
\kappa_{\phi\phi} &= -\frac{1}{\phi^2}\frac{n}{2}\frac{\nu}{\nu+3}, \quad \kappa_{\phi\phi}^{(\phi)} = -\frac{2}{\phi}\kappa_{\phi\phi}, \quad \kappa_{\phi\phi}^{(\phi\phi)} = \frac{6}{\phi^2}\kappa_{\phi\phi}, \quad \kappa_{\phi,\phi\phi} = \frac{6}{\phi}\frac{1}{\nu+5}\kappa_{\phi\phi},\\
\kappa_{\phi,\phi,\phi} &= \frac{1}{\phi}\frac{\nu-1}{\nu+5}\kappa_{\phi\phi}, \quad \kappa_{\phi\phi\phi} = -\frac{2}{\phi}\frac{(\nu+8)}{(\nu+5)}\kappa_{\phi\phi}, \quad \kappa_{\phi\phi\phi}^{(\phi)} = -\frac{6}{\phi^2}\frac{\nu+8}{\nu+5}\kappa_{\phi\phi},\\
\kappa_{\phi\phi\phi\phi} &= \frac{6}{\phi^2}\frac{(\nu+7)(\nu+8)+15}{(\nu+5)(\nu+7)}\kappa_{\phi\phi}, \quad \kappa_{ij}^{(\phi)} = \frac{1}{\phi}\kappa_{ij}, \quad \kappa_{\phi ij} = -\frac{1}{\phi}\frac{\nu+2}{\nu+5}\kappa_{ij},\\
\kappa_{\phi i} &= \kappa_{i\phi\phi} = \kappa_{i,\phi\phi} = \kappa_{i,\phi,\phi} = \kappa_{\phi ij}^{(k)} = \kappa_{\phi ij}^{(\phi)} = 0, \quad \kappa_{\phi,i,j} = \frac{1}{\phi}\frac{\nu-1}{\nu+5}\kappa_{ij}, \qquad (21)\\
\kappa_{\phi i,j} &= -\kappa_{\phi ij}, \quad \kappa_{\phi,ij} = \frac{3}{\phi}\frac{1}{\nu+5}\kappa_{ij}, \quad \kappa_{\phi\phi ij} = -\frac{6}{\phi}\frac{1}{\nu+7}\kappa_{\phi ij},\\
\kappa_{\phi i,\phi j} &= -\frac{1}{\phi}\frac{(\nu+2)(\nu+4)}{(\nu+5)(\nu+7)}\kappa_{ij}, \quad \kappa_{\phi,\phi,i,j} = \frac{4}{n}\frac{(\nu-3)(\nu^2+3\nu-1)}{\nu(\nu+5)(\nu+7)}\kappa_{\phi\phi}\kappa_{ij},\\
\kappa_{\phi\phi,i,j} &= \frac{12}{n}\frac{(\nu+1)^2-6}{\nu(\nu+5)(\nu+7)}\kappa_{\phi\phi}\kappa_{ij},
\end{aligned}$$

where κ_{ij} is given in (14). After lengthy algebra (see Appendix) we get

$$d = \frac{3}{2nr}(\rho_{ZZ} - \rho_{Z_2Z_2})h_1 + \frac{1}{n}h_3 + \frac{2p-r}{2n}h_4 \qquad (22)$$

$$c = \frac{3}{2nr}(3\rho_{ZZ} - 2\rho_{ZZ_2} - \rho_{Z_2Z_2})h_2 + \frac{1}{n}(p-r)h_5 + \frac{6}{n}h_6 + \frac{r+2}{2n}h_7, \qquad (23)$$

$$b = -\frac{9}{2nr(r+2)}(\rho_{ZZ} - 2\rho_{ZZ_2} + \rho_{Z_2Z_2})h_2 - \frac{1}{2n}h_7 \qquad (24)$$

$$a = 0, \qquad (25)$$

where h_1 and h_2 are defined in (19), $h_i = h_i(\nu)$, for $i = 3, \ldots, 7$, and

$$\begin{aligned}
h_3 &= \frac{(\nu+2)(\nu+3)(\nu^2+9\nu+2)}{\nu(\nu+5)^2(\nu+7)}, \quad h_4 = \frac{(\nu+2)^2(\nu+3)}{\nu(\nu+5)^2},\\
h_5 &= \frac{(\nu-1)(\nu+2)(\nu+3)}{\nu(\nu+5)^2}, \quad h_6 = \frac{(\nu+1)(\nu+2)(\nu+3)}{\nu(\nu+5)^2(\nu+7)} \qquad (26)\\
h_7 &= \frac{(\nu-1)^2(\nu+3)}{\nu(\nu+5)^2}.
\end{aligned}$$

It is interesting to notice that $a = 0$ regardless of whether or not ϕ is assumed to be known and that the expressions in (22)-(24) may be written as the corresponding expressions in (15)-(17), obtained for known ϕ, plus additional terms. Those extra terms are very simple and depend only on the sample size, the parameter ν, the total number of regression parameters β and the number of restrictions fixed by the null hypothesis. Moreover, for normal linear regression models ($\nu = \infty$), $h_i = 0$, for $i = 1, 2, 6$ and $h_i = 1$, for $i = 3, 4, 5, 7$ and therefore the modified statistics have the simple forms

$$LR^*_{(\infty)} = LR_{(\infty)}\Big\{1 - \frac{1}{2n}(2p - r + 2)\Big\}, \tag{27}$$

and

$$S^*_{(\infty)} = S_{(\infty)}\Big\{1 - \frac{1}{2n}(2p - r + 2 - S_{(\infty)})\Big\}. \tag{28}$$

(Here and from now on, we make use of the notation $LR_{(\nu)}$ and $S_{(\nu)}$ whenever a value for ν is specified; for instance, $LR_{(\infty)}$ and $LR^*_{(\infty)}$ denote the likelihood ratio statistic and its modified version for normal models.) It is noteworthy that, in this case, the correction terms only depend on the sample size, the total number of parameters and the number of parameters fixed at the null hypothesis. Consequently, the corrections are very easy to compute.

4. Some special cases

In this section we derive simple expressions for the corrections for the likelihood ratio and the score statistics from the results of the previous section for some special models. Since in practice the scale parameter ϕ is usually unknown, we concentrate on the formulae derived in Section 3.2 for unknown ϕ. Notice that for all cases presented below, if the model is normal, the corresponding modified statistics come directly from (27) and (28) with appropriate values for p and r.

Consider first the following simple model

$$y_l \overset{\text{ind}}{\sim} t(\beta, \phi, \nu), \quad l = 1, \ldots, n,$$

where β is an unknown scalar parameter and suppose that we are interested in testing the null hypothesis $H_0 : \beta = \beta^{(0)}$ against the alternative hypothesis $H : \beta \neq \beta^{(0)}$. For this testing problem $p = r = 1$, $\rho_{ZZ} = 1$ and $\rho_{ZZ_2} = \rho_{Z_2Z_2} = 0$. Now from (22)-(24) we obtain $d = (3h_1 + 2h_3 + h_4)/(2n)$, $c = 3(3h_2 + 4h_6 + h_7)/(2n)$ and $b = -(3h_2 + h_7)/(2n)$, where h_i, for $i = 1, 2, 3, 4, 6, 7$, are defined in (19) and (26). In particular, for Cauchy models ($\nu = 1$) we have $LR^*_{(1)} = LR_{(1)}\{1 - 7/(4n)\}$ and $S^*_{(1)} = S_{(1)}\{1 - (7 - S_{(1)})/(8n)\}$.

Now consider $p \geq 2$ populations having Student-t distribution with ν degrees of freedom, scale parameter ϕ, and location parameter $\mu_i = \beta + \beta_i$, for $i = 1, \ldots, p$, where $\beta_p = 0$, and assume that independent random samples of sizes $n_1, \ldots, n_p$, with $n_i \geq 1$ for $i = 1, \ldots, p$ are taken from these populations. Here β_i represents the effect of the response of the $i-$th population compared to the $p-$th population. For testing the homogeneity of the location parameter in the p populations we consider the null hypothesis $H_0 : \beta_1 = \ldots = \beta_{p-1} = 0$. For this testing problem we have $r = p - 1$, $\rho_{ZZ} = \rho = n \sum_{i=1}^{p} n_i^{-1}$, $\rho_{ZZ_2} = p$ and $\rho_{Z_2Z_2} = 1$. Using the equations (22)-(24) we obtain $d = \{3(\rho - 1)(p - 1)^{-1} h_1 + 2h_3 + (p + 1)h_4\}/(2n)$, $c = \{3(3\rho - 2p - 1)(p - 1)^{-1} h_2 + 2(h_5 + 6h_6) + (p + 1)h_7\}/(2n)$ and $b = -\{9(\rho - 2p + 1)(p^2 - 1)^{-1} h_1 + h_7\}/(2n)$, where h_i, for $i = 1, \ldots, 7$, are defined in (19) and (26) and $n = \sum_{i=1}^{p} n_i$. For Cauchy models the modified statistics have the simple forms $LR^*_{(1)} = LR_{(1)}\{1 - \{3\rho + 2p(p + 1) - 7\}/\{(4n(p - 1))\}\}$ and $S^*_{(1)} = S_{(1)}\{1 - \{(3\rho + 2p - 5)(p + 1) - 3(\rho - 2p + 1)S_{(1)}\}/\{8n(p^2 - 1)\}\}$.

Finally, let us consider the simple regression model

$$y_l \overset{\text{ind}}{\sim} t(\alpha + x_l\beta, \phi, \nu), \quad l = 1, \ldots, n, \tag{29}$$

where α and β are unknown scalar parameters and x_l is a scalar covariate. For testing the null hypothesis $H_0 : \beta = 0$ against the alternative hypothesis $H_0 : \beta \neq 0$ we have $p = 2$, $r = 1$, $\rho_{ZZ} = \gamma + 6$, $\rho_{ZZ_2} = 2$ and $\rho_{Z_2Z_2} = 1$, where $\gamma = (\bar{s}_4/\bar{s}_2^2) - 3$, with $\bar{s}_a = n^{-1} \sum_{l=1}^{n} (x_l - \bar{x})^a$, for $a = 2, 4$, and $\bar{x} = n^{-1} \sum_{l=1}^{n} x_l$. Notice that γ is the standard sample measure of kurtosis of the covariate x. Now using the equations (22)-(24) we have $d = \{3(\gamma + 3)h_1 + 6h_1 + 2h_3 + 3h_4\}/(2n)$, $c = \{9(\gamma + 3)h_2 + 12h_2 + 2h_5 + 12h_6 + 3h_7\}/(2n)$ and $b = -\{3(\gamma + 3)h_2 + h_7\}/(2n)$. For Cauchy models we obtain $LR^*_{(1)} = LR_{(1)}\{1 - \{3(\gamma + 3) + 14\}/(4n)\}$ and $S^*_{(1)} = S_{(1)}\{1 - \{3(\gamma + 3) + 8 - (\gamma + 3)S_{(1)}\}/(8n)\}$.

5. Simulation results

In this section we present two small simulation studies comparing the performance of the likelihood ratio and the score statistics and their modified versions. In the first simulation experiment we deal with the test of the homogeneity of the location parameter in two independent populations (see Section 4) based on two samples of sizes $n_1 = n_2 = n/2$. We carried out size and power simulations based on 2,000 replications and they are reported for total samples sizes 10, 20, 30 and 40, for nominal sizes of 10%, 5%, 2.5% and 1% and for $\nu = 1$ and 3. The response was generated assuming that $\beta = 1$ and $\sigma = \phi^{-1/2} = 2$. The power simulations were performed supposing that $\beta_1 = 1$ and they are based on estimated critical values rather than tabulated ones. This strategy was chosen since the tests do not have equal sizes. The results are reported in Tables 1 and 2.

[Tables 1 and 2 here]

From Tables 1 and 2 it is clear that the likelihood ratio test tends to reject the null hypothesis more often than expected based on the nominal sizes. In particular, for $n = 10$ and $\nu = 1$ the differences between the simulated and the nominal sizes are very large. For all the sample sizes, the Bartlett correction seems to be very effective in pushing the true sizes of the test towards the nominal levels. Tables 1 and 2 convey important information about the score test. It is clear that, under H_0, the score statistic has better chi-squared approximation than the likelihood ratio statistic for all the sample sizes. Even if n is not large, the simulated sizes for the score test are not very far from the nominal levels. Therefore, in this case the use of the correction for the score test is not as important as it is when the likelihood ratio statistic is used. Although in most cases the modified score test has simulated sizes closer to the nominal levels than the unmodified score test, the effect of the correction is not very strong since the unmodified score test has a good chi-squared approximation. In other words, the use of corrections leads to more substantial improvement in the sizes of the test when the likelihood ratio statistic is used.

Now let us make some comments on the power of the tests. First, it is important to notice that the modified statistics are increasing functions of their corresponding unmodified versions. Although S^* is a second degree polynomial of S, it is easy to show that, for the case considered in this simulation experiment, S^* becomes an increasing function of S for $S > 0$ (for the simulation experiment reported below the same happens). For this reason and since the power comparisons are based on estimated critical values rather than tabulated ones, the powers of the unmodified tests and their corresponding modified versions coincide. Tables 1 and 2 show that for $\beta_1 = 1$ and $\nu = 1$ and 3, the powers of all the tests are very close and, therefore it is not possible to distinguish between them based on their powers at least for the value of β_1 considered here.

The second simulation experiment is based on the simple regression model (27) and the test of the null hypothesis $H_0 : \beta = 0$ against the alternative hypothesis $H : \beta \neq 0$. The corrections for the likelihood ratio and the score statistics depend on the sample measure of kurtosis of the covariate (see Section 4). Therefore, it is expected that the performance of the tests are affected by the choice of values for the covariate. Following Cribari-Neto and Ferrari (1992), the way we choose values for x is as follows. For $n = 10$ and a given distribution attributed to the covariate, we set

$$x_l = \text{quantile}\left(\frac{l - 0.5}{10}\right), \; l = 1, \ldots, 10.$$

For other sample sizes (n = 20, 30, 40) the values for the covariate are obtained via replication. For instance, if for $n = 10$ the values for the covariate are $x_1, x_2, \ldots, x_n$ then for $n = 20$ the chosen values for the covariate are $x_1, x_1, x_2, x_2 \ldots, x_n, x_n$. This strategy is very convenient because it delivers the same coefficient of kurtosis for different sample sizes. The following distributions are used: uniform in the interval (0,1) [U(0,1)] and lognormal obtained as a transformation of a standard normal distribution [LN(0,1)]. The resulting coefficients of kurtosis γ of the covariate are respectively -1.22 and 1.42. The response was generated setting $\alpha = 1$ and $\sigma = \phi^{-1/2} = 2$ and for the power simulations we set $\beta = 1$. As in the first simulation study, the power simulations are based on estimated critical values and all the simulations are performed with 2,000 replications. The results are reported in Tables 3, 4, 5 and 6.

[Tables 3, 4, 5 and 6 here]

From Tables 3 and 4 we may see that the results obtained when the U(0,1) distribution is used are, to some extent, similar to those obtained in the first simulation experiment. On the other hand, when the LN(0,1) distribution is considered, the results reported on Tables 5 and 6 deserve special attention. The simulated sizes of the likelihood ratio test for $n = 10$ and $\nu = 1$ are well above the corresponding nominal levels and the use of the

Bartlett factor tends to overcorrect the size of the test. This undesirable feature of the Bartlett correction could be expected in this case since the numerical value of the correction factor (0.318) indicates that the correction would be very strong. Moreover, according to Cox (1988, p. 330) "a large adjustment [to the likelihood ratio statistic] would have to be interpreted partly as a warning, in particular to investigate the distribution more carefully, perhaps by simulation". For $\nu = 1$ with $n \geq 20$ and for $\nu = 3$, the Bartlett correction is very effective and the modified likelihood ratio test has good chi-squared approximation. When the score test is used, it does not make much difference whether one chooses the original score statistic or its modified version. In terms of powers, the likelihood ratio test presented simulated powers greater than the score test for $\beta = 1$ and $\nu = 1$ and 3.

It should be noticed that the two simulation experiments are closely related. The results of the former may be obtained from the later by setting $x_1 = \ldots = x_5 = 1$ and $x_6 = \ldots = x_{10} = 0$ for $n = 10$ and by using replicated values for the covariate for the other sample sizes. In this case the coefficient of kurtosis equals -2.

For both simulation experiments we made use of the EM algorithm for finding the maximum likelihood estimates $(\tilde{\beta}, \tilde{\phi})$ and $(\hat{\beta}, \hat{\phi})$ of the unknown parameters as described in Section 2.2. In most cases, the algorithm converged in a few iterations. However, for some samples, it did not converge in at least 50 iterations. Those samples were neglected and replaced by other generated samples. This problem arose more frequently for power simulations with $\nu = 1$ and $n = 10$. The worst case happened when we set the covariate x as the quantiles of the U(0,1) distribution, where, for 102 out of 2,000 samples, the algorithm (for obtaining either $(\tilde{\beta}, \tilde{\phi})$ or $(\hat{\beta}, \hat{\phi})$) did not converge. For the size simulations, this was also the worst case (we did not achieve convergence for 54 samples). For $n \geq 20$, the convergence was achieved for the great majority of the cases for $\nu = 1$. For $\nu = 3$, the algorithm worked very well for both simulation experiments and all the samples sizes. We did not get convergence in only 3 cases (two cases in the power simulations and one case in the size simulations, both with $n = 10$). Therefore, if n is small and the data follow the Cauchy distribution, perhaps another algorithm should be applied although convergence of algorithms, such as the scoring and Newton Raphson algorithms, is not garanteed.

Appendix

In this Appendix we present the computation of the quantities a, b, c and d involved in the corrections for the likelihood and the score statistics. Assume first that ϕ is known. From (8) and (14) we have $l_{ijkrst} = 0$ and hence we obtain from (6) and (7) $\epsilon_p = \sum' l_{ijkr}$, where l_{ijkr} simplifies to $\kappa^{ij}\kappa^{kr}\kappa_{ijkr}/4$. Now ϵ_p may be written as

$$\epsilon_p = \frac{3}{2}\frac{(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}\phi^2 {\sum}' \kappa^{ij}\kappa^{kr} \sum (ijkr)_l,$$

where $\sum'$ and $\sum$ are the summations over the parameters and over the sample, respectively. Now, interchanging the order of the summations we get

$$\epsilon_p = \frac{3}{2}\frac{(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}\phi^2 \sum_l \Big(\sum_{i,j}(i)_l\kappa^{ij}(j)_l\Big)\Big(\sum_{k,r}(k)_l\kappa^{kr}(r)_l\Big).$$

By noting that $\sum_{i,j}(i)_l\kappa^{ij}(j)_l = -(\nu+3)(\nu+1)^{-1}\phi^{-1}z_{ll}$ we finally get

$$\epsilon_p = \frac{3}{2}h_1 tr(Z_d Z_d),$$

where h_1 is given in (19). In the same way, we may show that ϵ_{p-r} is written as ϵ_p with the matrix Z replaced by Z_2. Now one may easily get the expression for d in (15).

From (14) we may see that the expressions for the A's (see (11)-(13)) are reduced to

$$A_1 = -6{\sum}'(\kappa_{i,j,k,r} + \kappa_{i,j,kr})a_{kr}m_{ij}, \quad A_2 = 3{\sum}'\kappa_{i,j,k,r}m_{ij}m_{kr}$$

and $A_3 = 0$. Again from (14) and after rearranging the order of the summations we obtain

$$A_1 = 72\phi^2\Big(\frac{\nu+1}{\nu+3}\Big)^2\frac{(\nu+2)^2-5}{\nu(\nu+5)(\nu+7)}\sum_l\Big(\sum_{i,j}(i)_l m_{ij}(j)_l\Big)\Big(\sum_{k,r}(k)_l a_{kr}(r)_l\Big)$$

and

$$A_2 = -54\phi^2\Big(\frac{\nu+1}{\nu+3}\Big)^2\frac{(\nu+2)^2-5}{\nu(\nu+5)(\nu+7)}\sum_l\Big(\sum_{i,j}(i)_l m_{ij}(j)_l\Big)\Big(\sum_{k,r}(k)_l m_{kr}(r)_l\Big).$$

By writting the terms $\sum_{i,j}(i)_l a_{ij}(j)_l$ and $\sum_{i,j}(i)_l m_{ij}(j)_l$ as the (l,l)-elements of the matrices $\phi(\nu+3)(\nu+1)^{-1}Z$ and $\phi(\nu+3)(\nu+1)^{-1}(Z-Z_2)$ respectively we obtain

$$A_1 = 72h_2\mathrm{tr}\{(Z_d - Z_{2d})Z_{2d}\}, \quad A_2 = -54h_2\mathrm{tr}\{(Z_d - Z_{2d})(Z_d - Z_{2d})\},$$

where h_2 is given in (19). Finally, from (10) we get the expressions for a, b and c given in (16)-(18).

Now we assume that ϕ is unknown. Firstly, it is important to notice that, since the numbers of unknown parameters under H_0 and H are respectively $p-r+1$ and $p+1$, the quantity d in (9) should be replaced by $(\epsilon_{p+1}-\epsilon_{p-r+1})/r$. The expression for ϵ_{p+1} is given by (6) but with the summation taking the range $1,\ldots,p,\phi$. It is easy to see that

$$\epsilon_{p+1}=\epsilon_{p;\beta}+\epsilon_{p+1;\phi,\beta},$$

where $\epsilon_{p;\beta,\phi}$ is given by (6) with the summation taking the range $1,\ldots,p$ and $\epsilon_{p+1;\phi,\beta}=\sum'_{\phi,\beta}(l_{ijkr}-l_{ijkrst})$ with $\sum'_{\phi,\beta}$ representing the summation that takes the range $1,\ldots,p,\phi$ but with at least one index fixed at ϕ. The first term, i.e. $\epsilon_{p;\beta}$, coincides with ϵ_p obtained when ϕ is assumed to be known. The remaining term is obtained by plugging the expressions of the cumulants (see (21)) into (7) and (8) and interchanging the order of the summations over the sample and over the parameters. Since several cumulants equal zero (some of them due to the orthogonality of β and ϕ) some terms involved in the summations vanish and therefore the computations are simplified. After some algebra we get

$$\epsilon_{p;\phi,\beta}=h+\frac{p}{n}h_3+\frac{p^2}{2n}h_4,$$

where h_3 and h_4 are given in (26)and $h=h(n,\nu)=l_{\phi\phi\phi\phi}-l_{\phi\phi\phi\phi\phi\phi}$ does not depend on p and Z.

Now by writting

$$\epsilon_{p+1}-\epsilon_{p-r+1}=(\epsilon_{p;\beta}-\epsilon_{p-r;\beta})+(\epsilon_{p-1;\phi,\beta}-\epsilon_{p-r+1;\phi,\beta})$$

and by noting that $\epsilon_{p-r+1;\phi,\beta}$ equals $\epsilon_{p+1;\phi,\beta}$ with p replaced by $p-r$, we get after some algebra the expression in (22).

We proceed in a similar way for obtaining a, b and c. First, we notice that the A's may be written as

$$A_i=A_{i;\beta}+A_{i;\phi,\beta},\quad i=1,2,3,$$

where $A_{i;\beta}$ are the corresponding A's obtained when ϕ is assumed to be known (see (16)-(18)) and $A_{i;\phi,\beta}$ are given by (11)-(13) with the summations taking the range $1,\ldots,p,\phi$ with at least one index fixed at ϕ. In order to obtain $A_{i;\phi,\beta}$, for $i=1,2,3$ it should be noticed that $m_{\phi\phi}=0$ and that the orthogonality of ϕ and β also implies that $a_{i\phi}=m_{i\phi}=0$, for

$i = 1, \ldots, p$. In addition to that, several cumulants involved in $A_{i;\phi,\beta}$, for $i = 1, 2, 3$, equal zero (see (14) and (21)) and it then follows, from (11)-(13) that

$$\begin{aligned} A_{1;\phi,\beta} \;=\; & -6a_{\phi\phi}^{2}\sum{}'(\kappa_{\phi\phi\phi} + 2\kappa_{\phi,\phi\phi})\kappa_{\phi,i,j}m_{ij} \\ & +6a_{\phi\phi}\sum{}'(\kappa_{i,\phi j} - \kappa_{i,\phi,j})(\kappa_{k\phi r} + \kappa_{k\phi,r})a_{jr}m_{ik} \\ & -6a_{\phi\phi}\sum{}'(\kappa_{ij\phi} + 2\kappa_{i,j\phi})\kappa_{\phi,k,r}a_{ij}m_{kr} \\ & +6a_{\phi\phi}\sum{}'(\kappa_{i,j\phi} - \kappa_{i,j,\phi})(\kappa_{kr\phi} + 2\kappa_{kr,\phi})a_{jr}m_{ik} \\ & -6a_{\phi\phi}\sum{}'(\kappa_{i,j,\phi,\phi} + \kappa_{i,j,\phi\phi})m_{ij}, \end{aligned}$$

$$A_{2;\phi,\beta} = -3a_{\phi\phi}\sum{}'\kappa_{i,j,\phi}\kappa_{k,r,\phi}m_{ij}m_{kr} - 6a_{\phi\phi}\sum{}'\kappa_{i,j,\phi}\kappa_{k,r,\phi}m_{ik}m_{jr}$$

and $A_{3;\phi,\beta} = 0$, where $\sum'$ is representing the sum over the β parameters which means that the indices, except ϕ, vary from 1 to p. Now proceeding in the same way as described above we get the expressions for c, b and a in (23)-(25).

5. Acknowledgments

The authors would like to thank Francisco Cribari-Neto, Heleno Bolfarine and Denise A. Botter for helpful comments and especially Anthony Davison for insightful conversations that led to the idea for this paper. The financial support of CNPq/Brazil is also greatefully acknowledged.

References

ALBERT, J., DELAMPADY, M. and POLASEK, W. (1991). A class of distributions for robustness studies. *Journal of Statistic Planning and Inference*, **28**, 291-304.

CHMIELEWSKI, M.A. (1981). Elliptically symmetric distribution: a review and bibliography. *Applied Statistcs*, **49**, 67-74.

CORDEIRO, G.M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society* B, **45**, 404-13.

CORDEIRO, G.M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265-74.

CORDEIRO, G.M. (1993). Bartlett correction and bias corrections for two heteroscedastic regression models. *Communications in Statistics - Theory and Methods*, **22**, 169-88.

CORDEIRO, G.M. and FERRARI, S.L.P. (1991). A modified score statistic having chi-squared distribution to order n^{-1}. *Biometrika*, **78**, 573-82.

CORDEIRO, G.M., FERRARI, S.L.P. and PAULA, G.A. (1993). Improved score tests for generalized linear models. *Journal of the Royal Statistical Society* B, **55**. To appear.

CORDEIRO, G.M. and PAULA, G.A. (1989). Improved likelihood statistics for exponential linear models. *Biometrika*, **76**, 93-100.

CORDEIRO, G.M., PAULA, G.A. and BOTTER,D.A. (1993). Improved likelihood ratio tests for dispersion models. *International Statistical Review.* To appear.

COX, D.R. (1988). Some aspects of conditional and asymptotic inference: A review. *Sankhya,* A, 50, 314-37.

CRIBARI-NETO, F. and FERRARI, S.L.P. (1992). An improved Lagrange multiplier statistic for the test of heteroskedasticity. Technical Report RT-MAE-9210, IME - Universidade de São Paulo.

DEMPSTER, A.P. LAIRD, N.M. and RUBIN, O.R. (1977). Maximum likelihood for incomplete data (with discussion). *Journal of the Royal Statistical Society*, B **39**, 1-38.

GHOSH, M. and SINHA, B.K. (1980). On the robustness of least squares procedures in regression models. *Journal of Multivariate Analysis*, **10**, 332-42.

HARRIS, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653-9.

HAYAKAWA, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics* A, **29**, 359-78.

HOGG, R.V. (1974). Adaptative robust procedures: a partial and some suggestions for future applications and theory (with discussion). *Journal of the American Statistical Association*, **69**, 909-26.

HUBER, P.J. (1981). *Robust Statistics.* New York. John Wiley.

LANGE, K.L., LITTLE, R.J.A. and TAYLOR,J.M.G. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, **84**, 881-96.

LAWLEY, D.N. (1956). A generalized method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **43**, 295-303.

LITTLE, R.J.A. (1988). Robust estimation of the mean and covariance matrix from data with missing values. *Applied Statistics*, **37**, 23-39.

MARONNA, R.A. (1976). Robust M-estimators of multivariate location and scatter. *Annals of Statistics*, **4**, 51-67.

MITCHELL, A.N.S. (1989). The information matrix, skewness tensor and α-connections for the general multivariate elliptic distribution. *Annals of the Institute of Statistical Mathemathics* **41**, 289-304.

SINGH, R.S. (1991). James-Stein rule estimators in linear regression models with multivariate-t distributed error. *Australian Journal of Statistics*, **33**, 145-55.

SUTRADHAR, B.C. and ALI, M.M. (1986). Estimation of parameters of a regression model with a multivariate t error variable. *Communications in Statistics - Theory and Methods*, **15**, 429-450.

TAYLOR, J.M.E. (1992). Properties of modelling the error distribution with an extra shape parameter. *Computational Statistics and Data Analysis*, **13**, 33-46.

ULLAH, A. and ZINDE-WALSH, V. (1981). On the robustness of LM, LR and W tests in regression models. *Econometrica*, **52**, 1055-66.

ULLAH, A. and ZINDE-WALSH, V. (1985). Estimation and testing in regression models with spherically symmetric errors. *Economics Letters*, **17**, 127-32.

YU, L. and HOGG, R. (1988). On adaptive M-regression. *Biometrics*, **44**, 433-45.

ZELLNER, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error terms. *Journal of the American Statistical Association*, **71**, 400-405.

Table 1. Size and Power Simulations for Testing the Homogeneity of the Location Parameter in Two Independent Populations ($\nu = 1$) (Percentage of Rejections for 2,000 Replications)

n	Nominal sizes	Sizes(*)				Powers(**)	
		(1)	(2)	(3)	(4)	(5)	(6)
10	10.0	23.1	9.5	12.7	10.7	14.5	14.4
	5.0	13.3	4.7	6.3	5.5	7.8	7.5
	2.5	8.6	2.3	3.4	2.8	3.1	3.6
	1.0	4.7	1.2	1.4	1.3	0.6	1.4
20	10.0	14.6	9.9	11.2	9.8	18.5	18.7
	5.0	7.7	4.6	5.3	4.8	11.3	11.5
	2.5	4.3	2.2	2.5	2.2	6.6	6.9
	1.0	1.9	0.9	0.9	0.8	3.4	3.7
30	10.0	13.6	10.4	11.2	10.8	21.5	21.3
	5.0	7.6	5.6	5.6	5.3	12.6	13.9
	2.5	4.3	2.2	2.4	2.4	8.6	9.4
	1.0	1.5	0.8	1.0	0.9	4.1	4.2
40	10.0	11.6	9.3	10.2	9.9	30.3	29.4
	5.0	5.7	4.3	4.7	4.5	20.5	20.1
	2.5	2.9	2.2	2.1	2.0	12.4	13.6
	1.0	1.4	0.7	0.7	0.6	8.3	9.0

(*) (1), (2), (3) and (4) correspond to the simulated sizes of the tests based on LR, LR^*, S and S^* respectively.

(**) (5) and (6) correspond to the simulated powers of the tests based on LR (or LR^*) and S (or S^*) respectively (for $\beta_1 = 1.0$).

Table 2. Size and Power Simulations for Testing the Homogeneity of the Location Parameter in Two Independent Populations ($\nu = 3$)

(Percentage of Rejections for 2,000 Replications)

n	Nominal	Sizes(*)				Powers(**)	
	sizes	(1)	(2)	(3)	(4)	(5)	(6)
10	10.0	16.5	10.8	13.5	10.6	15.6	16.2
	5.0	9.9	5.0	6.9	5.3	9.7	9.5
	2.5	5.2	2.6	3.5	2.7	5.4	5.2
	1.0	2.7	1.1	1.4	1.3	2.3	2.1
20	10.0	12.6	10.5	11.4	10.6	22.1	21.2
	5.0	7.3	5.8	6.1	5.6	12.3	13.7
	2.5	4.2	3.0	3.0	2.9	7.7	7.3
	1.0	2.0	1.3	1.3	1.3	3.8	3.8
30	10.0	11.7	9.6	10.8	9.8	30.6	30.6
	5.0	5.9	4.7	5.4	4.9	20.7	20.4
	2.5	3.1	2.2	2.4	2.3	14.3	14.3
	1.0	1.1	0.7	0.8	0.8	8.3	8.7
40	10.0	12.0	10.6	11.4	10.7	36.4	36.9
	5.0	5.5	4.8	5.1	4.7	26.1	26.7
	2.5	2.9	2.5	2.7	2.6	17.1	17.0
	1.0	1.4	0.9	1.1	1.1	10.1	9.8

(*) (1), (2), (3) and (4) correspond to the simulated sizes of the tests based on LR, LR^*, S and S^* respectively.

(**) (5) and (6) correspond to the simulated powers of the tests based on LR (or LR^*) and S (or S^*) respectively (for $\beta_1 = 1.0$).

Table 3. Size and Power Simulations when the Covariate is the Quantiles of the U(0,1) Distribution and $\nu = 1$ (Percentage of Rejections for 2,000 Replications)

n	Nominal sizes	Sizes(*)				Powers(**)	
		(1)	(2)	(3)	(4)	(5)	(6)
10	10.0	23.4	7.3	12.2	9.7	10.3	12.1
	5.0	14.4	3.4	6.0	5.0	5.5	6.2
	2.5	8.4	1.4	2.7	2.4	2.8	3.0
	1.0	4.6	0.4	0.8	0.7	1.0	1.0
20	10.0	14.1	9.3	10.8	9.8	13.7	12.9
	5.0	8.1	4.7	5.5	5.0	7.3	6.5
	2.5	4.8	2.4	3.1	2.7	3.6	3.0
	1.0	2.4	0.8	0.9	0.9	1.5	1.8
30	10.0	13.1	10.5	10.9	10.3	14.4	14.2
	5.0	7.7	4.9	5.2	4.9	8.8	8.7
	2.5	3.9	2.2	2.5	2.3	5.1	5.3
	1.0	1.7	1.0	0.9	0.9	1.8	2.5
40	10.0	13.6	10.8	11.7	10.9	15.9	15.0
	5.0	7.1	5.1	5.9	5.5	9.0	9.0
	2.5	3.9	2.5	2.5	2.4	5.2	5.7
	1.0	1.7	1.1	1.3	1.3	2.4	2.0

(*) (1), (2), (3) and (4) correspond to the simulated sizes of the tests based on LR, LR^*, S and S^* respectively.

(**) (5) and (6) correspond to the simulated powers of the tests based on LR (or LR^*) and S (or S^*) respectively (for $\beta = 1.0$).

Table 4. Size and Power Simulations when the Covariate is the Quantiles of the U(0,1) Distribution and $\nu = 3$ (Percentage of Rejections for 2,000 Replications)

n	Nominal sizes	Sizes(*)				Powers(**)	
		(1)	(2)	(3)	(4)	(5)	(6)
10	10.0	16.1	9.0	11.7	9.4	14.0	14.6
	5.0	8.7	4.1	6.1	4.8	8.3	7.9
	2.5	4.8	2.1	2.6	2.1	4.4	4.8
	1.0	2.4	0.8	0.6	0.6	2.4	2.0
20	10.0	12.6	10.1	11.3	9.5	15.3	16.1
	5.0	6.5	4.7	5.1	4.6	8.8	8.7
	2.5	3.5	2.0	2.5	2.3	4.9	4.7
	1.0	1.4	0.8	0.9	0.9	2.0	2.2
30	10.0	12.6	10.6	11.5	10.8	14.3	13.9
	5.0	7.0	5.5	5.9	5.4	8.1	7.8
	2.5	3.8	2.9	2.7	2.5	4.1	4.6
	1.0	1.6	1.3	1.2	1.2	1.6	1.7
40	10.0	12.4	10.9	11.2	10.8	15.4	16.0
	5.0	7.0	6.3	6.2	5.7	8.8	9.1
	2.5	3.6	2.6	2.9	2.8	5.6	5.4
	1.0	1.5	1.3	1.4	1.4	2.1	2.2

(*) (1), (2), (3) and (4) correspond to the simulated sizes of the tests based on LR, LR^*, S and S^* respectively.

(**) (5) and (6) correspond to the simulated powers of the tests based on LR (or LR^*) and S (or S^*) respectively (for $\beta = 1.0$).

Table 5. Size and Power Simulations when the Covariate is the Quantiles of the LN(0,1) Distribution and $\nu = 1$ (Percentage of Rejections for 2,000 Replications)

n	Nominal	Sizes(*)				Powers(**)	
	sizes	(1)	(2)	(3)	(4)	(5)	(6)
10	10.0	26.4	3.1	11.5	8.8	29.2	23.4
	5.0	17.3	0.7	4.3	3.9	18.1	14.3
	2.5	11.6	0.3	1.5	1.5	11.4	6.9
	1.0	6.5	0.1	0.2	0.3	5.4	3.5
20	10.0	16.1	8.2	10.5	9.6	59.7	46.5
	5.0	9.3	3.5	5.1	4.6	44.6	31.6
	2.5	5.6	1.5	2.1	2.1	35.3	21.8
	1.0	2.4	0.4	0.7	1.0	21.1	10.9
30	10.0	14.1	9.5	10.9	10.0	75.3	59.7
	5.0	8.3	4.2	5.0	4.6	63.0	46.6
	2.5	4.2	2.2	1.9	1.9	51.5	37.4
	1.0	2.1	0.5	0.7	0.8	38.4	24.1
40	10.0	12.4	9.7	10.2	9.6	86.3	73.1
	5.0	6.8	4.7	4.7	4.7	77.7	60.7
	2.5	3.6	1.8	1.9	1.9	71.4	53.4
	1.0	1.4	0.7	0.5	0.6	59.8	58.4

(*) (1), (2), (3) and (4) correspond to the simulated sizes of the tests based on LR, LR^*, S and S^* respectively.

(**) (5) and (6) correspond to the simulated powers of the tests based on LR (or LR^*) and S (or S^*) respectively (for $\beta = 1.0$).

Table 6. Size and Power Simulations when the Covariate is the Quantiles of the LN(0,1) Distribution and $\nu = 3$ (Percentage of Rejections for 2,000 Replications)

n	Nominal sizes	Sizes(*)				Powers(**)	
		(1)	(2)	(3)	(4)	(5)	(6)
10	10.0	19.1	9.2	12.6	9.2	46.6	39.1
	5.0	11.3	5.1	4.8	4.0	29.8	25.1
	2.5	7.2	2.5	1.4	1.5	19.1	14.4
	1.0	4.0	0.8	0.1	0.1	11.8	9.1
20	10.0	12.6	9.5	11.5	9.9	78.6	73.3
	5.0	7.2	4.6	5.5	4.9	65.5	58.0
	2.5	4.3	2.5	2.6	2.6	53.6	42.0
	1.0	2.2	1.0	0.8	1.1	37.4	25.5
30	10.0	12.2	10.0	10.5	10.0	88.9	84.8
	5.0	6.3	4.4	4.8	4.6	82.5	76.1
	2.5	2.8	1.8	2.1	2.1	76.1	68.1
	1.0	1.2	0.7	0.8	0.9	65.5	52.6
40	10.0	11.7	10.0	10.6	10.3	96.3	93.4
	5.0	6.6	5.8	5.6	5.3	91.0	87.8
	2.5	3.9	3.3	2.9	2.9	83.9	78.0
	1.0	2.0	1.7	1.5	1.6	70.9	62.6

(*) (1), (2), (3) and (4) correspond to the simulated sizes of the tests based on LR, LR^*, S and S^* respectively.

(**) (5) and (6) correspond to the simulated powers of the tests based on LR (or LR^*) and S (or S^*) respectively (for $\beta = 1.0$).

ULTIMOS RELATORIOS TECNICOS PUBLICADOS

1993

9301 - BUENO, V.C. & ARIZONO, H. Comparisons for Maintenance Policies Involving Complete and Minimal Repair Through Compensator Transform, 12p.

9302 - BUENO, V.C. Maintenance Comparisons Through Compensator Transform: Block Policies, 13p.

9303 - FERRARI, P.A. & FONTES, L.R.G. Shock Fluctuations in the Asymmetric Simple Exclusion Process, 17p.

9304 - CORDEIRO, G. M., BOTTER, D.A. & FERRARI, S.L.P. Nonnull Asymptotic Distributions of Three Classic Criteria in Generalized Linear Models, 21p.

9305 - TANAKA, N.I. & TAVARES, H.R. Introdução à Teoria da Percolação, 38p.

9306 - BOLFARINE, H., ZACKS, S. & SANDOVAL, M.C. On Predicting the Population Total Under Regression Models with Measurement Erros, 10p.

9307 - FERRARI, P.A. & FONTES, L.R.G. The Net Output Process of a System with Infinitely Many Queues, 20p.

9308 - HO, L.L. & BUSSAB, W.O. Decomposição de Dados Categóricos com Auxílio de Árvores Dicotômicas, 23p.

9309 - BOLFARINE, H. & VALLE, R.B.A. Robust Modelling in Measurement Error Models Using the t Distribution, 14p.

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

- Departamento de Estatística
  IME-USP
  Caixa Postal 20.570
  01498-970 - São Paulo, Brasil