Brazilian Journal of Probability and Statistics (2003), 17, pp. 75–90.
©Associação Brasileira de Estatística

Performance of the Bartlett and Bartlett-type corrections in some location-scale nonlinear models

Maurício S. Lordêlo[1] and **Gauss M. Cordeiro**[2]
[1] *Universidade Salvador*
[2] *Universidade Federal Rural de Pernambuco*

Abstract: Statistical inference based on the normal model is known to be vulnerable to outliers. Despite this fact and the considerable interest in robust procedures in the statistical literature, most applied statistical analysis continues to be based on the normal model. Our approach is to replace the normal model by a general location-scale family of nonlinear models which include several asymmetric distributions that have a wide range of practical applications for analysing univariate data. We focus on the second-order corrections to the likelihood ratio and score statistics, since they are the most commonly used large sample tests. We obtain simple formulae for the corrections in some special location-scale models. We use Monte Carlo simulation to show that the corrected likelihood ratio and score tests have empirical sizes closer to the nominal sizes than the classical uncorrected tests even when the scale parameter in replaced by a consistent estimate.

Key words: Asymptotic distribution; Bartlett correction; Bartlett-type correction; likelihood ratio statistic; location-scale model; maximum likelihood estimate; score statistic.

1 Location-scale nonlinear models

The random variable Y is said to have a general location-scale family of distributions with location parameter $\theta \in I\!R$ and scale parameter $\phi > 0$ if its density function is of the form

$$\pi(y;\theta,\phi) = \frac{1}{\phi} h\left(\frac{y-\theta}{\phi}\right), y \in I\!R, \tag{1.1}$$

for some function h which is independent of y, θ and ϕ, such that $h(u) > 0$ and $\int_{I\!R} h(z)dz = 1$. In this paper, any particular location-scale distribution in (1.1) will be denoted by $LS(\theta, \phi^2)$. Different $h(\cdot)$'s correspond to different distributions. The location parameter θ is generally chosen as a measure of central tendency. The parameter ϕ can allow for some uncertainty concerning the dispersion of the distribution. The probability density function of $Z = (Y-\theta)/\phi$ is $\pi(x) = \pi(x;0,1) = h(x)$, $x \in I\!R$, which does not involve the parameters θ and ϕ, i.e., $Z \sim LS(0,1)$. It is the standardized form of the location-scale family of distributions.

The class of location-scale distributions defined in (1.1) has been considered by several authors. Some distributions in (1.1) include normal, Cauchy, type I and type II logistic, extreme value and Student's distributions. If the function $h(.)$ is symmetric about zero then the family (1.1) reduces to the symmetric class of location-scale densities which are mainly distinguished by their tail length. The family of location-scale densities (1.1) retains the structure of the normal distribution while eliminating the specific form of the normal density.

Let $Y_1, ..., Y_n$ be a set of n independent but not necessarily identically distributed random variables such that

$$Y_i \sim LS(\theta_i, \phi^2) \tag{1.2}$$

for $i = 1, 2, ..., n$, where the scale parameter ϕ is common to all n variables. All Y_i's are assumed to have the same distribution in (1.1). The location parameter θ_i is not necessarily equal to the mean of Y_i. We consider the nonlinear regression structure which links the location parameter θ_i to the covariates x_i of interest

$$\theta_i = f(x_i; \beta), \tag{1.3}$$

where $x_i = (x_{i1}, ..., x_{im})^T$ is an $m \times 1$ vector of nonstochastic independent variables associated with the ith response, $\beta = (\beta_1, ..., \beta_p)^T$ is a $p \times 1$ vector of unknown parameters such that $\beta \in \Omega_\beta \subset I\!R^p$. We further assume that Ω_β is an open set in $I\!R^p$ and $f(x_i; \beta)$ is an injective and twice continuously differentiable function with respect to β. Then, inference about β and ϕ can be performed by likelihood methods analogously to those for the normal nonlinear model.

The log likelihood function for the parameter vector β in the location-scale nonlinear regression model (1.2)-(1.3) given the sample $y_1, ..., y_n$ is

$$\ell(\beta) = -n \log \phi + \sum_{i=1}^{n} \upsilon(z_i), \tag{1.4}$$

where $\upsilon(z_i) = \log h(z_i)$ with $z_i = \phi^{-1}(y_i - \theta_i)$ and θ_i given in (1.3). The $n \times p$ matrix $\partial\theta/\partial\beta^T$ of derivatives of $\theta = (\theta_1, ..., \theta_n)^T$ with respect to β is denoted as $\tilde{X} = \tilde{X}(\beta)$, and is assumed to be of full rank, i.e., rank $(\tilde{X}) = p$ for all $\beta \in \Omega_\beta$. For nonlinear models $\tilde{X}$ is called the local design matrix and is a function of β, whereas a linear model is parameterized such that $\tilde{X}$ is constant. We also assume that the usual regularity conditions for maximum likelihood estimation and large sample inference hold; see Cox and Hinkley (1974, Chapter 9).

Let $U = U(\beta)$ and $K = K(\beta)$ denote the total score function and Fisher's total information, respectively. We can show that $U = -\phi^{-1}\tilde{X}^T\upsilon^{(1)}$ and $K = \phi^{-2}\upsilon\tilde{X}^T\tilde{X}$, where $\upsilon^{(1)} = (\partial\upsilon(z_1)/\partial z_1, ..., \partial\upsilon(z_n)/\partial z_n)^T$ is an $n \times 1$ vector and υ is a scalar given by $\upsilon = E\{(\partial\upsilon(z)/\partial z)^2\}$. Consider now a partition of the $p \times 1$ vector β as $\beta = (\beta_1^T, \beta_2^T)$, where $\beta_1 = (\beta_1, ..., \beta_q)^T$ and $\beta_2 = (\beta_{q+1}, ..., \beta_p)^T$ for $q \leq p$, thus inducing a partition of the local model matrix as $\tilde{X} = (\tilde{X}_1, \tilde{X}_2)$. Our interest is in testing the null hypothesis $H_0 : \beta_1 = \beta_1^{(0)}$ against a two-sided alternative hypothesis. Here, β_2 is a $(p - q)$-vector of nuisance parameters

and $\beta_1^{(0)}$ is a q-vector of known constants. Usually, $\beta_1^{(0)} = 0$. In more realistic regression models, it is important to find good approximations for likelihood ratio (LR) and score statistics. This makes Bartlett-type corrections (Cordeiro and Ferrari, 1991) very helpful, both in theory and in practice. Let $\hat{\beta} = (\hat{\beta}_1^T, \hat{\beta}_2^T)^T$ be the unrestricted MLE of β and $\tilde{\beta}_2$ the restricted MLE of β_2 under $\beta_1 = \beta_1^{(0)}$. Functions evaluated at $\tilde{\beta} = (\beta_1^{(0)^T}, \tilde{\beta}_2^T)^T$ are distinguished by the addition of a tilde except in $\tilde{X}(\beta) = \partial\theta/\partial\beta$. The MLE $\hat{\beta}$ can be obtained by the method of iteratively reweighted least-squares from the equations

$$\beta^{(m+1)} = \beta^{(m)} - \frac{\phi}{\upsilon}\left(\tilde{X}^{(m)^T}\tilde{X}^{(m)}\right)^{-1}\tilde{X}^{(m)^T}\upsilon^{(1)(m)}, \tag{1.5}$$

which are easily implemented in GLIM (Baker and Nelder, 1978) and S-PLUS (see Becker at et., 1988). The MLE $\tilde{\beta}$ follows the same algorithm with $\beta_1 = \beta_1^{(0)}$. Except for normal models, the MLE $\hat{\beta}$ does depend in general on the scale parameter ϕ.

The LR and score statistics for testing $H_0\colon \beta_1 = \beta_1^{(0)}$ are given by $w = 2\{\ell(\hat{\beta}) - \ell(\tilde{\beta})\}$ and $S_R = \tilde{U}^T\tilde{K}^{-1}\tilde{U}$, respectively. The score statistic becomes $S_R = \tilde{s}^T\tilde{X}_1(\tilde{R}^T\tilde{R})^{-1}\tilde{X}_1^T\tilde{s}$, where $s = (s_1, ..., s_n)^T$ and s_i is given by $s_i = \upsilon^{-1/2}\partial\upsilon(z_i)/\partial z_i$ with $R = \tilde{X}_1 - \tilde{X}_2(\tilde{X}_2^T\tilde{X}_2)^{-1}\tilde{X}_2^T\tilde{X}_1$. Here $\tilde{X}_1$ and $\tilde{X}_2$ are evaluated at $\tilde{\beta}$. It is well known that under H_0 and mild regularity conditions both w and S_R converge in distribution to χ_q^2 distribution. Bartlett (1953) notes that generally the χ_q^2 approximation to the distribution of w does not hold to $O(n^{-1})$. Finite sample corrected LR tests can be performed by using a Bartlett corrected statistic $w^* = w/\tilde{c}_B$, where $\tilde{c}_B$ is the restricted MLE of $c_B = E(w)/q$, so that under H_0, $P(w^* \leq x) = P(\chi_q^2 \leq x) + O(n^{-2})$. Then, the improved LR test follows from the Bartlett correction c_B. Finite sample corrections for score tests can be found in full generality in Cordeiro and Ferrari (1991). The main goal of this paper is to present simulation results comparing the sizes and powers of the usual LR and score tests and their Bartlett corrected versions when the scale parameter is considered known and when it is replaced by a consistent estimate. It is a very difficult problem to obtain improved tests assuming unknown scale parameter, since this parameter is not orthogonal to β.

The plan of this paper is as follows. In Section 2 we review Bartlett and Bartlett-type corrections. In Section 3, based on Cordeiro and Montenegro (2001) and Cysneiros and Cordeiro (2002) works, we consider two special location-scale models and give simple formulae for the Bartlett and Bartlett-type corrections. These formulae can then be used analytically to obtain Bartlett and Bartlett-type corrections in a variety of important models. Finally, we show in Section 4 that both adjustments to the LR and score statistics bring the sizes of the modified tests closer to the nominal asymptotic levels. We also show that the corrected statistics deliver more precise inference than their uncorrected counterparts even when the scale parameter ϕ is replaced by its MLE under the null hypothesis. We also show that the powers of the modified tests are not very sensitive to the estimation of ϕ.

2 Bartlett and Bartlett-type corrections

Generally speaking, the main difficulty of testing a null hypothesis in regression models using the LR criterion lies not so much in deriving its closed-form expression - when it has one - but in finding its exact distribution, or at least a good approximation for it, when the null hypothesis is true. The Bartlett correction became widely known in order to obtain corrected LR tests, i.e., modified test statistics with better finite-sample behavior, and many papers have focused on deriving simple Bartlett corrections for specific regression problems. A general formula for the Bartlett correction is presented by Lawley (1956). Cordeiro (1983, 1987) derived closed-form Bartlett corrections in generalized linear models (Nelder and Wedderburn, 1972) and discussed improved LR tests. Several papers have focused on deriving closed-form Bartlett corrections for specific regression problems. Cordeiro, Paula and Botter (1994) derived corrections for the class of dispersion models proposed by Jørgensen (1987); and Chesher and Smith (1995) obtained Bartlett corrections for LR specification tests. A correction to the LR statistics in regression models with Student-t errors was obtained by Ferrari and Arellano-Valle (1996), and similar corrections to heteroskedastic linear models and multivariate regression were obtained by Cribari-Neto and Ferrari (1995a) and Cribari-Neto and Zarkos (1995), respectively.

Score tests are commonly used in statistics and econometrics to test a wide range of hypotheses in a large number of models. Finite-sample corrections for score tests can be found from the results given by Cordeiro and Ferrari (1991). They showed that, in regular problems, the score statistics S_R can be improved by a Bartlett-type correction which is not exactly a Bartlett correction because it involves a polynomial of second degree in the original statistic. The corrected score statistics S_R is given by

$$S_R^* = S_R\{1 - (c + bS_R + aS_R^2)\}, \tag{2.1}$$

where a, b and c are terms of order n^{-1} that come from Harris's (1985) expansion for the distribution function of S_R under the null hypothesis. The modified statistic S_R^* has a chi-squared distribution to order n^{-1} under the null hypothesis. The corrected score tests have been developed and studied extensively in recent years, as detailed in the references cited in Cribari-Neto and Cordeiro (1996). For generalized linear models, these coefficients were obtained by Cordeiro, Ferrari and Paula (1993) and Cribari-Neto and Ferrari (1995b) for testing that a subset of the regression coefficients equals a given vector of constants, assuming that the dispersion parameter is known and unknown, respectively.

3 Corrections to some special location-scale models

The Bartlett correction for improving the test of H_0: $\beta_1 = \beta_1^{(0)}$ follows as $c_B = 1 + (\epsilon_p - \epsilon_{p-q})/q$ where the terms ϵ_p and ϵ_{p-q} of order n^{-1} are obtained from

equations (3.4)-(3.6) given in matrix notation by Cordeiro and Montenegro (2001). Then, the modified LR statistic $w^* = 2\tilde{c}_B^{-1}\{\ell(\hat{\beta}) - \ell(\tilde{\beta})\}$ has cumulants which agree with those of χ^2_q apart from terms of order $O(n^{-2})$. On the other hand, the Bartlett-type corrections (2.1) follow from the general expressions (3.1)-(3.3) also given in matrix notation by Cysneiros and Cordeiro (2002). When the coefficients a, b and c involve unknown parameters they should be replaced by their maximum likelihood estimates under H_0 but this does not affect the order of approximation of the Bartlett-type corrections.

Expressing the Bartlett and Bartlett-type corrections in matrix notation has great advantage to obtain closed-form expressions for several special models of practical use for which the matrix $\tilde{X}^T\tilde{X}$ has a closed-form inverse. When this is not the case, the Bartlett and Bartlett-type corrections can be easily obtained using a computer algebra system such as Mathematica or Maple, or using a programming language with support for matrix operations, such as GAUSS, Ox or S-PLUS.

We now give Bartlett and Bartlett-type corrections for two special location-scale models for which some simulations are presented in Section 4. We consider the one-parameter nonlinear model and a two-parameter linear regression model.

3.1 The one-parameter nonlinear model

Consider a one-parameter location-scale model with a general nonlinear structure for $\theta_i = f_i(\beta)$. Using the notation $f'_\ell = d\theta_\ell/d\beta$ and $f''_\ell = d^2\theta_\ell/d\beta^2$, we obtain $d_\ell = f''_\ell V$, $b_{\ell m} = f''_\ell f''_m V^2$, $z_{\ell m} = f'_\ell f'_m V$ and $c_{\ell m} = f'^2_\ell f''_m V^2$, where $V = \left(\sum_{\ell=1}^n f'^2_\ell\right)^{-1}$. Using equations (3.4)-(3.6) of Cordeiro and Montenegro (2001), the Bartlett correction term $c_B = 1 + L_1 + NL_1$ can be decomposed into the sum of two components L_1 and NL_1 given by

$$L_1 = \phi^2 V^2 \left\{ \frac{h}{4e^2} \sum_{\ell=1}^n f'^4_\ell + \frac{5f^2}{12e^3} V \left(\sum_{\ell=1}^n f'^3_\ell \right)^2 \right\}$$

and

$$\begin{aligned} NL_1 \;=\; & \phi^2 V^2 \left\{ \frac{1}{4e} \sum_{\ell=1}^n f''^2_\ell + \frac{3f}{2e^2} \sum_{\ell=1}^n f'^2_\ell f''_\ell - \frac{V}{4e} \left(\sum_{\ell=1}^n f'^2_\ell f''_\ell \right)^2 \right. \\ & \left. - \frac{3fV}{2e^2} \left(\sum_{\ell=1}^n f'^3_\ell \right) \left(\sum_{\ell=1}^n f'_\ell f''_\ell \right) \right\}, \end{aligned}$$

where the scalars e, f and h are defined here by $e = -G_{21} = G_{12}$, $f = \phi^{-1}G_{31}$ and $h = \phi^{-2}G_{41}$ with $G_{ij} = E\left\{\partial^i v(z) \partial z^{ij}\right\}$ for $i = 1, ..., 4$ and $j = 1, 2$ and clearly depend only on the distribution proposed for the observations.

For testing the simple null hypothesis H_0: $\beta = \beta^{(0)}$, $\beta^{(0)}$ being a specified value for β, the quantities L_1 and NL_1 should be evaluated at $\beta^{(0)}$. We now consider

that $\beta^{(0)} = 1$ which yields the LR statistic $w = 2\sum_{i=1}^{n}\left\{v\left(\frac{y_i - e^{\hat{\beta}x_i}}{\phi}\right) - v\left(\frac{y_i - e^{x_i}}{\phi}\right)\right\}$ by assuming $f_i(\beta) = e^{\beta x_i}$. Table 1 gives the LR statistics and the right-hand side of equation (4.2) for obtaining the restricted MLE $\tilde{\phi}$ of ϕ.

Table 1 *LR statistic for testing H_0: $\beta = 1$ and the right-hand side of equation (4.2) for calculating $\tilde{\phi}$.*

Model	w
normal	$\frac{1}{\phi^2}\sum_{\ell=1}^{n}\left\{2y_\ell\left(e^{\hat{\beta}x_\ell} - e^{x_\ell}\right) + e^{2x_\ell} - e^{2\hat{\beta}x_\ell}\right\}$
logistic	$2\sum_{\ell=1}^{n}\left\{\frac{1}{\phi}\left(e^{\hat{\beta}x_\ell} - e^{x_\ell}\right) - 2\log\left(1 + e^{-(y_\ell - e^{\hat{\beta}x_\ell})/\phi}\right)\right.$ $\left.+2\log\left(1 + e^{-(y_\ell - e^{x_\ell})/\phi}\right)\right\}$
extreme value	$2\sum_{\ell=1}^{n}\left\{\frac{1}{\phi}\left(e^{\hat{\beta}x_\ell} - e^{x_\ell}\right) - e^{-\frac{1}{\phi}y_\ell}\left(e^{\frac{1}{\phi}e^{\hat{\beta}x_\ell}} - e^{\frac{1}{\phi}e^{x_\ell}}\right)\right\}$
Model	$\tilde{\phi}$
normal	$\left\{\frac{1}{n}\sum_{\ell=1}^{n}(y_\ell - e^{x_\ell})^2\right\}^{1/2}$
logistic	$\frac{1}{n}\sum_{\ell=1}^{n}\frac{(e^{x_\ell} - y_\ell)\left(1 - e^{-(y_\ell - e^{x_\ell})/\tilde{\phi}}\right)}{\left(1 + e^{-(y_\ell - e^{x_\ell})/\tilde{\phi}}\right)}$
extreme value	$\frac{1}{n}\sum_{\ell=1}^{n}(e^{x_\ell} - y_\ell)\left(1 + e^{-(y_\ell - e^{x_\ell})/\tilde{\phi}}\right)$

3.2 Simple linear regression model

We consider a simple linear regression model for the location parameter θ_i given by $\theta_i = \alpha + \beta x_i$ for $i = 1, ..., n$. For testing the composite null hypothesis H_0: $\beta = 0$ against H: $\beta \neq 0$ we can calculate the Bartlett corrections from Cordeiro and Montenegro (2001). We define the scalars e, f, g, h and i by $e = -\delta_{010}$, $f = \phi^{-1}\delta_{001}$, $g = \delta_{200} = -\delta_{010}$, $h = \phi^{-2}(\delta_{400} - 3\delta_{010}^2)$ and $i = \phi^{-2}(\delta_{400} - 2\delta_{010}^2 + \delta_{210})$, where $\delta_{abc} = E\{v^{(1)a}v^{(2)b}v^{(3)c}\}$ for $a, b, c = 0, 1, 2, 3, 4$ and $v^{(r)} = \partial^r v(z)/\partial z^r$. The Bartlett correction can be written just as

$$c_B = 1 + \frac{\phi}{12ne^3}\{3eh(\gamma_{2x} + 5) + 5f^2(\gamma_{1x}^2 + 3)\}, \tag{3.1}$$

where $\gamma_{1x} = s_3/s_2^{3/2}$ and $\gamma_{2x} = (s_4 - 3s_2^2)/s_2^2$ are the usual measures of the sample skewness and kurtosis of x and $\bar{s}_a = n^{-1}\sum_{\ell=1}^{n}(x_\ell - \bar{x})^a$ is the a-th sample central moment of the covariate x for $a = 2, 3, 4$.

Now we can find $s^T = g^{-1/2}(v_1^{(1)}, ..., v_n^{(1)})$, $X_1X_1^T = \{x_ix_j\}$, $(R^TR)^{-1} = \left\{\sum_{\ell=1}^{n}(x_\ell - \bar{x})^2\right\}^{-1}$ and then the score statistic to test H_0: $\beta = 0$ reduces to

$$S_R = \frac{\sum_{i=1}^{n}\left\{v^{(1)}(\frac{y_i - \tilde{\alpha}}{\phi})x_i\right\}^2}{ng\bar{s}_2},$$

where $\tilde{\alpha}$ comes iteratively from $\sum_{i=1}^{n}\left\{v^{(1)}(\frac{y_i-\tilde{\alpha}}{\phi})\right\}=0$. For the normal model, the score statistic becomes

$$S_R=\frac{\{\sum_{i=1}^{n}(x_iy_i-n\bar{x}\bar{y})\}^2}{n\phi^2\bar{s}_2},$$

which is exactly distributed as χ_1^2.

We now give the expressions for a, b and c to improve the score test of H_0: $\beta=0$ for other distributions in (1.1). From equations (3.1)-(3.3) of Cysneiros and Cordeiro (2002), we can obtain

$$a=\frac{\phi^2f^2}{9ng^3}\frac{\bar{s}_3^2}{\bar{s}_2^3}, \tag{3.2}$$

$$b=\frac{1}{36ng^2}\left\{-4\left[\frac{3\phi^2f^2}{g}+\frac{10\bar{s}_3^2}{\bar{s}_2^3}\right]+3\phi^2f\frac{\bar{s}_4}{\bar{s}_2^2}\right\} \tag{3.3}$$

and

$$c=\frac{1}{12ng^3}\left[4\phi^2f^2(6+\frac{5\bar{s}_3^2}{\bar{s}_2^3})-6\phi^2gi-3\phi^2gf\frac{\bar{s}_4}{\bar{s}_2^2}\right]. \tag{3.4}$$

The score statistics for testing H_0: $\beta=0$ for some distributions in (1.1) are given in Table 2.

Table 2 *Score statistics for testing H_0: $\beta=0$.*

Model	S_R
Cauchy	$\left(\frac{n\bar{s}_2}{2}\right)^{-1}\left[\sum_{\ell=1}^{n}\frac{2\left(\frac{y_\ell-\bar{y}}{\phi}\right)^2x_\ell}{1+\left(\frac{y_\ell-\bar{y}}{\phi}\right)^2}\right]^2$
logistic	$\left(\frac{n\bar{s}_2}{3}\right)^{-1}\left[\sum_{\ell=1}^{n}\left(\frac{2e^{-(y_\ell-\bar{y})/\phi}}{1+e^{-(y_\ell-\bar{y})/\phi}}-1\right)\left(\frac{y_\ell-\bar{y}}{\phi}\right)x_\ell\right]^2$
extreme value	$(n\bar{s}_2)^{-1}\left[\sum_{\ell=1}^{n}\left(e^{-(y_\ell-\bar{y})/\phi}\right)\left(\frac{y_\ell-\bar{y}}{\phi}\right)x_\ell\right]^2$
Student t	$\left(\frac{n\bar{s}_2}{(\nu+1)(\nu+3)}\right)^{-1}\left[\sum_{\ell=1}^{n}\frac{\left(\frac{y_\ell-\bar{y}}{\phi}\right)^2x_\ell}{\nu+\left(\frac{y_\ell-\bar{y}}{\phi}\right)^2}\right]^2$

For the Student t model $e=g=(\nu+1)/(\nu+3)$, $f=0$, $h=-18(\nu+1)^2(\nu^2+4\nu-1)/\{\phi^2\nu(\nu+3)^2(\nu+5)(\nu+7)\}$ and $i=-12(\nu+1)^2(\nu^2+4\nu-1)/\{\phi^2\nu(\nu+3)^2(\nu+5)(\nu+7)\}$.

4 Simulation results

We use the following equations to obtain the estimates $\tilde{\beta}$ and $\tilde{\phi}$ simultaneously

$$\beta^{(m+1)} = \beta^{(m)} - \frac{\phi^{(m)}}{g}(X^{(m)T}X^{(m)})^{-1}X^{(m)T}\upsilon^{(1)^{(m)}} \tag{4.1}$$

and

$$\phi^{(m+1)} = \frac{1}{n}\sum_{i=1}^{n}\upsilon^{(1)}\left(\frac{y_i - \theta_i^{(m)}}{\theta^{(m)}}\right)(\theta_i^{(m)} - y_i), \tag{4.2}$$

where $g = E\left\{\left(\frac{\partial\upsilon(z)}{\partial z}\right)^2\right\}$. Each loop through the iterative scheme (4.1)-(4.2) consists of a see-saw algorithm to optimize the log likelihood (1.4) which is now interpreted based on a one-parameter nonlinear model (discussed in Section 3.1) where the response is assumed to have normal, logistic and extreme value distributions. The test to be considered is that of the null hypothesis H_0: $\beta = 1$ against the alternative hypothesis H: $\beta \neq 1$.

The values of the independent variable were chosen as random draws from a uniform distribution on the interval $(0, 1)$ and were kept constant throughout the experiment. The simulations were performed by setting $\beta^{(0)} = 1$ for three distributions (normal, logistic and extreme value) and the sample sizes $n = 10, 20, 30$ and 40. The scale parameter was taken as $\phi = 2$ and $\phi = 5$.

We generate logistic and extreme value distributions from $y = \exp(\beta^{(0)}x) + \phi\log\left(\frac{u}{1-u}\right)$ and $y = \exp(\beta^{(0)}x) - \phi\log(-\log u)$, where the random variable u has a uniform $U(0, 1)$ distribution. For each data set, we compute the MLEs β and ϕ using the interative scheme (4.1)-(4.2). We compare the performance in terms of sizes of the usual LR test w, its Bartlett corrected version $w^* = w(1 + L_1 + NL_1)^{-1}$ assuming the true value for the scale parameter ϕ and the corrected statistic $\tilde{w}^*$ calculated using the restricted MLE $\tilde{\phi}$ (see Table 3). The calculation of the rejection rates of the tests based on the statistics w, w^* and $\tilde{w}^*$, i.e. the percentage of times that these statistics exceed the appropriate upper points of the χ_1^2 distribution, were performed using FORTRAN 90 with special IMSL subroutines. We carried out size simulations for these statistics based on $10,000$ replications. The rejection rates corresponding to the nominal sizes $\alpha = 5\%$ and 10% based on the χ_1^2 distribution are given in Table 3 for the normal, logistic and extreme value distributions.

The figures in this table convey important information. First, the LR test rejects the null hypothesis too often, especially when n is small. In fact, for all twenty four cases reported, the rejection rates of w are larger than the corresponding nominal levels. Second, the empirical sizes of the test based on the corrected LR statistics w^* are closer to the nominal levels than the empirical sizes of the uncorrected statistics w. Thus, the corrected tests outperform the original LR test in all cases suggesting that the χ_1^2 distribution provides a better approximate to the distribution of w^* than to the distribution of w. It is clear that the Bartlett correction is effective in pushing the rejection rates of the corrected

statistic toward to the nominal levels, thus correctly signalizing the direction of such corrections. Third, the original LR test has a poor finite-sample size performance when n is small. Further, the asymptotic χ^2_1 approximation for w works better for small values of ϕ. Fourth, the χ^2_1 approximation for both statistics w^* and $\tilde{w}^*$ is more accurate for larger sample sizes, which is in agreement with the large sample asymptotic theory. Fifth, the Bartlett correction is still very effective when ϕ is replaced by its MLE $\tilde{\phi}$ since the estimated sizes of w^* and $\tilde{w}^*$ are in good agreement.

Table 3 *Rejection rates (in percentage) for the tests based on w, w^* and $\tilde{w}^*$.*

				normal			
n	α	$\phi = 2$			$\phi = 5$		
		w	w^*	$\tilde{w}^*$	w	w^*	$\tilde{w}^*$
10	5	8.3	6.5	6.7	9.1	6.7	7.2
	10	13.4	11.9	12.6	14.7	12.1	12.8
20	5	7.9	5.6	5.9	8.5	6.1	6.7
	10	12.5	11.4	11.8	13.8	11.2	11.7
30	5	6.5	5.3	5.5	7.3	5.4	5.9
	10	11.8	10.8	11.3	12.1	10.7	11.2
40	5	5.4	5.1	5.2	5.6	5.1	5.4
	10	10.6	10.3	10.5	11.4	10.4	10.7
				logistic			
n	α	$\phi = 2$			$\phi = 5$		
		w	w^*	$\tilde{w}^*$	w	w^*	$\tilde{w}^*$
10	5	9.2	6.6	7.3	9.8	6.3	7.5
	10	15.4	12.4	13.5	16.1	11.8	12.7
20	5	8.4	5.9	6.2	9.1	5.9	6.4
	10	13.7	11.8	12.4	14.4	11.4	11.9
30	5	7.0	5.4	5.9	7.7	5.6	6.1
	10	12.2	10.6	11.3	12.5	10.5	10.8
40	5	5.7	5.3	5.5	5.9	5.3	5.6
	10	10.7	10.4	10.5	11.6	10.2	10.5
				extreme value			
n	α	$\phi = 2$			$\phi = 5$		
		w	w^*	$\tilde{w}^*$	w	w^*	$\tilde{w}^*$
10	5	6.5	5.8	6.1	6.8	6.1	6.5
	10	13.7	12.3	12.9	15.2	12.1	13.5
20	5	6.1	5.5	5.9	6.3	5.7	6.1
	10	12.4	11.7	12.1	13.1	11.4	12.3
30	5	5.7	5.3	5.6	5.9	5.5	5.8
	10	11.5	10.8	11.0	11.9	10.8	11.4
40	5	5.4	5.1	5.3	5.3	5.2	5.3
	10	10.9	10.3	10.6	11.3	10.4	10.8

The second experiment considers a two-parameter linear regression location-scale model defined by (1.1) and by the systematic component $\theta_\ell = \alpha + \beta x_\ell$, where

x_ℓ denotes the values of an explanatory variable, for $\ell = 1, ..., n$. For testing the composite null hypothesis H_0: $\beta = 0$ of *iid* observations against the hypothesis of violation of at least one equality, we compare the performance of the usual LR and score tests and their corrected versions in terms of sizes and powers. The independent variable x was chosen as random draws from a uniform distribution $U(0, 1)$ and its values were held constant throughout the simulations with equal sample sizes. We generate *iid* t distributed observations with $\nu = 2$ and 5 degrees of freedom. The number of observation was set at $n = 10, 20, 30$ and 40. The nuisance parameter was fixed at $\alpha = 5$ and ϕ was taken as 1. The simulations run size was $10,000$ in each case. In each simulation, we generate *iid* t distributed observations and compute the MLEs $\hat{\theta}$, $\tilde{\theta}$ and $\tilde{\phi}$, the uncorrected statistics w and S_R and the corrected statistics w^*, $\tilde{w}^*$, S_R^* and $\tilde{S}_R^*$ using equations (3.1)-(3.4), the forms $\tilde{w}^*$ and $\tilde{S}_R^*$ being calculated at the restricted MLE $\tilde{\phi}$ of ϕ. We calculated the rejection rates of these statistics at the nominal 5% and 10% levels of the reference χ_1^2 distribution out of $10,000$ simulations.

The figures in Table 4 reveal that the LR test tends to be quite oversized at least for n small. The size performance of the corrected test is much better than that of its uncorrected counterpart even when the scale parameter is estimated under the null hypothesis. The effect of the Bartlett correction to the LR statistics is remarkable. For example, for the t model with 2 degrees of freedom, $n = 30$ and a 10% nominal size, the simulated size of the LR test drops from 19.4% (almost two times the nominal size) to 12.0% and 13.7% when the Bartlett correction is used with the true and estimated values of the scale parameter, respectively. The simulation results for score statistics reveal that the usual score test is oversized in all cases and has a poor χ_1^2 approximation when n is not large. For all 8 cases reported there is some evidence of the superiority of the χ_1^2 approximation for both corrected statistics S_R^* and $\tilde{S}_R^*$ over the uncorrected score test. In fact, the empirical sizes of the corrected statistics S_R^* and $\tilde{S}_R^*$ are closer to the corresponding nominal levels than the empirical sizes of the uncorrected statistics S_R. Moreover, the estimated sizes of S_R^* and $\tilde{S}_R^*$ are in good agreement for the great majority of cases although the size distortions of S_R^* are smaller than those of $\tilde{S}_R^*$. On this basis, which confirms the theoretical results of Cordeiro and Ferrari (1991), the distributions of S_R^* and $\tilde{S}_R^*$ are better described by χ_1^2 than the distribution of S_R.

The power simulations for the six tests were performed at $\alpha = 5\%$ for the sample size fixed at $n = 30$ assuming $\beta = 0.10$, 0.15, 0.20 and 0.25. We also carried out power simulations based on 10,000 replications. Table 5 displays the estimated powers of all tests. The figures in Table 5 show that the tests based on the modified statistics w^* and S_R^* have simulated powers greater than the corresponding ones of the unmodified statistics. It is also clear that the estimation of ϕ leads to a small reduction in the powers of both corrected tests. The reduction is larger for the modified score statistic. However, the powers of the tests based on w^* and $\tilde{w}^*$ are approximately the same. Therefore, the scale parameter ϕ can in practice be replaced by its MLE under the null hypothesis since the corrections for both LR and score statistics which use the estimated value $\tilde{\phi}$ offer at least some improvement in terms of sizes and powers.

Table 4 *Rejection rates (in percentage) of the tests based on w, w^*, $\tilde{w}^*$, S_R, S_R^* and $\tilde{S}_R^*$.*

n	α	*t model with* $\nu = 2$					
		w	w^*	$\tilde{w}^*$	S_R	S_R^*	$\tilde{S}_R^*$
10	5	12.7	7.1	9.6	8.3	6.4	7.1
	10	38.5	13.7	15.2	14.3	12.3	13.8
20	5	11.5	6.2	7.2	8.0	6.2	6.7
	10	27.3	12.9	14.1	13.8	12.1	12.7
30	5	9.3	5.8	6.5	7.7	5.5	5.9
	10	19.4	12.0	13.7	12.9	10.8	11.4
40	5	7.8	5.2	6.0	5.8	5.2	5.5
	10	15.6	11.4	12.8	11.5	10.6	10.8
n	α	*t model with* $\nu = 5$					
		w	w^*	$\tilde{w}^*$	S_R	S_R^*	$\tilde{S}_R^*$
10	5	11.8	6.8	8.3	8.1	6.2	6.8
	10	35.3	12.9	14.1	13.7	12.1	13.3
20	5	9.6	5.7	6.9	7.7	5.9	6.5
	10	21.2	11.7	13.7	12.9	11.8	12.4
30	5	8.7	5.9	6.3	6.8	5.3	5.7
	10	16.3	11.3	12.3	11.4	10.7	10.9
40	5	6.9	5.4	5.8	5.5	5.2	5.3
	10	13.7	10.8	11.9	11.3	10.4	10.6

Table 5 *Powers of the statistics w, w^*, $\tilde{w}^*$, S_R, S_R^* and $\tilde{S}_R^*$ for $n = 30$.*

	t model with $\nu = 2$					
β	w	w^*	$\tilde{w}^*$	S_R	S_R^*	$\tilde{S}_R^*$
0.10	25.7	34.2	32.7	30.2	42.3	36.4
0.15	50.3	63.9	60.2	43.7	65.1	56.8
0.20	83.9	93.8	89.5	85.2	96.0	85.3
0.25	91.7	98.2	94.3	93.1	98.7	92.1
	t model with $\nu = 5$					
β	w	w^*	$\tilde{w}^*$	S_R	S_R^*	$\tilde{S}_R^*$
0.10	26.2	35.8	31.4	31.3	40.2	33.7
0.15	52.8	67.2	63.2	45.8	69.3	53.2
0.20	85.9	94.3	91.3	78.5	95.2	87.9
0.25	93.2	97.5	95.2	90.2	96.4	91.8

The final experiment analyzes the small-sample performance of the statistics S_R, S_R^* and $\tilde{S}_R^*$ in a more complex model. We use a nonlinear regression model

$$Y_\ell = \sum_{i=1}^{p-2} \beta_i x_i + \gamma \exp(\delta u_\ell) + \phi \epsilon_\ell, \tag{4.3}$$

where the errors ϵ_ℓ were simulated from a t_ν Student distribution. This is the classical exponential model which appears often in physical and biochemical applications. The null hypothesis considered here is H_0: $\delta = -1$ against the alternative hypothesis H_1: $\delta \neq -1$. The number of nuisance parameters is then $p-1$. For $n = 20$, the covariates $x_1, ..., x_6$ were chosen as the quantiles of the following distributions: $U(0,1)$, $U(0,4)$, $N(0,1)$, $N(1,4)$, χ_2^2 and χ_3^2, respectively. We replicate their values for $n = 20, 30, 40$ and 50. The covariate u was chosen randomly from a $F(2,3)$ distribution. The values of all independent variables were held constant throughout the simulations with equal sample sizes. The response Y was generated assuming that the nuisance parameters are $\beta_1 = \beta_2 = 1$, $\beta_3 = \beta_4 = 2$, $\beta_5 = 3$, $\beta_6 = -2$ and $\gamma = 3$, and ϕ fixed at one.

In each simulation, we generate t Student data with ν degrees of freedom ($\nu = 1, 2, 4, \infty$), fit the exponential model and compute the score statistics S_R by

$$S_R = \tilde{s}^T \tilde{X}_1 (\tilde{R}^T \tilde{R})^{-1} \tilde{X}_1^T \tilde{s}, \tag{4.4}$$

where here $X_2 = (x_1, \ldots, x_{p-2})$ is a $n \times (p-2)$ matrix of covariates and $\tilde{X}_1 = -\gamma e^{-u}$ and $R = \tilde{X}_1 - \tilde{X}_2 (\tilde{X}_2^T \tilde{X}_2)^{-1} \tilde{X}_2^T \tilde{X}_1$ are $n \times 1$ vectors and $s = (s_1, ..., s_n)^T$ is an $n \times 1$ vector whose components are given by $s_i = (\nu+1) g^{-1/2} z_i / (\nu + z_i^2)$. The tildes in (4.4) indicate that functions are evaluated at the restricted MLEs obtained from (4.1) and (4.2). Then, we computed the adjusted statistics S_R^* and $\tilde{S}_R^*$. For the statisitc S_R^*, the Bartlett-type correction was calculated from the general formulae given in Section 4 by Cysneiros and Cordeiro (2002) with the β's being estimated under the null hypothesis. For the statistic $\tilde{S}_R^*$, we estimated under the null hypothesis β and ϕ simultaneously from equations (4.1) and (4.2). The simulations run-size was $10,000$ in each case and the calculations were performed using FORTRAN 90 with special IMSL subrontines. We report in Tables 6, 7 and 8 the rejection rates of the statistics S_R, S_R^* and $\tilde{S}_R^*$ at the nominal $\alpha = 5\%$ and 10% levels of the reference χ_1^2 distribution for $p = 4$, $p = 6$ and $p = 8$, respectively, the degrees of freedom ν being equal to $1, 2, 4$ and ∞ in all cases.

Table 6 *Rejection rates (in percentage) of S_R, S_R^* and $\tilde{S}_R^*$ for testing H_0: $\delta = -1$ with $p = 4$.*

n	α	$\nu = 1$			$\nu = 2$		
		S_R	S_R^*	$\tilde{S}_R^*$	S_R	S_R^*	$\tilde{S}_R^*$
20	5	8.3	6.1	7.0	7.9	5.8	6.7
	10	16.8	11.7	14.4	16.4	11.4	15.4
30	5	8.0	5.8	6.9	7.4	5.7	5.9
	10	15.2	11.3	13.8	14.5	11.0	13.1
40	5	7.5	5.8	6.3	6.7	5.6	5.8
	10	14.6	11.6	12.5	14.2	10.8	12.4
50	5	6.9	5.5	5.7	6.3	5.7	5.9
	10	13.8	10.9	11.8	13.1	11.0	11.4
n	α	$\nu = 4$			$\nu = \infty$		
		S_R	S_R^*	$\tilde{S}_R^*$	S_R	S_R^*	$\tilde{S}_R^*$
20	5	7.5	5.8	6.3	6.9	5.4	5.7
	10	14.8	10.9	13.7	12.5	10.5	12.8
30	5	7.0	5.6	6.7	6.3	5.4	5.9
	10	14.1	10.4	12.9	12.2	10.7	11.5
40	5	6.2	5.4	5.9	5.8	5.2	5.4
	10	13.5	10.6	11.8	11.8	10.4	10.8
50	5	5.9	5.3	5.7	5.5	5.1	5.3
	10	12.6	10.5	11.9	11.1	10.3	10.9

Table 7 *Rejection rates (in percentage) of S_R, S_R^* and $\tilde{S}_R^*$ for testing H_0: $\delta = -1$ with $p = 6$.*

n	α	$\nu = 1$			$\nu = 2$		
		S_R	S_R^*	$\tilde{S}_R^*$	S_R	S_R^*	$\tilde{S}_R^*$
20	5	10.1	6.7	7.3	9.3	5.9	6.8
	10	21.7	12.4	14.8	20.8	12.9	15.8
30	5	9.2	6.1	7.2	8.6	5.8	6.1
	10	17.8	12.8	14.0	17.4	12.4	13.7
40	5	8.5	5.9	6.2	8.1	5.4	6.1
	10	16.5	11.9	13.1	16.3	11.3	12.9
50	5	7.6	5.4	5.6	7.1	5.3	5.8
	10	15.5	11.3	12.7	15.1	11.8	12.5
n	α	$\nu = 4$			$\nu = \infty$		
		S_R	S_R^*	$\tilde{S}_R^*$	S_R	S_R^*	$\tilde{S}_R^*$
20	5	9.1	5.9	6.2	7.3	5.7	6.0
	10	17.5	12.1	14.1	14.7	11.8	13.1
30	5	8.3	6.1	6.4	7.1	5.3	5.9
	10	16.8	11.9	12.8	11.9	10.8	11.3
40	5	8.0	5.7	5.9	6.9	5.4	5.5
	10	16.1	11.4	12.3	12.7	11.1	10.7
50	5	7.0	5.8	6.0	5.8	5.1	5.4
	10	14.3	11.6	12.2	11.3	10.6	10.9

Table 8 *Rejection rates of (in percentage) S_R, S_R^* and $\tilde{S}_R^*$ for testing H_0: $\delta = -1$ with $p = 8$.*

n	α	$\nu = 1$			$\nu = 2$		
		S_R	S_R^*	$\tilde{S}_R^*$	S_R	S_R^*	$\tilde{S}_R^*$
20	5	11.7	6.5	7.2	10.2	6.3	7.0
	10	24.2	13.8	15.2	22.3	13.6	16.1
30	5	10.1	6.4	7.4	9.8	5.6	6.5
	10	21.8	13.3	14.1	21.0	13.1	14.8
40	5	9.4	5.9	6.3	8.4	5.7	6.1
	10	19.7	12.8	13.4	18.5	12.5	13.7
50	5	8.3	5.6	5.9	7.7	5.6	5.9
	10	17.5	12.1	12.9	16.7	11.9	12.6
n	α	$\nu = 4$			$\nu = \infty$		
		S_R	S_R^*	S_R^*	S_R	S_R^*	S_R^*
20	5	9.7	5.7	6.4	7.7	5.9	6.5
	10	20.6	13.1	14.6	15.3	12.7	13.2
30	5	9.4	5.8	6.1	7.4	5.5	5.9
	10	19.3	12.8	13.5	13.7	11.5	12.4
40	5	8.5	5.4	5.6	6.6	5.6	5.7
	10	17.6	12.3	12.9	12.7	11.1	11.9
50	5	7.4	5.7	6.2	6.2	5.3	5.5
	10	15.9	11.2	12.4	11.8	10.7	11.1

Simulation results shown in Tables 6-8 indicate that the uncorrected score test based on S_R is oversized in all cases especially when the number $p-1$ of nuisance parameters is large and n is small. For example, for $p = 8$ and $n = 20$, the rejection rates of S_R are more than two times the nominal sizes. The χ_1^2 approximation for the score statistic S_R improves when ν increases but such size reduction not happens always for the corrected statistics S_R^* and $\tilde{S}_R^*$. On the other hand, the χ_1^2 approximation for both corrected statistics S_R^* and $\tilde{S}_R^*$ works well even for large p. Clearly, the Bartlett-type corrections for S_R^* and $\tilde{S}_R^*$ become more necessary as p increases and serve to warn of situations in which the standard χ^2 approximation for the score statistic may be poor. The corrected statistic $\tilde{S}_R^*$ seems to perform roughly as well as the corrected statistic S_R^* and clearly both statistics have better performance than the unmodified statistic S_R in all 48 cases reported in these tables. Finally, this simulation experiment suggests that the first-order asymptotics usually employed with score tests can be misleading in small samples. More importantly, they show that analytical corrections based on second-order asymptotics can then be used to obtain improved tests with more reliable finite sample size behavior.

Acknowledgements

The authors wish to thank two referees and an Associate Editor for constructive comments and suggestions which have improved the paper.

(Received January, 2003. Accepted June, 2003.)

References

Baker, R. J. and Nelder, J. A. (1978). The GLIM Manual: Release 3. Oxford: Numerical Algorithms Group.

Bartlett, M. S. (1953). Approximate confidence intervals II. More than one unknown parameter. *Biometrika*, **40**, 306-317.

Becker, R. A., Chambers, J. M., Wilks, A. R. and Pagan, A. R. (1988). The New S Language: A Programming Environment for Data Analysis and Graphics. *Wadsworth and Brooks/Cole, Pacific Groove.*

Chesher, A. and Smith, R. J. (1995). Bartlett corrections to the likelihood ratio tests. *Biometrika*, **82**, 433-436.

Cordeiro, G. M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, **45**, 404-413.

Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265-274.

Cordeiro, G. M. and Ferrari, S. L. P. (1991). A modified score statistic having chi-squared distribution to order n^{-1}. *Biometrika*, **78**, 573-582.

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**, 661-674.

Cordeiro, G. M. and Montenegro, L. C. C. (2001). Bartlett corrected likelihood ratio tests in location-scale nonlinear models. *Communications in Statistics - Theory and Methods*, **30**, 1353-1372.

Cordeiro, G. M., Paula, G. A. and Botter, D. A. (1994). Improved likelihood ratio tests for dispersion models. *International Statistical Review*, **62**, 257-276.

Cox, D. R. and Hinkley D. V. (1974). *Theoretical Statistics.* New York: Chapman and Hall.

Cribari-Neto, F. and Cordeiro, G. M. (1996). On Bartlett and Bartlett-type corrections. *Econometric Reviews*, **15**, 339-367.

Cribari-Neto, F. and Ferrari, S. L. P. (1995a). Bartlett-corrected tests for heteroskedastic linear models. *Economics Letters*, **48**, 113-118.

Cribari-Neto, F. and Ferrari, S. L. P. (1995b). Second-order asymptotics for score tests in generalized linear models. *Biometrika*, **82**, 426-432.

Cribari-Neto, F. and Zarkos, S. (1995). Improved test statistics for multivariate regression. *Economics Letters*, **49**, 113-120.

Cysneiros, A. H. M. A. and Cordeiro, G. M. (2002). On improving the χ^2 approximation of score tests in location-scale nonlinear models. *Communications in Statistics - Theory and Methods*, **31**, 1709-1732.

Ferrari, S. L. P. and Arellano-Valle, R. B. (1996). Modified likelihood ratio and score tests in linear regression models using the t distribution. *Brazilian Journal of Probability and Statistics*, **10**, 15-33.

Harris, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653-659.

Jørgensen B.(1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society B*, **49**, 127-162.

Lawley, D. N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **71**, 233-244.

Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, **135**, 370-384.

Maurício S. Lordêlo
Universidade Salvador
Departamento de Ciências Exatas
Rua Dr. José Peroba, Edf. Civil Empresarial 251, Stiep
41770-235, Salvador, BA, Brazil.

Gauss M. Cordeiro
Departamento de Física e Matemática
Universidade Federal Rural de Pernambuco
Rua Dom Manoel de Medeiros s/n
52171-900, Recife, PE, Brazil.
E-mail: gauss@ufrpe.br