An Acad Bras Cienc (2022) 94(Suppl. 3): e20200568 DOI 10.1590/0001-3765202220200568
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

MATHEMATICAL SCIENCES

Bartlett and Bartlett-type corrections in heteroscedastic symmetric nonlinear regression models

MARIANA C. ARAÚJO, AUDREY H.M.A. CYSNEIROS & LOURDES C. MONTENEGRO

Abstract: This paper provides general expressions for Bartlett and Bartlett-type correction factors for the likelihood ratio and gradient statistics to test the dispersion parameter vector in heteroscedastic symmetric nonlinear models. This class of regression models is potentially useful to model data containing outlying observations. Furthermore, we develop Monte Carlo simulations to compare size and power of the proposed corrected tests to the original likelihood ratio, score, gradient tests, corrected score test, and bootstrap tests. Our simulation results favor the score and gradient corrected tests as well as the bootstrap tests. We also present an empirical application.

Key words: Bartlett corrections, Bartlett-type corrections, bootstrap, gradient test, large-sample test statistics.

INTRODUCTION

The symmetric class of models has received increasing attention in the literature. Besides the normal distribution, the symmetric family covers both light and heavy tailed distributions including the Cauchy, Student—*t*, generalized Student—*t* and power exponential, among others. The symmetric models provide a very useful extension of the normal model, since using a heavy tailed distribution for the error component reduces the influence of extreme observations and enables carrying out a more robust statistical analysis (Lange et al. 1989). An extensive range of practical applications considering symmetric distributions can be found in various fields, such as engineering, biology and economics, among others. Symmetric regression models have been the subject of several studies (e.g., Lin et al. 2009, Cysneiros et al. 2010, Lemonte 2012, Maior & Cysneiros 2018).

Constant dispersion is often a standard assumption when symmetric data are fitted. However, in many practical situations this condition is not satisfied, requiring verification, since the inference strategies change when one observes variable dispersion of the observations. The likelihood ratio (LR), Wald and score are the large-sample tests commonly used for this purpose. The gradient test proposed by Terrel (2002), whose statistic shares the same first-order asymptotic properties with the LR, Wald and score statistics (Lemonte & Ferrari 2012a), has been the subject of many studies in the past few years (e.g., Lemonte 2011, 2013, Lemonte & Ferrari 2012b, Medeiros & Ferrari 2017). Compared to the Wald and score statistics, the gradient statistic does not depend on the information matrix, either expected or observed, and is also simpler to compute.

The four statistics for testing hypothesis in regression models have a null asymptotic χ^2_q distribution, where q is the difference between the dimensions of the parameter space under the two hypotheses being tested, up to an error order of n^{-1}. Relying on inference in tests based on such statistics has less justification when dealing with small and moderately sized samples. A strategy to improve the χ^2 approximation for the exact distributions of the LR, score and gradient statistics is to multiply them by a correction factor. For the LR statistic, Bartlett (1937) proposed a correction factor known as the Bartlett correction, which was put into a general framework later by Lawley (1956), while for the score statistic, Cordeiro & Ferrari (1991) proposed a Bartlett-type correction. Based on the results of Cordeiro & Ferrari (1991), Bartlett-type correction for the Wald statistic was derived by Cribari-Neto & Ferrari (1995) in the heteroscedastic linear models. Furthermore, dos Santos & Cordeiro (1999) obtained corrected Wald test statistics for one-parameter exponential family models. Gradient statistic was proposed in a general framework by Vargas et al. (2013). The corrected versions of the test statistics have the same χ^2_q null distribution with approximation error of order n^{-2}. Cordeiro & Cribari-Neto (2014) provided additional details on Bartlett and Bartlett-type corrections. Improved tests have been discussed in some recent articles, in particular Lemonte et al. (2012), Bayer & Cribari-Neto (2013), Vargas et al. (2014) and Medeiros et al. (2017).

Considering the class of heteroscedastic symmetric nonlinear models (HSNLM) proposed by Cysneiros et al. (2010), Cysneiros (2011) derived a Bartlett-type correction for the score statistic, and carried out a numerical study to test the regression coefficients in the dispersion parameter. In this paper, our main goal is to derive Bartlett and Bartlett-type corrections to improve inference of the dispersion parameter based on the LR and gradient statistics, respectively, for the class of HSNLM considering the parameterization presented in Cysneiros et al. (2010). In other words, we deal only with one aspect of the high-order asymptotic theory which aims to obtain adjustments of test statistics. Furthermore, we consider a partition of the dispersion parameter, which is an advantage since in some cases we are not interested in making inferences of all parameters of the model. One of the main results presented in this paper, the Bartlett correction factor for the *LR* statistic, is not the same as presented in Araújo & Montenegro (2020), since the aforementioned dealt with two aspects of the high-order asymptotic theory: first we obtained the adjustment for the profiled likelihood (first aspect) and then we obtained the adjustment for the test statistic based on the profiled likelihood (second aspect).

In order to achieve our aim, we adopt a regression structure to model the dispersion parameter vector so that under the null hypothesis the dispersion is constant. In other words, the null hypothesis delivers the symmetric nonlinear regression model. Our results provide a new class of tests that can be used in practical applications, mainly those involving small datasets.

We perform a Monte Carlo simulation study to evaluate the performance of the proposed tests. For comparison purposes, besides the proposed tests and the usual score and gradient tests, we also consider in the Monte Carlo experiment the improved score test (Cysneiros 2011), the modified score tests proposed by Kakisawa (1996) and Cordeiro et al. (1998) and bootstrap-based tests. Our simulation results show that the improved gradient test proposed in this paper is an interesting alternative to the classic large-sample tests, delivering accurate inferences, mainly when dealing with small datasets. We are unaware of any simulation study in the literature drawing a comparison between the performance of the proposed tests in the class of models considered, so this paper fills this gap.

The remainder of this paper is organized as follows. In the next Section we present the class of HSNLM, explaining inferential aspects. After that, we derive Bartlett and Bartlett-type corrections to improve the LR and gradient tests for investigating varying dispersion in the model class of interest. We also conduct a Monte Carlo study to evaluate and compare the performance of the proposed tests. An application to real data is presented. Some concluding remarks are given in the last Section.

MODEL SPECIFICATION

Let y be a random variable with symmetric distribution. Its density function is given by

$$\pi(y;\mu,\phi) = \frac{1}{\sqrt{\phi}} g(u),\ y,\ \mu \in \mathbb{R},\ \phi > 0, \tag{1}$$

where μ is a location parameter, ϕ is a dispersion parameter, $u = (y-\mu)^2/\phi$, $g : \mathbb{R} \to [0,\infty)$ is the density generator (see, for example, Fang et al. (1990)). We then denote $y \sim S(\mu,\phi,g)$. Cysneiros et al. (2005) presented the density generator function $g(\cdot)$ for some symmetric distributions. In some symmetric distributions the density generator function, $g(\cdot)$, depends on an additional shape parameter, say ν, which controls the kurtosis. This parameter can be estimated from the data or can be kept fixed. Villegas et al. (2013) presented aspects of the symmetric distributions relating the issue of robustness to the estimation of the parameter ν. Thus, in this work we keep ν fixed.

Assume $y_1,\dots,y_n$ are independent random variables where each y_ℓ has a symmetric distribution (1) with location parameter μ_ℓ and dispersion parameter ϕ_ℓ. Also, consider that the components of $\boldsymbol{\mu} = (\mu_1,\dots,\mu_n)^\top$ and $\boldsymbol{\phi} = (\phi_1,\dots,\phi_n)^\top$ vary across observations through nonlinear regression structures. The heteroscedastic symmetric nonlinear regression model $y_\ell \sim S(\mu_\ell,\phi_\ell,g),\ \ \ell = 1,\dots,n$, proposed by Cysneiros et al. (2010) is defined by (1) and by the systematic components for the mean vector response $\boldsymbol{\mu}$ and the dispersion parameter vector $\boldsymbol{\phi}$ described as follows:

$$\mu_\ell = f(\boldsymbol{x_\ell};\boldsymbol{\beta}) \ \text{ and } \ \phi_\ell = h(\tau_\ell),$$

where $f(\cdot;\cdot)$ is a possibly nonlinear function in the second argument which is continuous and differentiable in $\boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_1,\dots,\beta_p)^\top$ ($p < n$ and $\boldsymbol{\beta} \in \mathbb{R}^p$) is a vector of unknown parameters to be estimated, and $\boldsymbol{x_\ell} = (x_{\ell 1},\dots,x_{\ell m})^\top$ is an $m \times 1$ vector of known explanatory variables associated with the ℓth observation. The matrix of derivatives of $\boldsymbol{\mu}$ with respect to $\boldsymbol{\beta}$, $\tilde{\boldsymbol{X}} = \partial\boldsymbol{\mu}/\partial\boldsymbol{\beta}$, is assumed to have full rank for all $\boldsymbol{\beta}$. Moreover, $h(\cdot)$ is a known continuous bijective function and differentiable in $\boldsymbol{\delta}$. Furthermore, $\tau_\ell = \boldsymbol{\omega_\ell}^\top\boldsymbol{\delta}$ is a linear predictor where $\boldsymbol{\omega_\ell} = (1,\omega_{\ell 1},\dots,\omega_{\ell k-1})^\top$ is a vector of explanatory variables whose components are not necessarily different from $\boldsymbol{x_\ell}$ and $\boldsymbol{\delta} = (\delta_0,\dots,\delta_{k-1})^\top$ ($\boldsymbol{\delta} \in \mathbb{R}^k$) is a vector of unknown parameters.

It is further assumed that if a value $\boldsymbol{\delta_0}$ of $\boldsymbol{\delta}$ exists, then $h(\boldsymbol{\omega_\ell}^\top\boldsymbol{\delta_0}) = 1$ for all ℓ, so the terms $y_\ell' s$ have constant dispersion if $\boldsymbol{\delta} = \boldsymbol{\delta_0}$. The function $h(\cdot)$ should be positively valued, and a possible choice is $h(\cdot) = \exp(\cdot)$, which is adopted in several papers (e.g., Cook & Weisberg 1983, Verbyla 1993, Simonoff & Tsai 1994, Barroso & Cordeiro 2005). Furthermore, considering $h(\tau_\ell) = \exp(\tau_\ell) = \exp(\boldsymbol{\omega_\ell}^\top\boldsymbol{\delta})$, it is not necessary impose any restriction on the components of $\boldsymbol{\omega}_\ell$ (Cook & Weisberg 1983, Lin et al. 2009). It is important to note that the meaning of heteroscedasticity we use in this work refers to varying dispersion, that is, when $\phi_1 = \phi_2 = \dots = \phi_n$ we have a homoscedastic model; without this we have a heteroscedastic model.

Let $l(\boldsymbol{\theta})$ denote the total log-likelihood function for the parameter of vector $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ given $y_1, \dots, y_n$. We have $l(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{\ell=1}^{n}\log(\phi_\ell) + \sum_{\ell=1}^{n} t(z_\ell)$, with $t(z_\ell) = \log g(z_\ell^2)$ and $z_\ell = \sqrt{u_\ell} = \frac{(y_\ell - \mu_\ell)}{\sqrt{\phi_\ell}}$. We assume that the function $l(\boldsymbol{\theta})$ is regular (Cox & Hinkley 1974, Chap 9) with respect to all $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ derivatives to the fourth order. The total Fisher information matrix for $\boldsymbol{\theta}$ has a block diagonal structure, i.e., $\boldsymbol{K_\theta} = \mathrm{diag}\{\boldsymbol{K_\beta}, \boldsymbol{K_\delta}\}$, where $\boldsymbol{K_\beta} = -\alpha_{2,0}\tilde{\boldsymbol{X}}\boldsymbol{\Lambda}^{-1}\tilde{\boldsymbol{X}}$ and $\boldsymbol{K_\delta} = \boldsymbol{W}^\top \boldsymbol{V}\boldsymbol{W}$, with $\boldsymbol{\Lambda} = \mathrm{diag}\{1/\phi_1, \dots, 1/\phi_n\}$, $\boldsymbol{W} = \boldsymbol{\partial\tau/\partial\delta}$ and $\boldsymbol{V} = \mathrm{diag}\{v_1, \dots, v_n\}$, such that $v_\ell = ((1-\alpha_{2,0}){h'}_\ell^2)/4\phi_\ell^2$, where $h' = \partial\phi_\ell/\partial\tau_\ell$ and $\alpha_{r,s} = E\{t(z_\ell)^{(r)}z_\ell^s\}$ for $r, s \in \{1,2,3,4\}$ and $t(z_\ell)^{(k)} = \partial^k t(z_\ell)/\partial z_\ell^k$, for $k = 1,2,3,4$ and $\ell = 1, \dots, n$. For some symmetric distributions, the quantities $\alpha_{r,s}$ are given in Uribe-Opazo et al. (2008). The parameters $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are globally orthogonal, so their respective maximum likelihood estimators (MLEs), $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\delta}}$, are asymptotically independent. In order to obtain the MLEs $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\delta}}$ iteratively, the scoring method can be applied. This procedure is described in detail in Cysneiros et al. (2010).

Our interest is to test heteroscedasticity in symmetric nonlinear regression models. The null and alternative hypotheses considered are, respectively, $H_0 : \boldsymbol{\delta_1} = \boldsymbol{\delta_1^{(0)}}$ and $H_1 : \boldsymbol{\delta_1} \neq \boldsymbol{\delta_1^{(0)}}$, where $\boldsymbol{\delta}$ is partitioned as $\boldsymbol{\delta} = (\delta_0, \boldsymbol{\delta_1}^\top)^\top$, with δ_0 a scalar and $\boldsymbol{\delta_1} = (\delta_1, \dots, \delta_{k-1})^\top$. Here, $\boldsymbol{\delta_1^{(0)}}$ is a fixed column vector of dimension $k-1$ such that $h(\boldsymbol{\omega_\ell^\top}\boldsymbol{\delta_1^{(0)}}) = 1$ and δ_0 and $\boldsymbol{\beta}$ are considered nuisance parameters. Actually, we are testing the dispersion parameters in the HSNLM, considering that under the null hypothesis this model boils down to the symmetric nonlinear regression model. The partition previously considered for $\boldsymbol{\delta}$ induces the corresponding partitions: $\boldsymbol{W} = (\boldsymbol{W_0}, \boldsymbol{W_1})$, where $\boldsymbol{W_0}$ is an $n \times 1$ vector with all ones and $\boldsymbol{W_1} = \partial\boldsymbol{\tau}/\partial\delta_1$,

$$\boldsymbol{K_\delta} = \begin{bmatrix} K_{\delta_0\delta_0} & \boldsymbol{K_{\delta_0\delta_1}} \\ & \\ \boldsymbol{K_{\delta_1\delta_0}} & \boldsymbol{K_{\delta_1\delta_1}} \end{bmatrix},$$

with $K_{\delta_0\delta_0} = \boldsymbol{W_0^\top V W_0}$, $\boldsymbol{K_{\delta_0\delta_1}^\top} = \boldsymbol{K_{\delta_1\delta_0}} = \boldsymbol{W_1^\top V W_0}$ e $\boldsymbol{K_{\delta_1\delta_1}} = \boldsymbol{W_1^\top V W_1}$. The likelihood ratio ($S_{LR}$), score ($S_r$) and gradient ($S_g$) statistics for testing H_0 can be expressed, respectively, as

$$\begin{aligned} S_{LR} &= 2\{l(\hat{\boldsymbol{\delta}}_1, \hat{\delta}_0, \hat{\boldsymbol{\beta}}) - l(\boldsymbol{\delta_1^{(0)}}, \tilde{\delta}_0, \tilde{\boldsymbol{\beta}})\}, \\ S_r &= \frac{1}{4}[\boldsymbol{W_1}\tilde{\boldsymbol{\Lambda}}(\tilde{\boldsymbol{S}}\tilde{\boldsymbol{F_1}}\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{F_1}}\boldsymbol{\iota})]^\top(\tilde{\boldsymbol{R}}^\top\tilde{\boldsymbol{V}}\tilde{\boldsymbol{R}})^{-1}[\boldsymbol{W_1}\tilde{\boldsymbol{\Lambda}}(\tilde{\boldsymbol{S}}\tilde{\boldsymbol{F_1}}\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{F_1}}\boldsymbol{\iota})] \text{ and} \\ S_g &= \frac{1}{2}[\boldsymbol{W_1}\tilde{\boldsymbol{\Lambda}}(\tilde{\boldsymbol{S}}\tilde{\boldsymbol{F_1}}\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{F_1}}\boldsymbol{\iota})]^\top(\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta_1^{(0)}}), \end{aligned}$$

where $(\hat{\boldsymbol{\beta}}, \hat{\delta_0}, \hat{\boldsymbol{\delta_1}})$ and $(\tilde{\boldsymbol{\beta}}, \tilde{\delta_0}, \boldsymbol{\delta_1^{(0)}})$ are, respectively, the unrestricted and restricted (under H_0) MLEs of $(\boldsymbol{\beta}, \delta_0, \boldsymbol{\delta_1})$, $\boldsymbol{\iota}$ is an $n \times 1$ vector of ones and $\boldsymbol{R} = \boldsymbol{W_1} - \boldsymbol{W_0 C}$, with $\boldsymbol{C} = (\boldsymbol{W_0}^\top \boldsymbol{V W_0})^{-1}(\boldsymbol{W_0}^{-1}\boldsymbol{V W_1})$. Under the null hypothesis, these statistics have an asymptotic χ_{k-1}^2 distribution up to an error of order n^{-1}.

IMPROVED TEST INFERENCE

In order to obtain a more accurate inference when dealing with small and moderately sized samples, some procedures based on second-order asymptotic theory have been developed in the literature. For the HSNLM, a Bartlett-type correction factor for the score statistic was derived by Cysneiros (2011). To provide another improved statistics to test varying dispersion in the HSNLM class, we derive Bartlett and Bartlett-type correction factors for the LR and gradient statistics, respectively, considering the general procedures developed by Lawley (1956) and Vargas et al. (2014). The Bartlett and Bartlett-type correction factors are very general and need to be obtained for every model of interest, since they involve complex functions of the moments of log-likelihood derivatives up to the fourth order. Details about the derivation of the Bartlett and Bartlett-type correction factors are given in the supplementary material.

To test $H_0 : \boldsymbol{\delta_1} = \boldsymbol{\delta_1^{(0)}}$ in HSNLM considering $h(\boldsymbol{\omega_l}^\top \boldsymbol{\delta}) = \exp(\boldsymbol{\omega_l}^\top \boldsymbol{\delta})$, i.e., the case of heteroscedasticity with multiplicative effects, the Bartlett-corrected LR statistic is given by

$$S_{LR^*} = \frac{S_{LR}}{1 + c/(k-1)},$$

where $c = \epsilon(\boldsymbol{\delta}) + \epsilon(\boldsymbol{\beta}, \boldsymbol{\delta}) - \epsilon(\delta_0) - \epsilon(\boldsymbol{\beta}, \delta_0)$,

$$\begin{aligned}
\epsilon(\boldsymbol{\delta}) &= N_1 tr\{\boldsymbol{Z_{\delta_d}}^{(2)}\} + N_2 \boldsymbol{\iota}^\top \boldsymbol{Z_\delta}^{(3)} \boldsymbol{\iota} + N_3 \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_\delta}^{(3)} \boldsymbol{\iota} + N_4 \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_\delta}^{(3)} \boldsymbol{\Lambda} \boldsymbol{\iota} \\
&+ N_5 \boldsymbol{\iota}^\top Z_{\boldsymbol{\delta_d}}{}^{(2)} \boldsymbol{Z_\delta} \boldsymbol{\iota} + N_6 \boldsymbol{\iota}^\top Z_{\boldsymbol{\delta_d}}{}^{(2)} \boldsymbol{Z_\delta} \boldsymbol{\Lambda} \boldsymbol{\iota} + (N_7 + N_8) \boldsymbol{\iota}^\top \boldsymbol{\Lambda} Z_{\boldsymbol{\delta_d}}{}^{(2)} \boldsymbol{Z_\delta} \boldsymbol{\iota},
\end{aligned}$$

$$\begin{aligned}
\epsilon(\boldsymbol{\beta}, \boldsymbol{\delta}) &= -N_{15} tr\{\boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} \boldsymbol{Z_{\delta_d}}\} - (N_{10} + N_{12}) \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} \boldsymbol{Z_\delta} \boldsymbol{Z_{\delta_d}} \boldsymbol{\iota} \\
&+ N_{14} \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} \boldsymbol{Z_\delta} \boldsymbol{Z_{\beta_d}} \boldsymbol{\Lambda} \boldsymbol{\iota} - (N_{11} + N_{13}) \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} \boldsymbol{Z_\delta} \boldsymbol{Z_{\delta_d}} \boldsymbol{\Lambda} \boldsymbol{\iota} \\
&+ N_9 \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_\delta} \boldsymbol{Z_\beta}^{(2)} \boldsymbol{\Lambda} \boldsymbol{\iota},
\end{aligned}$$

$$\begin{aligned}
\epsilon(\delta_0) &= N_1 tr\{\boldsymbol{Z_{\delta_{0d}}}^{(2)}\} + N_2 \boldsymbol{\iota}^\top \boldsymbol{Z_{\delta_0}}^{(3)} \boldsymbol{\iota} + N_3 \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\delta_0}}^{(3)} \boldsymbol{\iota} + N_4 \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\delta_0}}^{(3)} \boldsymbol{\Lambda} \boldsymbol{\iota} \\
&+ N_5 \boldsymbol{\iota}^\top Z_{\boldsymbol{\delta_{0d}}}{}^{(2)} Z_{\boldsymbol{\delta_0}} \boldsymbol{\iota} + N_6 \boldsymbol{\iota}^\top Z_{\boldsymbol{\delta_{0d}}}{}^{(2)} Z_{\boldsymbol{\delta_0}} \boldsymbol{\Lambda} \boldsymbol{\iota} \\
&+ (N_7 + N_8) \boldsymbol{\iota}^\top \boldsymbol{\Lambda} Z_{\boldsymbol{\delta_{0d}}}{}^{(2)} Z_{\boldsymbol{\delta_0}} \boldsymbol{\iota} \quad \text{and}
\end{aligned}$$

$$\begin{aligned}
\epsilon(\boldsymbol{\beta}, \delta_0) &= -N_{15} tr\{\boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} Z_{\boldsymbol{\delta_{0d}}}\} - (N_{10} + N_{12}) \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} Z_{\boldsymbol{\delta_0}} Z_{\boldsymbol{\delta_{0d}}} \boldsymbol{\iota} \\
&+ N_{14} \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} Z_{\boldsymbol{\delta_0}} \boldsymbol{Z_{\beta_d}} \boldsymbol{\Lambda} \boldsymbol{\iota} - (N_{11} + N_{13}) \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_{\beta_d}} Z_{\boldsymbol{\delta_0}} Z_{\boldsymbol{\delta_{0d}}} \boldsymbol{\Lambda} \boldsymbol{\iota} \\
&+ N_9 \boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_\delta} \boldsymbol{Z_\beta}^{(2)} \boldsymbol{\Lambda} \boldsymbol{\iota},
\end{aligned}$$

where $\boldsymbol{Z_\beta} = \tilde{\boldsymbol{X}}(\tilde{\boldsymbol{X}}^\top \boldsymbol{\Lambda} \tilde{\boldsymbol{X}})^{-1} \tilde{\boldsymbol{X}}^\top$, $\boldsymbol{Z_\delta} = \boldsymbol{W}(\boldsymbol{W}^\top \boldsymbol{V} \boldsymbol{W})^{-1} \boldsymbol{W}^\top$, $\boldsymbol{Z_{\delta_0}} = \boldsymbol{W_0}\ (\boldsymbol{W_0}^\top \boldsymbol{V} \boldsymbol{W_0})^{-1} \boldsymbol{W_0}^\top$, $\boldsymbol{Z_\beta}^{(2)} = \boldsymbol{Z_\beta} \odot \boldsymbol{Z_\beta}$, $\boldsymbol{Z_\delta}^{(2)} = \boldsymbol{Z_\delta} \odot \boldsymbol{Z_\delta}$, $\boldsymbol{Z_{\delta_0}}^{(2)} = \boldsymbol{Z_{\delta_0}} \odot \boldsymbol{Z_{\delta_0}}$, $\boldsymbol{Z_\delta}^{(3)} = \boldsymbol{Z_\delta}^{(2)} \odot \boldsymbol{Z_\delta}$, $\boldsymbol{Z_{\delta_0}}^{(3)} = \boldsymbol{Z_{\delta_0}}^{(2)} \odot \boldsymbol{Z_{\delta_0}}$, $\odot$ denotes the Hadamard (elementwise) product of matrices, and $(\cdot)_d$ indicates that the off-diagonal elements of the matrix are set equal to zero. For the sake of brevity, the elements $N_i,\ i = 1, \ldots, 15$ are presented in the supplementary material.

The improved gradient statistic is obtained by multiplying its original statistic by a polynomial in the original statistic itself. The corrected gradient statistic continues to have a chi-squared distribution under the null hypothesis but its asymptotic approximation error decreases from n^{-1} to n^{-2}, providing a more accurate inference. To test $H_0 : \boldsymbol{\delta_1} = \boldsymbol{\delta_1^{(0)}}$ in HSNLM when $h(\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta}) = \exp(\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta})$, the corrected gradient statistic is given by

$$S_{g^*} = S_g\{1 - (c_g + b_g S_g + a_g S_g^2)\},$$

where $a_g = \frac{A_3^g}{12(k-1)((k-1)+2)((k-1)+4)}$, $b_g = \frac{A_2^g - 2A_3^g}{12(k-1)((k-1)+2)}$, $c_g = \frac{A_1^g - A_2^g + A_3^g}{12(k-1)}$, with

$$\begin{aligned}
A_1^g &= 12\alpha_{2,0}Q_2\boldsymbol{\iota}^\top \boldsymbol{\Lambda} \boldsymbol{Z_\beta}^{(2)} \odot (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})\boldsymbol{\Lambda\iota} + 3Q_2^2\boldsymbol{\iota}^\top \boldsymbol{\Lambda Z_{\beta_d}}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\beta_d}}\boldsymbol{\Lambda\iota} \\
&+ 6Q_2^2\boldsymbol{\iota}^\top \boldsymbol{\Lambda}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}}) \odot \boldsymbol{Z_\beta}^{(2)}\boldsymbol{\Lambda\iota} + 3Q_1Q_2\boldsymbol{\iota}^\top \boldsymbol{\Lambda Z_{\beta_d}}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\delta_{0d}}}\boldsymbol{\iota} \\
&+ 3Q_1Q_2\boldsymbol{\iota}^\top Z_{\boldsymbol{\delta_{0d}}}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\beta_d}}\boldsymbol{\Lambda\iota} + 3Q_1Q_2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\beta_d}}\boldsymbol{\Lambda\iota} \\
&+ 6Q_1Q_2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d Z_{\boldsymbol{\delta_0}}Z_{\boldsymbol{\beta_d}}\boldsymbol{\Lambda\iota} + 3Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\delta_{0d}}}\boldsymbol{\iota} \\
&+ 6Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d Z_{\boldsymbol{\delta_0}}Z_{\boldsymbol{\delta_{0d}}}\boldsymbol{\iota} + 3Q_1^2\boldsymbol{\iota}^\top Z_{\boldsymbol{\delta_{0d}}}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\delta_{0d}}}\boldsymbol{\iota} \\
&+ 6Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}}) \odot \boldsymbol{Z_{\delta_0}}^{(2)}\boldsymbol{\iota} + 6Q_3 tr\{Z_{\boldsymbol{\delta_{0d}}}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d\} \\
&- 12Q_5 tr\{\boldsymbol{\Lambda}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d Z_{\boldsymbol{\beta_d}}\} + 6Q_4 tr\{\boldsymbol{\Lambda}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d Z_{\boldsymbol{\beta_d}}\},
\end{aligned}$$

$$\begin{aligned}
A_2^g &= -3Q_1Q_3\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\beta_d}}\boldsymbol{\Lambda\iota} \\
&- 3Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})Z_{\boldsymbol{\delta_{0d}}}\boldsymbol{\iota} \\
&- 3Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d Z_{\boldsymbol{\delta_0}}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d\boldsymbol{\iota} - 6Q_1^2\boldsymbol{\iota}(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})^{(2)} \odot \boldsymbol{Z_{\delta_0}}\boldsymbol{\iota} \\
&- \frac{9}{4}Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d\boldsymbol{\iota} - \frac{3}{2}Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})^{(3)}\boldsymbol{\iota} \\
&- 3Q_3 tr\{(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d^{(2)}\} \text{ and}
\end{aligned}$$

$$A_3^g = \frac{3}{4}Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d\boldsymbol{\iota} + \frac{1}{2}Q_1^2\boldsymbol{\iota}^\top (\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})^{(3)}\boldsymbol{\iota},$$

where $(\boldsymbol{Z_\delta} - \boldsymbol{Z_{\delta_0}})_d = \boldsymbol{Z_{\delta_d}} - \boldsymbol{Z_{\delta_{0d}}}$ and $Q_i, \;\; i = 1, \ldots, 5$ are given in the supplementary material.

The correction factors that improve the LR and gradient statistics are not easy to interpret, although they involve only simple matrix operations and can be easily implemented in any programming environment that performs linear algebra operations, such as MAPLE, Ox, R, etc. Also, they depend on the distribution in (1) only through the α's and also depend on the number of nuisance parameters, the dimension of the hypothesis tested and the matrices $\boldsymbol{X}$ and $\boldsymbol{W}$ of covariates. Finally, all unknown parameters in the correction factors are replaced by their restricted MLEs.

NUMERICAL EVIDENCE

The simulation experiments are based on the heteroscedastic symmetric nonlinear regression model

$$y_\ell = \beta_0 + \exp\{\beta_1 x_{\ell 1}\} + \sum_{s=2}^{p} \beta_s x_{s\ell} + \epsilon_\ell, \; \ell = 1, \ldots, n,$$

where $\epsilon_\ell \sim S(0, \exp\{\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta}\}, g)$. The response variable was generated assuming that $\beta_0 = \ldots = \beta_{p-1} = 1$, $\delta_0 = 0.1, \delta_2 = 0.3, \delta_3 = 0.5$ and $\delta_4 = \delta_5 = \delta_6 = 1$ and different values of p and k were considered. The covariates $x_1, \ldots, x_{p-1}$ and $\omega_1, \ldots, \omega_k$ were generated as random samples of the $U(0,1)$ distribution and were kept fixed throughout the simulations. The null hypothesis tested is $H_0 : \delta_1 = \ldots = \delta_{k-1} = 0$, i.e., $\exp\{\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta}\} = \exp\{\delta_0\}$, so that under H_0 we have constant dispersion. We report the null rejection rates of the tests based on the following statistics: the original likelihood ratio, score and gradient statistics ($S_{LR}, \; S_r, \; S_g$), their respective Bartlett and Bartlett-type corrected versions ($S_{LR^*}, \; S_{r^*}, \; S_{g^*}$) and the monotonic versions of the corrected score statistic proposed by Kakisawa (1996) and Cordeiro et al. (1998) ($S_{r_1^*}, \; S_{r_2^*}$), respectively. The bootstrap versions of the of the LR, score and gradient tests, with S_{LR}^{boot}, S_r^{boot} and S_g^{boot} being their respective test statistics, were also included. For the bootstrap-based tests, we followed the steps described as in Araújo & Montenegro (2020).

The simulation results are based on the Student-t (with $\nu = 5$) and power exponential (with $\nu = 0.3$) models. The following nominal levels and sample sizes were considered: $\alpha = 1\%, 5\%$ and 10%, and $n = 20,\ 30,$ and 40, respectively. All results were obtained using 10,000 Monte Carlo replications and 500 bootstrap samples. The bootstrap sampling was performed parametrically under the null hypothesis. The simulations were carried out using the Ox matrix programming language (Doornik 2006). All entries are percentages.

Tables I and II show results for different sample sizes while keeping fixed (varying) the number of nuisance (interest) parameters. The results clearly show that the LR test is notably liberal (i.e., it over-rejects the null hypotheses), especially when the numbers of interest parameters and nuisance parameters increase (the results with varying number of nuisance parameters are not shown to save space). It also can be noted that the gradient test behaves similarly to the LR test, but is less size distorted, while the usual score test performs much better than the other two uncorrected ones, although it is a bit liberal in a few cases. Considering $\alpha = 1\%$ and $n = 30$ for the Student-t model (see Table I), the null rejection rates of the LR test are 3.2% ($k = 3$), 5.0% ($k = 4$) and 6.5% ($k = 5$), while for the gradient test they are 2.4% ($k = 3$), 3.7% ($k = 4$) and 5.3% ($k = 5$), and for the score test they are 0.8% ($k = 3$), 1.0% ($k = 4$) and 1.0% ($k = 5$).

The simulation results also showed that the corrected tests based on the S_{LR^*}, S_{r^*} and S_{g^*} statistics outperformed their uncorrected versions, regardless of the sample size and the number of interest or nuisance parameters. Additionally, as shown in Tables I and II, the corrected versions of the LR and gradient tests are very sensitive to increasing the number of parameters in the model, whether they are interest or nuisance parameters. Otherwise, the corrected score test is not influenced by the increase in the number of parameters in the model and among the improved tests, the one based on the S_{r^*} statistic presents the best performance, exhibiting null rejection rates very close to the nominal level in most cases. For example, considering the power exponential model (Table II), if $k = 3, n = 20$ and $\alpha = 10\%$, the null rejection rate of the tests based on S_{LR^*}, S_{g^*} and S_{r^*} are, respectively, 16.7%, 11.3% and 10.3%, while considering the same scenario with $k = 4$, the null rejection rates for the tests based on S_{LR^*}, S_{g^*} and S_{r^*} are, respectively, 21.0%, 15.7% and 9.6%. Now considering the tests based on the monotonic versions of the corrected score statistics $S_{r_1^*}$ and $S_{r_2^*}$ proposed by Kakisawa (1996) and Cordeiro et al. (1998), the simulation results show that the performance of the tests based on those statistics are very similar to the corrected score test, presenting the same null rejection rate in most cases.

Tables I and II also show that the bootstrap-based tests are less size distorted than the corresponding uncorrected tests. Also, for the LR and gradient tests, their bootstrap versions outperform the corrected ones. On the other hand, the bootstrap score test behaves, in general, similarly to the monotonic and non-monotonic corrected ones. For small samples ($n = 20$), the bootstrap score test presents null rejection rates closer to the nominal level than the monotonic and non-monotonic corrected versions. For example, considering $p = 3,\ k = 3,\ n = 20$ and $\alpha = 5\%$, the null rejection rates are 17.7% (S_{LR}), 8.2% (S_{LR^*}), 4.8% ($S_{LR}^{boot},\ S_r^{boot}$), 5.7% ($S_r$), 5.5% ($S_{r^*},\ S_{r_1^*},\ S_{r_2^*}$), 15.2% ($S_g$), 5.6% ($S_{g^*}$) and 5.2% ($S_g^{boot}$) for the Student-$t$ model and 16.3% (S_{LR}), 9.6% (S_{LR^*}), 4.8% ($S_{LR}^{boot},\ S_r$), 5.4% ($S_{r^*},\ S_{r_1^*},\ S_{r_2^*}$), 14.6% ($S_g$), 6.2% ($S_g^*$), 4.9% ($S_r^{boot}$) and 5.1% ($S_g^{boot}$) for the power exponential model.

Table I. Null rejection rates $(\%)$ **for** $H_0 : \delta_1 = \ldots = \delta_k = 0$ **with** $p = 3$; t_5 **model.**

n	*Stat*	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
		k			k			k		
		3	4	5	3	4	5	3	4	5
20	S_{LR}	26.7	37.1	41.4	17.7	26.6	30.2	6.5	11.8	14.7
	S_{LR^*}	14.7	18.1	18.3	8.2	10.7	10.7	2.0	3.1	3.0
	S_r	11.0	11.3	12.2	5.7	5.7	6.1	0.9	1.1	1.0
	S_{r^*}	10.0	10.3	11.2	5.5	5.2	5.8	1.1	1.0	1.0
	$S_{r_1^*}$	10.1	10.3	11.2	5.5	5.2	5.8	1.1	1.0	1.0
	$S_{r_2^*}$	10.1	10.3	11.2	5.5	5.2	5.8	1.1	1.0	1.0
	S_g	24.1	33.1	36.5	15.2	22.6	26.1	5.3	9.2	11.4
	S_{g^*}	10.6	16.9	18.7	5.6	10.2	11.5	1.6	3.2	4.3
	S_{LR}^{boot}	10.1	10.0	9.8	4.8	4.8	5.0	0.9	1.1	1.1
	S_r^{boot}	9.5	10.1	9.9	4.8	4.7	4.8	1.1	1.0	0.9
	S_g^{boot}	10.1	10.3	9.6	5.2	4.9	4.6	1.0	1.1	0.9
30	S_{LR}	18.4	23.0	26.0	11.0	14.1	17.3	3.2	5.0	6.5
	S_{LR^*}	12.0	12.7	13.6	6.4	6.7	7.8	1.3	1.7	2.0
	S_r	10.4	10.2	11.0	5.0	5.2	5.5	0.8	1.0	1.0
	S_{r^*}	9.8	9.9	10.2	4.8	5.1	5.1	0.8	1.1	1.0
	$S_{r_1^*}$	9.8	9.9	10.2	4.8	5.1	5.1	0.8	1.1	1.0
	$S_{r_2^*}$	9.8	9.9	10.2	4.8	5.1	5.1	0.8	1.1	1.0
	S_g	17.1	20.1	24.0	9.7	11.9	15.5	2.4	3.7	5.3
	S_{g^*}	10.5	11.1	14.6	5.4	6.0	8.3	1.0	1.5	2.6
	S_{LR}^{boot}	10.4	10.0	10.6	5.4	4.8	5.5	1.4	0.9	1.3
	S_r^{boot}	10.3	10.4	9.9	5.7	5.3	5.0	1.2	0.9	0.9
	S_g^{boot}	10.2	10.2	9.8	5.4	5.0	5.2	1.1	0.9	1.0
40	S_{LR}	15.8	17.6	19.1	9.9	10.7	11.1	2.6	3.1	3.2
	S_{LR^*}	11.7	11.2	10.9	6.1	6.0	5.2	1.2	1.3	1.2
	S_r	10.8	10.5	10.6	5.5	5.2	5.1	0.9	0.9	0.9
	S_{r^*}	10.2	10.0	9.9	5.3	5.1	4.8	0.9	0.9	0.9
	$S_{r_1^*}$	10.2	10.1	10.0	5.3	5.1	4.8	0.9	0.9	0.9
	$S_{r_2^*}$	10.2	10.0	9.9	5.3	5.1	4.8	0.9	0.9	0.9
	S_g	15.2	16.6	17.8	8.8	9.9	10.1	2.1	2.5	2.5
	S_{g^*}	9.9	10.8	10.6	4.9	5.5	5.4	1.1	1.1	1.2
	S_{LR}^{boot}	9.9	10.2	9.6	4.7	5.1	4.8	1.0	1.0	1.1
	S_r^{boot}	9.8	9.9	10.1	5.1	4.9	5.2	0.9	1.0	1.0
	S_g^{boot}	9.9	9.9	9.7	5.0	4.9	4.8	0.9	0.9	0.9

Table II. Null rejection rates $(\%)$ **for** $H_0 : \delta_1 = \ldots = \delta_k = 0$ **with** $p = 3$; **power exponential** $\nu = 0.3$ **model.**

n	Stat	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
		k			k			k		
		3	4	5	3	4	5	3	4	5
20	S_{LR}	25.2	34.0	35.9	16.3	23.7	25.3	6.2	9.8	10.7
	S_{LR^*}	16.7	21.0	20.0	9.6	12.5	12.0	2.9	3.5	3.5
	S_r	9.7	9.9	10.3	4.8	5.2	5.7	0.9	1.7	1.6
	S_{r^*}	10.3	9.6	10.0	5.4	3.6	4.5	1.2	1.1	0.6
	$S_{r_1^*}$	10.3	10.1	10.1	5.4	3.6	4.5	1.2	1.1	0.6
	$S_{r_2^*}$	10.3	10.0	10.1	5.4	4.3	4.7	1.2	1.1	0.6
	S_g	23.0	30.1	31.5	14.6	20.4	21.4	5.1	7.6	8.7
	S_{g^*}	11.3	15.7	15.4	6.2	8.5	9.0	1.3	2.2	2.7
	S_{LR}^{boot}	10.1	10.4	10.1	4.8	5.3	5.4	1.0	1.1	1.3
	S_r^{boot}	9.8	10.4	10.1	4.9	5.0	5.2	1.0	0.9	1.3
	S_g^{boot}	10.4	10.1	10.0	5.1	5.1	5.2	0.9	0.9	1.0
30	S_{LR}	18.6	21.3	22.8	11.1	13.1	14.2	3.4	4.2	4.5
	S_{LR^*}	13.9	14.3	14.3	7.7	8.0	7.5	1.9	2.0	1.6
	S_r	9.9	10.3	10.1	5.2	5.3	5.4	1.4	1.4	1.3
	S_{r^*}	10.2	10.2	10.0	5.5	4.5	4.8	1.4	1.2	0.7
	$S_{r_1^*}$	10.2	10.3	10.0	5.5	4.7	4.8	1.4	1.2	0.7
	$S_{r_2^*}$	10.2	10.3	10.0	5.5	4.7	4.8	1.4	1.3	0.7
	S_g	18, 0	20.2	21.5	10.5	12.2	13.0	3.2	3.5	3.7
	S_{g^*}	11.4	12.1	12.4	6.0	6.4	6.2	1.2	1.4	1.4
	S_{LR}^{boot}	10.2	10.7	10.5	5.3	5.4	5.5	1.3	1.0	1.2
	S_r^{boot}	10.3	10.3	9.6	5.2	5.1	5.0	0.9	1.1	1.1
	S_g^{boot}	10.3	10.3	9.7	5.2	5.2	4.9	0.9	1.0	0.9
40	S_{LR}	14.4	16.4	18.3	8.2	9.3	10.9	2.3	2.9	3.4
	S_{LR^*}	12.6	11.7	12.7	6.6	6.4	7.3	1.8	1.7	1.8
	S_r	10.4	9.3	10.1	5.4	4.9	5.3	1.2	1.1	1.4
	S_{r^*}	10.2	9.4	10.7	4.8	4.8	5.7	0.5	0.8	1.5
	$S_{r_1^*}$	10.3	9.4	10.7	4.9	4.8	5.7	0.8	0.9	1.5
	$S_{r_2^*}$	10.3	9.4	10.7	4.9	4.8	5.7	0.8	0.9	1.5
	S_g	13.9	15.5	17.3	7.8	8.7	10.1	2.1	2.5	2.8
	S_{g^*}	9.9	10.5	11.7	5.0	5.6	6.4	1.2	1.5	1.5
	S_{LR}^{boot}	9.7	10.5	9.7	5.5	5.0	4.9	1.1	0.9	1.2
	S_r^{boot}	10.2	10.1	10.2	5.0	4.8	5.3	1.1	0.8	1.2
	S_g^{boot}	10.2	9.9	10.1	5.1	4.8	5.3	1.1	0.9	1.2

Finally, we also point out that all corrected and uncorrected tests present null rejection rates very close to the corresponding nominal level as the sample size increases, as expected.

Completing our simulations, we also obtained rejection rates under the alternative hypothesis (heteroscedasticity) for $n = 30$, $p = 3$, different values of $\delta_1 = \delta_2 = \delta_3 = \delta$ and at the 10% nominal level. It is noteworthy that these power simulations correspond to the setting in Tables I and II for $n = 30$. We did not include the likelihood ratio test, corrected likelihood ratio test, and gradient test in the power comparison since they are too liberal to be recommended. All other tests studied in this paper were considered. It should be noted in Figure 1 that the powers of the tests are similar, although there is an ordering, being the bootstrap-based tests and the corrected gradient test the ones which were slightly more powerful. As expected, the power tends to one when $|\delta|$ grows.

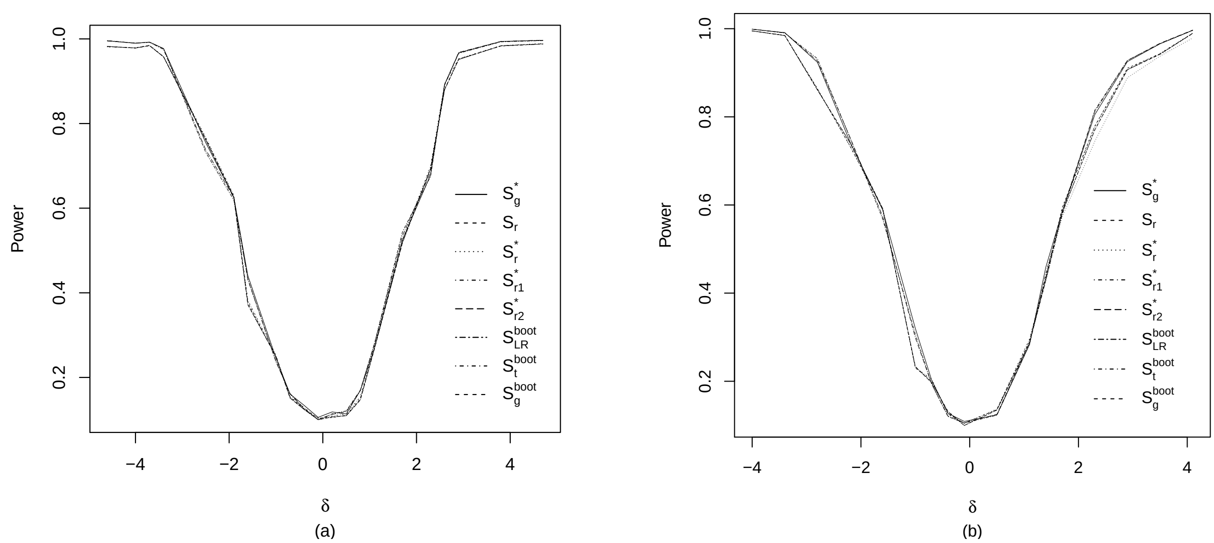

Figure 1. Power functions of the tests for $H_0 : \delta_1 = \ldots = \delta_3 = \delta$ with $p = 3$, $n = 30$, $\alpha = 10\%$; (a) t_5 and (b) power exponential $\nu = 0.3$ models.

In summary, the simulation results presented in this section show that the *LR* and gradient tests are considerably oversized (liberal) and the analytical Bartlett and Bartlett-type corrections for these tests are effective in reducing the size distortion. Also, their bootstrap versions outperformed the corrected ones, presenting results closer to the nominal levels considered. The score test is the best performing uncorrected test. Its (monotonic or not) corrected versions perform the same, as well its bootstrap-based test which, in some cases, performed better than the corrected versions.

REAL DATA APPLICATION

In this section, we consider a dataset analysed by Rawlings (1988) and Brazzale (2000). The dataset comes from a study executed by the Botany Department of North Carolina State University, where an experiment was carried out with the aim of performing biochemical analysis of intracellular storage and calcium transport across the plasma membrane. For this purpose, the cells studied were suspended in a solution of radioactive calcium for a certain period of time x (in minutes), after which was measured the amount of calcium absorbed by the cells y (in nmol/mg). The studied sample contains 27 observations and the model used is given by

$$y_\ell = \beta_0\{1 - \exp(-\beta_1 x_\ell)\} + u_\ell, \quad \ell = 1, \ldots, 27, \tag{2}$$

where $Cov(u_\ell, u_m) = 0$ for all $\ell \neq m$. Initially, we assume that $u_\ell \sim N(0, \exp\{\delta_0\})$. The maximum likelihood parameter estimates and their asymptotic standard errors in brackets are $\widehat{\beta}_0 = 4.3094\,(0.2901)$ $\widehat{\beta}_1 = 0.2084\,(0.0376)$ and $\widehat{\delta}_0 = -1.286\,(0.2722)$.

Residual analysis of the fitted homoscedastic normal model suggest an evidence of heteroscedasticity. In addition, we detect the presence of two observations #21 and #22 with large residuals (in absolute value) outside range [-2,2] (see Figure 2a). Due to the evidence of heteroscedasticity, we assume that $u_\ell \sim N(0, \exp\{\delta_0 + \delta_1 x_\ell\})$, named here, heteroscedastic normal model.

Our main interest lies in testing $H_0 : \delta_1 = 0$ (homoscedasticity) against $H_1 : \delta_1 \neq 0$. Rejection of the null hypothesis would suggest that the nonconstant response variance should be modeled as well. Table III presents the observed values for the test statistics and their respective $p-$values. Considering the 10% nominal level, the score, the standard likelihood ratio tests (and their corrected versions) and the gradient test, lead to rejecting H_0, suggesting the presence of heteroscedasticity. While the tests based on S_g^*, S_{LR}^{boot}, S_r^{boot} and S_g^{boot} statistics may not reject the null hypothesis at the same nominal level.

The studentized residual plot versus fitted values for the fitted heteroscedasticity normal model (see Figure 2b) still presents a slight indication of heteroscedasticity and residuals in the limit of the range $[-2, 2]$. Furthermore, we may suspect the presence of aberrant points, which we propose a symmetrical model with heavy tails as Student-t model. Lange et al. (1989) suggest that the degrees of freedom should be fixed for small-sized samples and indicate that $\nu = 4$ has worked well for some applications.

Figure 2c from the fitted homoscedastic Student-t model does not highlight aberrant points but still suggests little evidence of heteroscedasticity. Finally, Figure 2d from the heteroscedastic Student-t model does not present evidence of heteroscedasticity and does not highlight aberrant points. It follows that all proposed tests lead to the same decision of homoscedasticity at the nominal level $\alpha = 10\%$ (see Table III). Therefore, based on the parsimony principle we suggest the Student-t model.

CONCLUSIONS

In this paper we derive Bartlett and Bartlett-type corrections to improve hypothesis testing of the dispersion parameters for the HSNLM class of proposed by Cysneiros et al. (2010) and compare in a simulation study the performance of the proposed tests with the score test, its Bartlett-type corrected version and the uncorrected LR and gradient tests. We also consider for the simulation study monotonic versions of the Bartlett-type corrected score test and bootstrapped tests.

The numerical evidence suggests that the usual LR and gradient tests have similar performance, being oversized, mainly if the sample size is small or even moderate. The Bartlett and Bartlett-type

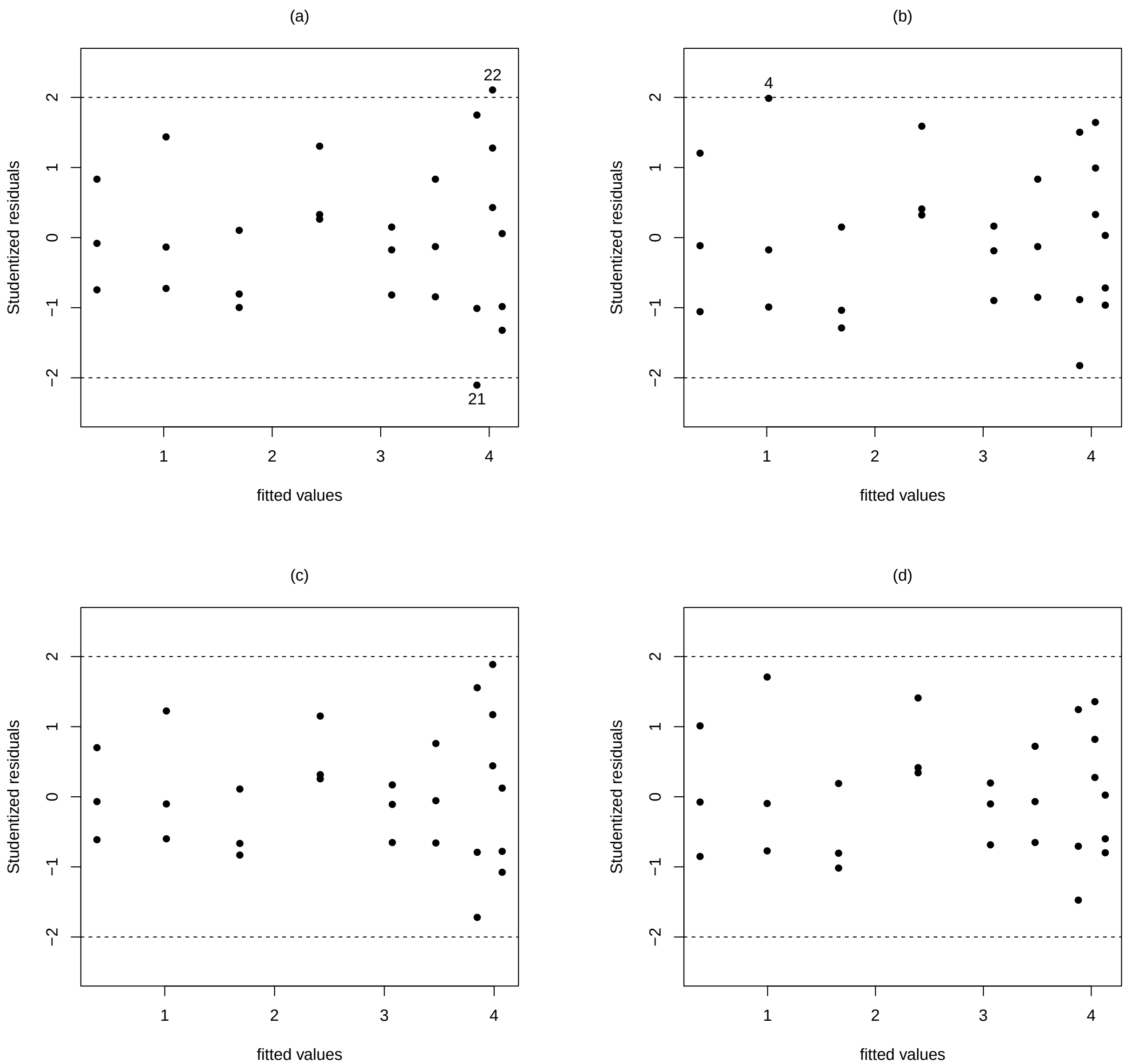

Figure 2. Calcium data. Plots of the studentized residual versus fitted values (a) homoscedastic normal model, (b) heteroscedastic normal model (c) homoscedastic Student-t model, (d) heteroscedastic Student-t with $\nu = 4$ model.

corrections attenuate this tendency, although the Bartlett corrected LR test still presents very liberal behavior while the Bartlett-type corrected gradient test produces results comparable to those of the usual and (monotonic or not) Bartlett-type corrected score tests. Additionally, the corrected score test and the bootstrapped tests perform the best overall. An advantage of the analytically corrected tests in relation to the bootstrapped tests is that they do not demand as much computational burden. Moreover, it is important to note that the corrected tests deliver more reliable inferences than their uncorrected versions when dealing with small or even moderate sized samples. We hence encourage practitioners to use the Bartlett-type corrected score and gradient or bootstrapped tests.

Table III. Test statistics and $p-$values for testing $H_0 : \delta_1 = 0$ in normal model and t_4 model.

	Normal model			t_4 model	
Stat	Observed value	$p-$value		Observed value	$p-$value
S_{LR}	3.097	0.078		2.283	0.131
S_{LR^*}	2.908	0.088		2.174	0.140
S_r	3.029	0.082		2.634	0.105
S_r^*	3.456	0.063		2.515	0.113
$S_{r_1}^*$	3.491	0.063		2.518	0.113
$S_{r_2}^*$	3.473	0.062		2.517	0.113
S_g	3.030	0.082		2.277	0.131
S_g^*	2.710	0.100		2.038	0.153
S_{LR}^{boot}		0.108			0.183
S_r^{boot}		1.000			0.945
S_g^{boot}		0.398			0.276
S_g^{boot}		0.398			0.276

Acknowledgments

We are thankful for the financial support of Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and FACEPE. The research of Lourdes C. Montenegro was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Grant 6796/14-1).

REFERENCES

ARAÚJO MC & MONTENEGRO AHMACLC. 2020. Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models. Stat Pap 61: 167-188.

BARROSO LP & CORDEIRO GM. 2005. Bartlett corretions in heteroskedastic *t* regression models. Stat Probabil Lett 75: 86-96.

BARTLETT MS. 1937. Properties of suficiency and statistical tests. P R Soc London 160: 268-282.

BAYER FM & CRIBARI-NETO F. 2013. Bartlett corrections in beta regression models. J Stat Plan Infer 143: 531-547.

BRAZZALE AR. 2000. Practical Small-Sample Parametric Inference. Ph.D. thesis. Unpublished PhD Dissertation. Department of Mathematics, Swiss Federal Institute of Technology Lausanne.

COOK D & WEISBERG S. 1983. Diagnostics for heteroscedasticity diagnostics in regression. Biometrika 70: 1-10.

CORDEIRO GM & CRIBARI-NETO F. 2014. An Introduction to Bartlett Correction and Bias Reduction. New York: Springer.

CORDEIRO GM & FERRARI SLP. 1991. A modified score test statistic having chi-squared distribution to order n^{-1}. Biometrika 78: 573-582.

CORDEIRO GM, FERRARI SLP & CYSNEIROS AHMA. 1998. A formula to improve score test statistics. J Stat Comput Sim 62: 123-136.

COX DR & HINKLEY DV. 1974. Theoretical Statistics. London: Chapman and Hall.

CRIBARI-NETO F & FERRARI S. 1995. Bartlett-corrected tests for heteroskedastic linear models. Economics Letters 48: 113-118.

CYSNEIROS AHMA. 2011. Bartlett-type Correction in Heteroscedastic Symmetric Nonlinear Models. In: WSM 2011- Proceedings of the 26th International Workshop on Statistical Modelling. Vol. 1. p. 156-159. Valencia, Spain.

CYSNEIROS FJA, PAULA GA & GALEA M. 2005. Modelos Simetricos Aplicados. Sao Paulo: ABE - XI Escola de Modelos de Regressao.

CYSNEIROS FJA, CORDEIRO GM & CYSNEIROS AHMA. 2010. Corrected maximum likelihood estimators in heteroscedastic symmetric nonlinear models. J Stat Comput Sim 80: 451-461.

DOORNIK JA. 2006. Ox: An Object-Oriented Matrix Programming Language 4^{th} ed. London: Timberlake Consultants Ltd.

DOS SANTOS SJP & CORDEIRO GM. 1999. Corrected Wald test statistics for one parameter exponential family models. Commun Statist - Theory Meth 28: 1391-1414.

FANG KT, KOTZ S & NG KW. 1990. Symmetric Multivariate and Related Distributions. London: Chapman and Hall.

KAKISAWA Y. 1996. Higher order monotone Bartlett-type adjustment for some multivariate test statistics. Biometrika 71: 233-244.

LANGE KL, LITTLE RJA & TAYLOR JMG. 1989. Robust statistical modeling using the *t* distribution. J Am Stat Assoc 84: 881-896.

LAWLEY DN. 1956. A general method for approximating to the distribution of the likelihood ratio criteria. Biometrika 43: 295-303.

LEMONTE AJ. 2011. Local power of some asymptotic tests in expontial nonlinear regression models. J Stat Plan Infer 141: 1981-1989.

LEMONTE AJ. 2012. Local power properties of some asymptotic tests in symmetric linear regression models. J Stat Plan Infer 142: 1178-1188.

LEMONTE AJ. 2013. Nonnull asymptotic distributions of the LR, Wald, score and gradient statistics in generalized linear models with dispersion covariates. Statistics 47: 1249-1265.

LEMONTE AJ & FERRARI SLP. 2012a. The local power of gradient test. Ann I Stat Math 64: 373-381.

LEMONTE AJ & FERRARI SLP. 2012b. Local power and size properties of the LR, Wald, score and gradient tests in dispersion models. Stat Methodol 9: 537-554.

LEMONTE AJ, CORDEIRO GM & MORENO G. 2012. Bartlett corretions in Birnbaum-Saunders nonlinear regression models. J Stat Comput Sim 82: 927-935.

LIN JG, ZHU LX & XIE FG. 2009. Heteroscedasticity diagnostics for *t* linear regression models. Metrika 70: 59-77.

MAIOR VQS & CYSNEIROS FJA. 2018. SYMARMA: a new dynamic model for temporal data on conditional symmetric distribution. Stat Pap 59: 75-97.

MEDEIROS FMC & FERRARI SLP. 2017. Small-sample testing inference in symmetric and log-symmetric linear regression models. Stat Neerl 71: 200-224.

MEDEIROS FMC, FERRARI SLP & LEMONTE AJ. 2017. Improved inference in dispersion models. Appl Math Model 51: 317-328.

RAWLINGS JO. 1988. Applied regression analysis. Wadsworth and Brooks/Probability series.

SIMONOFF JS & TSAI CH. 1994. Use of modified profile likelihood for improved tests of constancy of variance in regression. Appl Stat-J Roy St C 43: 357-370.

TERREL GR. 2002. The gradient statistic. Comp Sci Stat 34: 206-215.

URIBE-OPAZO MA, FERRARI SLP & CORDEIRO GM. 2008. Improved score test in symmetric linear regression model. Commun Stat A-Theor 37: 261-276.

VARGAS TM, FERRARI SLP & LEMONTE AJ. 2013. Gradient statistic: higher order asymptotics and Bartlett-type corrections. Electron J Stat 7: 43-61.

VARGAS TM, FERRARI SLP & LEMONTE AJ. 2014. Improved likelihood inference in generalized linear models. Comput Stat Data An 74: 110-124.

VERBYLA AP. 1993. Modelling variance heterogeneity: residual maximum likelihood and diagnostics. J Roy Stat Soc B Met 55: 509-521.

VILLEGAS C, PAULA GA, CYSNEIROS FJA & GALEA M. 2013. Influence diagnostics in generalized symmetric linear models. Comput Stat Data An 24: 161-170.

How to cite

ARAÚJO MC, CYSNEIROS AHMA & MONTENEGRO LC. 2022. Bartlett and Bartlett-type corrections in heteroscedastic symmetric nonlinear regression models. An Acad Bras Cienc 94: e20200568. DOI 10.1590/0001-3765202220200568.

Manuscript received on April 16, 2020;
accepted for publication on April 12, 2021

MARIANA C. ARAÚJO[1]
https://orcid.org/0000-0002-5682-520X

AUDREY H.M.A. CYSNEIROS[2]
https://orcid.org/0000-0002-8920-4004

LOURDES C. MONTENEGRO
https://orcid.org/0000-0003-0867-3931

[1]Universidade Federal do Rio Grande do Norte, Departamento de Estatística, Lagoa Nova, Av. Senador Salgado Filho, 3000, 59078-970 Natal, RN, Brazil
[2]Universidade Federal de Pernambuco, Departamento de Estatística, Cidade Universitária, Av. Prof. Moraes Rego, 1235, 50740-540 Recife, PE, Brazil
[3]Universidade Federal de Minas Gerais, Departamento de Estatística, Pampulha, Av. Pres. Antônio Carlos, 6627, 31270-901 Belo Horizonte, MG, Brazil

Correspondence to: **Mariana C. Araújo**

E-mail: mariana.araujo@ufrn.br

Author contributions

The three authors jointly developed the present paper. The first author derived all analytical results and performed the Monte Carlo simulations. The second author introduced the idea, checked, refined the notations and edited the improved test inference section. The third author carried out the empirical application, and wrote the abstract and the introduction section. All authors worked together to refine the manuscript in consultation with each other.