CrossMark

Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models

Mariana C. Araújo[1] · **Audrey H. M. A. Cysneiros**[2] ·
Lourdes C. Montenegro[3]

Abstract In this paper we address the issue of testing inference of the dispersion parameter in heteroscedastic symmetric nonlinear regression models considering small samples. We derive Bartlett corrections to improve the likelihood ratio as well modified profile likelihood ratio tests. Our results extend some of those obtained in Cordeiro (J Stat Comput Simul 74:609–620, 2004) and Ferrari et al. (J Stat Plan Inference 124:423–437, 2004), who consider a symmetric nonlinear regression model and normal linear regression model, respectively. We also present the bootstrap and bootstrap Bartlett corrected likelihood ratio tests. Monte Carlo simulations are carried out to compare the finite sample performances of the three corrected tests and their uncorrected versions. The numerical evidence shows that the corrected modified profile likelihood ratio test, the bootstrap and bootstrap Bartlett corrected likelihood ratio test perform better than the other ones. We also present an empirical application.

Keywords Bartlett correction · Bootstrap · Likelihood ratio test · Modified profile likelihood ratio test

✉ Mariana C. Araújo
mariana@ccet.ufrn.br

[1] Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Lagoa Nova, Natal, RN 59078-970, Brazil

[2] Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, Recife, PE 50740-40, Brazil

[3] Departamento de Estatística, Universidade Federal de Minas Gerais, Av. Antonio Carlos 6627, Pampulha, Belo Horizonte, MG 31270-010, Brazil

1 Introduction

The normal model, although very attractive, is not always appropriate to fit a dataset, especially if the data present extreme or outlying observations. Due to this problem, new regression models which are not so easily affected by extreme or outlying observations have been developed in the statistical literature. The symmetric family of distributions supplies an extension of the normal distribution, including other ones with heavier and lighter tails than the normal, such as Cauchy, Student-*t*, generalized Student-*t*, logistic I and II, generalized logistic, power exponential, Kotz distribution and generalized Kotz distribution, among others. This family of distributions provides a vast source of alternative models for analyzing data containing outlying observations. These models have been widely studied in the statistical literature. There are several recent articles considering the symmetric distribution; see Cordeiro et al. (2000), Cordeiro (2004), Cysneiros et al. (2010a), Cysneiros et al. (2010b), Vanegas et al. (2013), Maior and Cysneiros (2016). Further details about the symmetric family of distributions can be seen in Fang et al. (1990) and Fang and Anderson (1990).

When the dispersions are not constant over the observations, the inference strategies for regression parameters are different. Thus, it is extremely important to test whether variable dispersion is present in the data. One way to do this is to model the dispersion parameter as a function which depends on regressors and an unknown parameter vector in such a way that for a specific value of the parameter vector, the function corresponds to varying dispersion. Following this approach, one can formulate a hypothesis test in which the null hypothesis leads to constant dispersion. To that end, the likelihood ratio test is commonly used. Under the null hypothesis, the likelihood ratio statistic (LR) is asymptotically chi-square (χ^2) distributed up to an error of order n^{-1}, where n is the sample size. However, it is well known that for small samples this test can provide very distorted rejection rates. This happens due to the approximation of the null distribution of the LR statistic to the χ^2 distribution, which is not accurate. One way to improve the approach of the LR statistic distribution to the χ^2 distribution, and consequently reduce its distortion, is to multiply the LR statistic by a Bartlett correction factor (Bartlett 1937). This method was generalized later by Lawley (1956). The resulting statistic has a χ_k^2 null distribution up to an error of order n^{-2}, where k is the difference between the dimensions of the parameter space under the two hypothesis being tested.

Another factor that must be considered is the presence of nuisance parameters, which can have a profound impact on inference. Many approaches have been proposed to eliminate or reduce their impact. In the presence of nuisance parameters, generally it is feasible to perform inference based on profile likelihood, which is a function where the nuisance parameters are replaced by consistent estimates in the likelihood function, resulting in a function that depends only on the parameters of interest. The profile likelihood has some properties of the usual likelihood (Pace and Salvan 1997, Chapter 4), but for problems with large numbers of nuisance parameters, this procedure results in inconsistent or inefficient estimates. Another problem caused by the number of nuisance parameters is the poor approximation of the LR statistic distribution to the χ^2 distribution. To reduce the effect of the nuisance parameters, Cox and Reid (1987, 1992) proposed a modified profile likelihood. The modified profile likelihood ratio

statistic (LR_m) has asymptotic χ_k^2 null distribution up to an error of order n^{-1} and it can also be corrected by Bartlett correction (DiCiccio and Stern 1994), producing a more accurate inference, as can be seen in Ferrari et al. (2004, 2005), Cysneiros and Ferrari (2006) and Melo et al. (2009).

To improve the large-sample χ^2 approximation to the null distribution of the LR and LR_m statistics in many parametric models, the Bartlett correction is widely used. Although the Bartlett correction factors are somewhat complex to obtain, they can be readily implemented in computer programs. Moreover, this is a worthwhile practice, since Bartlett corrections generally provide a considerable improvement. In recent years, interest in Bartlett corrections has resurfaced and some articles have been published considering this issue. See for example Fujita et al. (2010), Lemonte et al. (2012), Bayer and Cribari-Neto (2013), Stein et al. (2014).

The main purpose of this article is to derive Bartlett correction factors to improve the inference of the dispersion parameter based on the likelihood ratio statistic and modified profile likelihood ratio statistic in the class of heteroscedastic symmetric nonlinear models (HSNLM), a class of models proposed by Cysneiros et al. (2010a), when the number of the observations available is small. To that end, we will follow the approach of modeling the dispersion parameter vector as a function of regressors and unknown parameters such that under the null hypothesis the function is constant, that is, the null hypothesis leads to the symmetric nonlinear regression model. Our results extend some of those obtained in Cordeiro (2004), since we consider a regression structure for the dispersion parameter while they assume the dispersion parameter as a scalar in the class of symmetric nonlinear regression models. We also extend the results obtained in Ferrari et al. (2004), who improved likelihood-based tests for heteroscedasticity in linear regression models. In this work, we also consider the bootstrap Bartlett correction introduced by Rocke (1989) as a numerical alternative to analytical Bartlett correction. A Monte Carlo simulation study is performed to evaluate the performance of the corrected tests and their uncorrected versions. We expected the proposed tests to deliver more trustworthy inferences in small samples and our Monte Carlo simulation results show that, in fact, this happens. That is, the corrected likelihood ratio and corrected modified profile likelihood ratio tests proposed here are attractive alternatives to the usual likelihood ratio test in the HSNLM class when the sample size is small. We are unaware of any simulation study in the statistical literature comparing the performance of the proposed tests in HSNLM. Thus, this paper fills this gap.

The article is organized as follows. In Sect. 2, we define the model and present some inferential aspects such as estimation and hypothesis testing of regression parameters. In Sect. 3, we discuss Bartlett corrections to improve the usual likelihood ratio and modified profile likelihood ratio tests in HSNLM. We also present the bootstrap Bartlett correction for likelihood ratio statistic. A Monte Carlo simulation is performed in Sect. 4 to evaluate the performance of the studied tests. An application using real data is considered in Sect. 5. Conclusions about the results obtained are presented in Sect. 6. Finally, appendices with technical details are presented at the end.

2 Heteroscedastic symmetric nonlinear models

Let $y_1, \ldots, y_n$ be n independent random variables. Each y_ℓ, $\ell = 1, \ldots, n$, follows a continuous symmetric distribution with location parameter $\mu_\ell \in \mathbb{R}$ and dispersion parameter $\phi_\ell > 0$ if its probability density function is of the form

$$\pi(y_\ell; \mu_\ell, \phi_\ell) = \frac{1}{\sqrt{\phi_\ell}} g(u_\ell), \quad y_\ell \in \mathbb{R} \tag{1}$$

where $g : \mathbb{R} \to [0, \infty)$ is generally known as the density generator, such that, $\int_0^\infty g(u)du < \infty$, with $u_\ell = (y_\ell - \mu_\ell)^2/\phi_\ell$. In what follows, we will denote $y_\ell \sim S(\mu_\ell, \phi_\ell, g)$. The density generation function $g(\cdot)$ for some symmetric distributions is given in Table 1.

The heteroscedastic symmetric nonlinear regression model proposed by Cysneiros et al. (2010a) is defined as:

$$y_\ell = \mu_\ell + \sqrt{\phi_\ell} e_\ell, \quad \ell = 1, \ldots, n, \tag{2}$$

where $\mu_\ell = f(x_\ell; \boldsymbol{\beta})$ is a continuous and twice differentiable nonlinear regression structure with respect to the components of a vector of unknown regression parameters $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top, (p < n)$, $\boldsymbol{x}_\ell = (x_{\ell 1}, \ldots, x_{\ell P})^\top$ is a vector of known explanatory variables associated with the ℓth observation and $e_\ell \sim S(0, 1, g)$. Moreover, we assume that $\boldsymbol{\beta}$ is defined in a subset $\boldsymbol{\Omega_\beta} \in \mathbb{R}^p$ $(p < n)$ such that the $n \times p$ matrix of derivatives of $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ with respect to $\boldsymbol{\beta}$, denoted by $\tilde{X} = \partial\boldsymbol{\mu}/\partial\boldsymbol{\beta}$, has rank p for all $\boldsymbol{\beta}$. In addition we consider $\phi_\ell = \sigma^2 m(\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta})$, where $m(\cdot) > 0$ is any known one-to-one continuously differentiable function, where $\boldsymbol{\omega}_\ell = (\omega_{\ell 1}, \ldots, \omega_{\ell k})^\top$ is a vector of explanatory variables that may have components in common with $\boldsymbol{x}_\ell$, $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_k)^\top$ is a vector of unknown parameters to be estimated, and $\sigma^2 \in (0, +\infty)$ is an unknown constant. We also assume that a unique value $\boldsymbol{\delta_0}$ of $\boldsymbol{\delta}$ exists such that $m(\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta_0}) = 1$ for all ℓ. Consequently, $\phi_\ell = \sigma^2$, if $\boldsymbol{\delta} = \boldsymbol{\delta_0}$, which implies that the y_ℓ's have constant dispersion. Note that the concept of heteroscedasticity in this context refers to varying dispersion, i.e., we say that the model is homoscedastic

Table 1 Density generation function $g(\cdot)$ for some symmetric distributions

Distributions	$g(u)$
Normal	$\frac{1}{\sqrt{2\pi}} \exp\{-u/2\}$
Cauchy	$\frac{1}{\pi}(1+u)^{-1}$
Student-t	$\frac{\nu^{\frac{\nu}{2}}}{B(1/2, \nu/2)}(\nu+u)^{-\frac{\nu+1}{2}}$
Type I logistic	$c\frac{\exp\{-u\}}{(1+\exp\{-u\})^2}$ [a]
Type II logistic	$\frac{\exp\{-u^{1/2}\}}{(1+\exp\{-u^{1/2}\})^2}$
Power exponential	$c(k)\exp\left\{-\frac{1}{2}u^{-(1+k)}\right\}$ [b]

[a] $c \approx 1,484300029$ is a normalizing constant which follows from $\int_0^\infty u^{-1/2} g(u) du = 1$.
[b] $c(k)^{-1} = \Gamma(1 + \frac{1+k}{2})2^{1+(1+k)/2}$, $-1 < k \le 1$

when all dispersion parameters $\phi_1, \ldots, \phi_n$ are equal; otherwise we say that the model is heteroscedastic. To perform the procedure, an explicit form for m must be chosen. A possible and common choice is to consider $m(\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta}) = \exp(\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta})$, since this functional form for m does not impose any restriction on the components of $\boldsymbol{\omega_\ell}$ (Cook and Weisberg 1983; Lin et al. 2009).

In this way, we are interested in assessing the constancy of the dispersion parameter in model (2) by testing the null hypothesis $H_0 : \boldsymbol{\delta} = \boldsymbol{\delta_0}$ against the alternative hypothesis $H_1 : \boldsymbol{\delta} \neq \boldsymbol{\delta_0}$, where $\boldsymbol{\delta_0}$ is a $k \times 1$ vector of specified constant such that $m_\ell = m(\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta_0}) = 1$ for $\ell = 1, \ldots, n$. In other words, we are performing a test for heteroscedasticity in symmetric nonlinear regression models, since under the null hypothesis, model (2) reduces to the aforementioned class of models. The number of parameters of interest is k and the number of nuisance parameters is $p + 1$. The total log-likelihood function for the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top, \sigma^2)^\top$ given $y_1, \ldots, y_n$ in model (2) is expressed by:

$$l(\boldsymbol{y}; \boldsymbol{\theta}) = -\frac{n}{2} \log \sigma^2 - \frac{1}{2} \sum_{\ell=1}^{n} \log(m_\ell) + \sum_{\ell=1}^{n} t(z_\ell),$$

where $t(z_\ell) = \log g(z_\ell^2)$ with $z_\ell = \sqrt{u_\ell} = \frac{(y_\ell - \mu_\ell)}{\sqrt{\phi_\ell}}$.

Note that to obtain the maximum likelihood estimator (MLE) of $\boldsymbol{\delta}$ we maximize the profile log-likelihood function

$$l_p(\boldsymbol{\delta}) = l\left(\boldsymbol{y}; \boldsymbol{\delta}, \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\sigma}_{\boldsymbol{\delta}}^2\right),$$

where $\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}$ and $\hat{\sigma}_{\boldsymbol{\delta}}^2$ are MLEs of $\boldsymbol{\beta}$ and σ^2 given $\boldsymbol{\delta}$, respectively. Under usual regularity conditions, $\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}$ and $\hat{\sigma}_{\boldsymbol{\delta}}^2$ are solutions of the equations $\boldsymbol{U_\beta} = \boldsymbol{0}$ and $U_{\sigma^2} = \boldsymbol{0}$, respectively, which cannot be obtained in closed form. Thus, $\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}$ and $\hat{\sigma}_{\boldsymbol{\delta}}^2$ are derived from computationally iterative restricted maximization techniques. Further details of these techniques can be obtained from Nocedal and Wright (1999).

The likelihood ratio statistic (LR) for test H_0 can be written as

$$LR = 2\{l_p(\hat{\boldsymbol{\delta}}) - l_p(\boldsymbol{\delta_0})\},$$

where $\hat{\boldsymbol{\delta}}$ is the MLE of $\boldsymbol{\delta}$. Asymptotically and under the null hypothesis, LR has χ_k^2 distribution.

When replacing the nuisance parameters by their maximum likelihood estimates, we are in a way treating them as known, and as a consequence, the profile log-likelihood function may present biases in the score and information function (Ferrari et al. 2005). This procedure is also known to provide inconsistent or inefficient estimates for problems with large numbers of nuisance parameters. Cox and Reid (1987) proposed a modified version of the profile likelihood function in order to attenuate the impact of the number of nuisance parameters on the resulting inference. However, that version requires orthogonality between parameters of interest and nuisance ones. Therefore, in our case $\boldsymbol{\delta}$ should be orthogonal to the remaining parameters. For this,

a transformation $(\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \sigma^2)^\top \to (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \gamma)^\top$ is necessary and sufficient, where $E[-\partial l/\partial \delta_a \partial \gamma] = 0$, $a = 1, \ldots, k$. Following Cox and Reid (1987, Eq. 4), the desired transformation is obtained by solving

$$\frac{n}{2\sigma^4}\frac{\partial \sigma^2}{\partial \delta_a} = -\frac{1}{2\sigma^2}\sum_{\ell=1}^{n}\frac{\partial m_\ell}{\partial \delta_a}\frac{1}{m_\ell},$$

which has solution (Simonoff and Tsai 1994)

$$\sigma^2 = \frac{\gamma}{(\prod_{\ell=1}^{n} m_\ell)^{1/n}}.$$

Considering the reparameterized model, the modified profile log-likelihood function for $\boldsymbol{\delta}$ (Cox and Reid 1987) is given by

$$l_{CR}^*(\boldsymbol{\delta}) = l_p^*(\boldsymbol{\delta}) - \frac{1}{2}\log\{\det[j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})]\}, \tag{3}$$

where $l_p^*(\boldsymbol{\delta}) = l^*(\mathbf{y}; \boldsymbol{\delta}, \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}}) = -\frac{n}{2}\log\gamma + \sum_{\ell=1}^{n} t(z_\ell)$ correspond to the profile log-likelihood function for $\boldsymbol{\delta}$, $j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})$ denotes the block of the observed information matrix for the nuisance parameters $(\boldsymbol{\beta}^\top, \gamma)^\top$ evaluated at $(\boldsymbol{\delta}, \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})$, and $\hat{\gamma}_{\boldsymbol{\delta}} = \hat{\sigma}_{\boldsymbol{\delta}}^2 \left(\prod_{\ell=1}^{n} m_\ell\right)^{1/n}$. The matrix $j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})$ is shown in Appendix A.

The modified profile likelihood ratio statistic (LR_m) for testing H_0 against H_1 is given by

$$LR_m = 2\{l_{CR}(\hat{\boldsymbol{\delta}}) - l_{CR}(\boldsymbol{\delta_0})\},$$

where $\hat{\boldsymbol{\delta}}$ is the MLE of $\boldsymbol{\delta}$. Under the null hypothesis, the LR_m statistic has asymptotic χ_k^2 distribution.

3 Bartlett corrections

For large samples, the null distributions of LR and LR_m statistics are approximated by the χ^2 distribution. On the other hand, if the sample size is not large enough, it is well known that these approximations cannot be satisfactory, leading to size-distorted tests. In order to improve such approximations, some correction factors for incorporation in LR and LR_m statistics have been proposed in the literature, yielding corrected test statistics whose null distributions are better approximated by the reference χ^2 distribution, that is, the approximation error from the corrected test statistic distribution to the χ^2 one reduces from order n^{-1} to n^{-2}. The ideas of transforming the LR and LR_m statistics to make their distributions better approximated by the chi-squared distribution are due to Bartlett (1937) and DiCiccio and Stern (1994), respectively. These correction factors do not depend on a particular parametric model, i.e., they are very general and need to be obtained for each problem of interest.

3.1 Bartlett correction for the LR statistic

It is known that for large samples and under the null hypothesis, the LR statistic has chi-square distribution up to error of order n^{-1}. Bartlett (1937) proposed to multiply the LR statistic by a correction factor, denoted by $(1 + c/k)^{-1}$, resulting in a corrected statistic LR^* given by

$$LR^* = \frac{LR}{1 + c/k},$$

where c is a constant of order n^{-1} that can be estimated under H_0. Moreover, c can be written in terms of moments of likelihood derivatives up to the fourth order, see Lawley (1956). Particularly, $P(LR^* \leq x) = P(\chi_k^2 \leq x) + O(n^{-2})$[1] under the null hypothesis. Further details on Bartlett corrections can be seen in Cordeiro and Cribari-Neto (2014).

In what follows, we shall consider the case of heteroscedasticity with multiplicative effects, that is, the special case where $m_\ell = \exp(\omega_\ell^\top \delta)$. Thus, to test $H_0 : \delta = \delta_0$ against $H_1 : \delta \neq \delta_0$ in the class of HSNLM, the constant c from the Bartlett correction factor for the LR statistic can be expressed as

$$c = \epsilon_k(\delta) + \epsilon_{p,k}(\beta, \delta) + \epsilon_{p,k}(\delta, \gamma) + \epsilon_{p,k}(\beta, \delta, \gamma), \tag{4}$$

where

$$\epsilon_k(\delta) = \frac{\Delta_2}{4}\text{tr}(H_d^2) + \frac{\Delta_1^2}{6}\iota^\top H^{(3)}\iota + \frac{\Delta_1^2}{4}\iota^\top H H_d H\iota,$$

$$\begin{aligned}
\epsilon_{p,k}(\beta, \delta) = & -\frac{\Delta_6}{4\delta_{(0,1,0,0,0)}}\iota^\top Q H_d Z_{\beta d}\iota - \frac{\Delta_7}{4\delta_{(0,1,0,0,0)}}\iota H_d Z_{\beta d}\iota \\
& + \frac{\delta_{(0,0,1,0,1)}}{2\delta_{(0,1,0,0,0)}}\iota Q H_d Z_{\beta d}\iota + \iota^\top Q H_d Z_{\beta d}\iota \\
& - \frac{\Delta_6}{4\delta_{(0,1,0,0,0)}}\iota^\top Q H_d Z_{\beta d}\iota - \frac{\Delta_7}{4\delta_{(0,1,0,0,0)}}\iota H_d Z_{\beta d}\iota \\
& + \left(\frac{\Delta_7^2}{2(\delta_{(0,1,0,0,0)})^2} - \frac{\Delta_7}{\delta_{(0,1,0,0,0)}}\right)\iota^\top Q Z_\beta \odot H \odot Z_\beta Q\iota \\
& + \frac{\Delta_7^2}{4(\delta_{(0,1,0,0,0)})^2}\iota^\top Q Z_{\beta d} H Z_{\beta d} Q\iota \\
& + \frac{\Delta_1 \Delta_7}{2\delta_{(0,1,0,0,0)}}\iota^\top Q Z_{\beta d} H H_d\iota,
\end{aligned}$$

---

[1] Let $\{a_n\}$ and $\{b_n\}$ be real sequences. a_n is said to be at most of order equal to b_n, denoted by $a_n = O(b_n)$, if $K \in \mathbb{R}^+$ exists and $n_0(K)$, such that $|a_n/b_n| \leq K$, $\forall n \geq n_0(K)$.

$$\epsilon_{p,k}(\delta, \gamma) = -\frac{\Delta_4 \Delta_8}{2}\text{tr}(\boldsymbol{H}_d) - \Delta_1 \Delta_4 \text{tr}(\boldsymbol{H}_d) - \frac{\Delta_1^2 \Delta_4}{2}\boldsymbol{\iota}^\top \boldsymbol{H}^{(2)} \boldsymbol{\iota}$$
$$- \frac{\Delta_1^2 \Delta_4}{4}(\text{tr}(\boldsymbol{H}_d))^2 + \left(\frac{\Delta_1 \Delta_5}{4} + \frac{\Delta_1(\Delta_5 - 4\Delta_3)}{4}\right)\Delta_4^2 \text{tr}(\boldsymbol{H}_d)$$

and

$$\epsilon_{p,k}(\boldsymbol{\beta}, \delta, \gamma) = -\frac{\Delta_1 \Delta_4 \Delta_7}{4\delta_{(0,1,0,0,0)}}\left[(\boldsymbol{\iota}^\top \boldsymbol{W} \boldsymbol{Z}_{\beta d} \boldsymbol{\iota}) \odot (\boldsymbol{\iota}^\top \boldsymbol{H}_d \boldsymbol{\iota}) + \boldsymbol{\iota}^\top \boldsymbol{Q} \boldsymbol{H}_d \boldsymbol{Z}_{\beta d} \boldsymbol{\iota}\right],$$

where Δ_ℓ, $\ell = 1, \ldots, 8$, are scalars expressed as

$$\Delta_1 = -\frac{1}{8}\{\delta_{(0,0,1,0,3)} + 3\delta_{(0,1,0,0,2)} + \delta_{(1,0,0,0,1)}\},$$

$$\Delta_2 = \frac{1}{16}\{\delta_{(0,0,0,1,4)} + 6\delta_{(0,0,1,0,3)} + 7\delta_{(0,1,0,0,2)} + \delta_{(1,0,0,0,1)}\},$$

$$\Delta_3 = -\frac{n}{2}\{2 + 3\delta_{(1,0,0,0,1)} + \delta_{(0,1,0,0,2)}\},$$

$$\Delta_4 = \frac{4}{n\{2 + \delta_{(0,1,0,0,2)} + 3\delta_{(1,0,0,0,1)}\}},$$

$$\Delta_5 = -\frac{n}{8}\{8 + 15\delta_{(1,0,0,0,1)} + 9\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)}\},$$

$$\Delta_6 = \frac{1}{4}\{\delta_{(0,0,0,1,2)} + 3\delta_{(0,0,1,0,1)}\},$$

$$\Delta_7 = \frac{1}{2}\{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}\} \text{ and}$$

$$\Delta_8 = \frac{1}{16}\{\delta_{(0,0,0,1,4)} + 8\delta_{(0,0,1,0,3)} + 13\delta_{(0,1,0,0,2)} + 3\delta_{(1,0,0,0,1)}\}.$$

The δ's correspond to $\delta_{(a,b,c,d,e)} = E\{t(z_\ell)^{(1)a}t(z_\ell)^{(2)b}t(z_\ell)^{(3)c}t(z_\ell)^{(4)d}z_\ell^e\}$ for $a, b, c, d, e \in \{1, 2, 3, 4\}$ and $t(z_\ell)^{(k)} = \frac{\partial^k t(z_\ell)}{\partial z_\ell^k}$ for $k = 1, 2, 3, 4$. Some δ's values for symmetric distributions studied in the literature can be found in Uribe-Opazo et al. (2008). In addition, $\boldsymbol{H} = \{h_{\ell s}\} = -(\boldsymbol{W} - \bar{\boldsymbol{W}})[(\boldsymbol{W} - \bar{\boldsymbol{W}})^\top \boldsymbol{V}(\boldsymbol{W} - \bar{\boldsymbol{W}})]^{-1}(\boldsymbol{W} - \bar{\boldsymbol{W}})^\top$, with $(\boldsymbol{W} - \bar{\boldsymbol{W}}) = (\boldsymbol{w}_1 - \bar{\boldsymbol{w}}, \ldots, \boldsymbol{w}_n - \bar{\boldsymbol{w}})^\top$, $\boldsymbol{V}$ representing the diagonal matrix of order n with $v_\ell = (1 - \delta_{(0,1,0,0,2)})/4$ and $\ell, s = 1, \ldots, n$. We also have $\boldsymbol{H}^{(2)} = (h_{\ell s}^2)$, $\boldsymbol{H}^{(3)} = (h_{\ell s}^3)$, $\boldsymbol{Q} = \text{diag}(q_1, \ldots, q_n)$, with $q_\ell = \exp\{-(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})^\top \boldsymbol{\delta}\}$ and $\bar{\boldsymbol{\omega}} = (\bar{\omega}_1, \ldots, \bar{\omega}_k)^\top$, $\boldsymbol{Z}_\beta = \tilde{\boldsymbol{X}}(\tilde{\boldsymbol{X}}^\top \boldsymbol{Q} \tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^\top$, an $n \times p$ matrix $\boldsymbol{W}$ with the ijth element given by ω_{ij} and an $n \times 1$ vector of ones represented by $\boldsymbol{\iota}$. The subscripts d in some matrices indicate that only diagonal elements of these matrices were considered. Finally, the symbol $\odot$ denotes the Hadamard (elementwise) product of matrices.

Therefore, we can observe that the constant c involves only simple matrix operations and can be easily implemented in symbolic computation packages and programming languages that allow performing basic linear algebra operations, such as Ox and R. Details about the derivation of the expression c are presented in Appendices B and C.

3.2 Bartlett correction for the LR_m statistic

The modified profile likelihood ratio statistic, LR_m, as well as the usual likelihood ratio statistic, LR, has null asymptotic χ_k^2 distribution up to an error of order n^{-1}. DiCiccio and Stern (1994) proposed a Bartlett correction for LR_m, reducing the approximation error from its distribution to the reference χ^2 distribution to order n^{-2}. The corrected statistic is defined by

$$LR_m^* = \frac{LR_m}{1 + c_m/k},$$

where c_m is a constant of order n^{-1} such that under the null hypothesis the expected value of the corrected modified test statistic is written as $E(LR_m^*) = k + O(n^{-3/2})$. The general expression of c_m was defined in DiCiccio and Stern (1994, Eq. 25). In this context, Ferrari et al. (2004, Eq. 5) obtained an equation to calculate c_m from normal linear regression models which can be used in any class of models that uses the partition of the parameter vector we are using and where orthogonality holds. To test H_0 in the HSNLM class considering heteroscedasticity with multiplicative effect, the constant c_m for the correction factor for the LR_m statistic is used, which is presented in details in Appendix D. It is written in matrix notation as

$$c_m = \frac{1}{4}\Delta_2 \text{tr}(\boldsymbol{H_d^2}) - \frac{1}{4}\Delta_1^2\Delta_4[\text{tr}(\boldsymbol{H_d})]^2 + \frac{1}{4}\Delta_1^2\boldsymbol{\iota}^\top \boldsymbol{H_d}\boldsymbol{H}\boldsymbol{H_d}\boldsymbol{\iota} + \frac{1}{6}\Delta_1^2\boldsymbol{\iota}^\top \boldsymbol{H}^{(3)}\boldsymbol{\iota}$$
$$- \Delta_1\Delta_4\text{tr}(\boldsymbol{H_d}) - \Delta_1\Delta_3\Delta_4^2\text{tr}(\boldsymbol{H_d}) - \frac{1}{2}\Delta_1^2\Delta_4\boldsymbol{\iota}^\top \boldsymbol{H}^{(2)}\boldsymbol{\iota}. \tag{5}$$

We can see that the constant c_m involves only simple matrix operations, such as the constant c in (4) for the correction factor of the LR statistic. We also observe that c_m in (5) involves only matrix $\boldsymbol{W}$ of covariates (defined in 4), the number of unknown parameters in ϕ_ℓ and the number of observations, not depending on unknown parameters or the number of nuisance parameters. Both c and c_m depend on the symmetric distribution considered, since the $\delta's$ values change from one distribution to another.

3.3 Bootstrap Bartlett correction

As an alternative to the asymptotic likelihood ratio and modified profile likelihood ratio tests, we can carry out the inference based on a test with critical values ($p-$values) obtained from the bootstrap technique introduced by Efron (1979). The bootstrap likelihood ratio test (LR_{boot}) offers a reliable inference and does not involve complex calculations. However, it requires very intensive computing. Considering the bootstrap a numerical alternative to the Bartlett correction factor for the LR statistic, deriving the bootstrap Bartlett technique, Rocke (1989) proposed corrected likelihood ratio statistic (LR_{boot}^*), which is obtained as follows. Initially, we generate B bootstrap resamples $(y_1^*, \ldots, y_B^*)$ from the assumed model under H_0, replacing the unknown parameter vector with its respective restricted estimates (i.e., the estimates obtained under the null hypothesis), calculated using the original sample $(y_1, \ldots, y_n)$. After

that, we calculate the LR statistic for each pseudo sample $y_1^*, \ldots, y_B^*$, denoted by LR_{boot}^b, $b = 1, \ldots, B$. The bootstrap Bartlett corrected likelihood ratio statistic is obtained by

$$LR_{boot}^* = \frac{LR}{\overline{LR}_{boot}^*} k,$$

where $\overline{LR}_{boot}^* = \frac{1}{B} \sum_{b=1}^B LR_{boot}^b$ and k is the number of restrictions imposed by H_0. Under the null hypothesis, LR_{boot}^* has asymptotic χ_k^2 distribution (Rocke 1989).

The LR_{boot} statistic does not follow the χ^2 distribution. Instead it, is based on this statistic, performed as described below: for a fixed nominal level α, we calculate the $1 - \alpha$ percentile of LR_{boot}, which is estimated by $\hat{q}_{(1-\alpha)}$ such that #$\{LR_{boot} \leq \hat{q}_{(1-\alpha)}\}/B = 1 - \alpha$, with # denoting the cardinality set. Then, we reject the null hypothesis if $LR > \hat{q}_{(1-\alpha)}$. Alternatively, the decision rule can be written based on the bootstrap p-value given by $p^* = \#\{LR_{boot}^b \geq LR\}/B$.

Recent works have developed inference based on these tests, see for example Bayer and Cribari-Neto (2013), Cribari-Neto and Queiroz (2014), and Loose et al. (2016). An advantage of using the bootstrap Bartlett correction instead of the usual bootstrap technique is its computational efficiency. To obtain a critical value using the bootstrap Bartlett correction requires a resample with smaller size than the one needed when using the bootstrap technique, which implies that the bootstrap Bartlett correction is computationally more efficient than the usual bootstrap technique, see (Rocke 1989).

4 Simulation results

In this section, we present Monte Carlo simulation results to compare the performance of seven tests in HSNLM: the likelihood ratio test (LR); the modified profile likelihood ratio test (LR_m); the bootstrap likelihood ratio test (LR_{boot}); their corrected versions, denoted by LR^*, LR_m^* and LR_{boot}^*, respectively; and the score test (S_r) in small and moderate-sized samples. As known, an advantage of the S_r test in relation to the others is its easy implementation, since it involves estimations under the null hypothesis only. The number of Monte Carlo replications was 10,000 and for each Monte Carlo replication we performed 1000 bootstrap replications. All simulations were performed using the programming language Ox (Doornik 2006). We considered the heteroscedastic symmetric nonlinear regression model given by

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

where $\epsilon_\ell \sim S(0, \sigma^2 \exp\{\boldsymbol{\omega_\ell}^\top \boldsymbol{\delta}\}, g)$. The covariates $x_1, \ldots, x_{p-1}$ and $\omega_1, \ldots, \omega_q$ were obtained as random draws from the uniform $U(0, 1)$ and their values were kept fixed during the simulations. We considered two symmetric distributions for the errors, namely Student-t with 5 degrees of freedom (ν) and power exponential with shape parameter $\kappa = 0.3$. The test to be considered is $H_0 : \delta_1 = \cdots = \delta_q = 0$ against

Table 2 Null rejection rates: t_5 model with $p = 3, 5, k = 3$ and several values for n

α	Test	$p = 3$				$p = 5$			
		n				n			
		30	40	50	100	30	40	50	100
$\alpha = 10\%$	LR	21.2	19.4	18.3	16.3	33.0	25.5	19.7	14.4
	LR^*	13.2	12.5	11.4	10.7	21.1	15.9	13.8	11.6
	LR_m	9.2	9.4	9.4	10.0	8.8	10.2	9.9	9.9
	LR_m^*	9.6	9.7	9.6	10.3	9.3	10.4	10.1	10.1
	LR_{boot}	10.2	9.8	10.1	9.8	10.3	10.8	10.2	10.3
	LR_{boot}^*	10.2	9.8	10.0	9.9	10.3	10.7	10.2	10.2
	S_r	10.5	10.3	10.4	10.6	11.4	11.4	11.2	11.0
$\alpha = 5\%$	LR	13.6	11.9	10.7	9.5	22.6	16.8	12.3	7.7
	LR^*	7.0	6.6	5.9	5.4	12.9	8.8	7.4	5.8
	LR_m	4.7	4.7	4.9	5.1	4.2	5.0	4.5	5.0
	LR_m^*	5.0	4.9	5.0	5.3	4.5	5.2	4.6	5.2
	LR_{boot}	5.2	4.9	5.1	4.7	5.2	5.5	5.2	5.2
	LR_{boot}^*	5.4	4.8	5.2	4.7	5.1	5.7	5.2	5.0
	S_r	5.4	5.1	5.2	5.2	5.5	5.5	5.4	5.1
$\alpha = 1\%$	LR	4.7	4.0	3.4	2.6	9.5	6.5	3.8	1.9
	LR^*	1.6	1.4	1.2	1.1	3.9	2.4	1.9	1.1
	LR_m	0.9	0.9	1.0	1.0	1.0	0.9	1.0	1.0
	LR_m^*	1.0	0.9	1.0	1.0	1.0	1.0	1.0	1.0
	LR_{boot}	1.2	1.1	1.0	0.9	0.9	1.0	1.1	1.2
	LR_{boot}^*	1.2	1.1	0.9	0.8	0.9	1.0	1.0	1.0
	S_r	1.0	0.8	1.2	1.0	1.0	1.2	1.0	1.0

$H_1 : \delta_i \neq 0$ *for at least one* i, $i = 1, \ldots, q$. The true values of the parameters for the simulations were taken as $\beta_0 = \cdots = \beta_{p-1} = 1$, $\sigma^2 = 1$, $\delta_1 = 0.1, \delta_2 = 0.3, \delta_3 = 0.5$ and $\delta_4 = \delta_5 = 1.0$. The null hypothesis was tested for sample sizes 30, 35, 40, 50 and 100 considering the three nominal levels $\alpha = 10, 5$ and 1%.

The null rejection rates of the seven tests for different sample sizes are presented in Tables 2 and 3. In these tables, it can be noted that the likelihood ratio test is substantially oversized, for example, in Table 2 when $p = 5, \alpha = 5\%$ and considering all sample sizes ($n = 30, 40, 50$ and 100), the null rejection rates of the LR test are 22.6, 16.8, 12.3 and 7.7%, respectively. The corrected version of the likelihood ratio test tends to attenuate the oversized behavior of the usual likelihood ratio test, but still presents higher rejection rates. For example, considering again $p = 5, \alpha = 5\%$ and all sample sizes, the LR^* test presents the following null rejection rates: 12.9, 8.8, 7.4 and 5.8%, respectively. In general, for both tests, as the sample size increases, the distortion of the tests decreases.

The simulation results for the t_5 model shown in Table 2 indicate that the S_r test and the corrected tests LR_m^* and LR_{boot}^* present better results than the other ones.

Table 3 Null rejection rates: power exponential model with $\kappa = 0.3$, $p = 3, 5$, $k = 3$ and several values for n.

α	Test	$p = 3$				$p = 5$			
		n				n			
		30	40	50	100	30	40	50	100
$\alpha = 10\%$	LR	22.6	17.2	14.9	12.1	32.0	22.2	20.2	14.2
	LR^*	14.9	12.8	11.6	10.4	22.4	18.6	16.2	13.7
	LR_m	7.7	8.9	9.3	9.7	6.1	8.3	9.1	10.2
	LR_m^*	8.8	9.7	10.2	10.0	7.2	9.1	9.9	10.6
	LR_{boot}	10.2	10.1	10.1	10.1	9.9	10.6	9.3	9.6
	LR_{boot}^*	10.1	10.1	9.9	10.1	9.8	10.7	9.3	9.7
	S_r	10.4	10.0	9.8	9.8	11.1	10.4	10.8	10.3
$\alpha = 5\%$	LR	14.1	10.1	8.6	6.4	22.8	14.2	12.0	7.8
	LR^*	8.5	6.9	6.1	5.5	14.0	10.8	9.3	7.4
	LR_m	3.5	4.4	4.7	4.8	2.6	3.8	4.2	4.7
	LR_m^*	4.2	5.0	5.2	5.1	3.3	4.3	4.8	4.9
	LR_{boot}	5.1	5.0	5.1	5.0	5.1	5.5	4.8	4.8
	LR_{boot}^*	5.2	4.9	5.1	5.1	4.9	5.5	4.8	4.6
	S_r	5.4	5.2	5.0	5.0	6.0	5.1	5.5	5.3
$\alpha = 1\%$	LR	4.7	3.0	2.4	1.0	9.6	4.8	3.6	1.9
	LR^*	2.2	1.5	1.4	1.1	4.7	3.2	2.8	1.9
	LR_m	0.4	0.8	1.2	1.0	0.4	0.6	0.7	0.9
	LR_m^*	0.6	1.0	1.3	1.1	0.5	0.7	0.8	1.0
	LR_{boot}	0.9	1.1	1.2	1.1	1.1	1.2	1.1	1.1
	LR_{boot}^*	0.9	1.1	0.8	1.1	1.0	1.1	1.0	1.0
	S_r	1.6	1.3	1.0	1.1	1.3	1.2	1.2	1.2

Furthermore, for $p = 5$ and $\alpha = 5\%$, the rejection rates for the S_r test considering all four sample sizes are, respectively, 5.5, 5.5, 5.4 and 5.1% and the corresponding rates for the LR_m^* test are 4.5, 5.2, 4.6 and 5.2% and for the LR_{boot}^* test are 5.8, 4.8, 5.2 and 4.7%. Table 3 presents the results for the power exponential model. We can observe that, in general, the tests LR_{boot} and LR_{boot}^* perform better than the others, followed by the S_r test. For example, when $p = 5$ and $\alpha = 10\%$, the rejection rates for the LR_{boot} test considering all four sample sizes are 9.9, 10.6, 9.3 and 9.6%, while for the LR_{boot}^* test they are 9.8, 10.7, 9.3 and 9.7%, and for the S_r test they are 11.1, 10.4, 10.8 and 10.3%.

In Table 4 we present the null rejection rates of the tests and evaluate the effect of the number of nuisance parameters on the performance of the tests, fixing the sample size ($n = 35$), the number of parameters of interest ($k = 3$) and varying the number of nuisance parameters ($p = 2, 3, 4$ and 5). For both considered models, the usual LR test and its corrected version are quite distorted, and when the number of nuisance parameters increases, the distortion of the LR and LR^* tests also increases. In

Table 4 Null rejection rates: the t_5 model and power exponential $\kappa = 0.3$ with $n = 35, k = 3$ and multiple values for p

α	Test	t_5 model				Power exponential model			
		p				p			
		2	3	4	5	2	3	4	5
10%	LR	16.3	19.4	24.5	29.1	15.4	19.0	23.0	28.6
	LR^*	11.4	12.9	15.4	17.9	12.3	13.7	15.8	18.8
	LR_m	9.4	9.4	9.7	9.0	8.5	8.3	7.6	7.5
	LR_m^*	9.7	9.7	10.0	9.3	9.4	9.6	8.5	8.7
	LR_{boot}	10.1	10.4	9.8	9.8	10.0	10.4	10.2	10.0
	LR_{boot}^*	10.0	10.3	9.8	9.8	10.0	10.3	10.3	9.9
	S_r	10.2	10.4	8.4	11.0	9.7	9.8	8.6	10.8
5%	LR	9.3	11.9	15.8	19.4	8.9	11.5	14.7	19.1
	LR^*	5.9	7.0	8.8	10.5	6.4	7.5	8.8	11.4
	LR_m	4.5	4.7	4.5	4.3	4.2	3.8	3.4	3.6
	LR_m^*	4.8	4.9	4.7	4.5	5.0	4.6	4.1	4.3
	LR_{boot}	4.9	5.1	5.0	5.1	5.0	5.3	5.2	5.0
	LR_{boot}^*	4.9	5.1	4.8	4.9	5.0	5.3	5.2	5.0
	S_r	4.7	5.2	4.3	5.3	5.5	5.0	4.9	5.4
1%	LR	2.8	4.0	5.8	7.5	2.7	3.9	5.3	7.9
	LR^*	1.2	1.8	2.3	3.2	1.5	1.1	2.4	3.8
	LR_m	0.7	0.9	1.0	0.9	0.8	0.7	0.6	0.6
	LR_m^*	0.8	0.9	1.0	1.0	1.0	0.9	0.7	0.7
	LR_{boot}	1.0	1.0	0.9	1.1	1.1	1.1	1.1	1.5
	LR_{boot}^*	0.9	0.9	0.9	1.1	0.9	1.2	1.0	1.3
	S_r	0.8	0.9	1.0	1.1	1.8	1.1	1.5	1.2

contrast, the number of nuisance parameters is indifferent to the other tests, and the bootstrap tests LR_{boot} and LR_{boot}^* performed best. For example, for model t_5 when $p = 4$ and $\alpha = 5\%$, the null rejection rates of the tests are 15.8% (LR), 8.8% (LR^*), 4.5% (LR_m), 4.7% (LR_m^*), 5.0% (LR_{boot}), 4.8% (LR_{boot}^*) and 4.3% (S_r). Assuming the same scenario, we have for the power exponential model the following null rejection rates: 14.7% (LR), 8.8% (LR^*), 3.4% (LR_m), 4.1% (LR_m^*), 5.2% (LR_{boot}), 5.2% (LR_{boot}^*) and 4.9% (S_r).

Table 5 reports results for the situation where $n = 35$, $p = 3$ and $k = 2, 3, 4$ and 5. The tests' performances are similar to that shown in Table 4, with the LR_{boot}^* test outperforming the other ones, especially in the power exponential model. For example, considering the scenario where $k = 3$ and $\alpha = 1\%$, the null rejection rates for the power exponential model are 3.9% (LR), 1.1% (LR^*), 0.7% (LR_m), 0.9% (LR_m^*), 1.1% (LR_{boot}), 1.2% (LR_{boot}^*) and 1.1% (S_r). Considering the same scenario for the t_5 model, we have 4% (LR), 1.8% (LR^*), 0.8% (LR_m), 0.8% (LR_m^*), 1.1% (LR_{boot}), 0.9% (LR_{boot}^*) and 0.9% (S_r).

Table 5 Null rejection rates: t_5 model and power exponential $\kappa = 0.3$ with $n = 35$, $p = 3$ and different values of k

α	Test	t_5 model				Power exponential model			
		k				k			
		2	3	4	5	2	3	4	5
10%	LR	16.0	19.5	21.5	23.9	16.4	18.9	20.0	21.8
	LR^*	11.4	12.3	13.7	15.2	12.0	13.7	14.1	15.4
	LR_m	9.1	9.1	9.7	10.2	9.0	8.3	8.0	7.3
	LR_m^*	9.4	9.4	10.0	10.3	10.0	9.5	9.0	8.5
	LR_{boot}	9.9	9.8	10.4	10.5	9.7	10.4	9.7	10.7
	LR_{boot}^*	9.8	9.9	10.2	10.4	9.6	10.3	9.6	10.8
	S_r	10.3	10.4	10.4	11.5	10.1	9.8	10.1	10.5
5%	LR	9.2	11.8	13.0	15.0	9.6	11.5	12.1	13.4
	LR^*	6.0	7.0	7.2	8.6	6.2	7.5	7.7	8.9
	LR_m	4.7	4.4	4.5	5.4	4.8	3.8	3.7	3.5
	LR_m^*	5.0	4.6	4.7	5.4	5.0	4.6	4.4	4.2
	LR_{boot}	5.0	4.9	5.1	5.3	4.9	5.2	5.0	5.6
	LR_{boot}^*	4.9	4.9	5.0	5.2	4.8	5.3	4.8	5.6
	S_r	4.9	5.2	5.2	5.7	5.1	5.0	6.2	5.8
1%	LR	2.9	4.0	4.1	5.6	2.9	3.9	4.0	4.6
	LR^*	1.4	1.8	1.7	2.2	1.5	1.1	2.0	2.5
	LR_m	0.9	0.8	0.9	1.0	0.8	0.7	0.7	0.6
	LR_m^*	1.0	0.8	1.0	1.0	1.1	0.9	0.9	0.8
	LR_{boot}	1.1	1.1	1.2	1.3	0.9	1.1	1.1	1.3
	LR_{boot}^*	0.9	0.9	1.0	1.2	0.8	1.2	0.9	1.3
	S_r	1.1	0.9	0.9	1.1	1.2	1.1	2.3	1.8

The numerical results presented in Tables 2, 3, 4 and 5 show that the corrected tests outperform the uncorrected tests in small and moderate sample sizes, except for the S_r test, which performs as good as the corrected tests in the indicated scenarios. Moreover, for some cases the S_r test also outperforms the LR_m^* test, one of the best performing tests. The simulation results showed that the LR_{boot}^* and LR_m^* tests are the best performing corrected tests, followed by the S_r test. The LR^* test attenuates the oversized behavior of the LR test, but still presents distorted rejection rates, especially when the number of parameters in the models increases.

Table 6 presents the test non-null rejection rates, i.e., their power. The data generation was carried out using different values of δ. We only considered the tests LR_m, LR_m^*, LR_{boot}, LR_{boot}^* and S_r. The usual and Bartlett-corrected likelihood ratio test are not included in the power comparison since they are considerably size distorted. The results in Table 6 show what is expected, that is, the tests are more powerful as δ moves away from zero. The difference in powers is very small. For the t_5 model, we have that the LR_{boot}, LR_{boot}^* and S_r tests are slightly more powerful than the others. For

Table 6 Non-null rejection rates: t_5 model and power exponential $\kappa = 0.3$ with $n = 35$, $\alpha = 10\%$, $p = 3$ and $k = 3$

δ	t_5 model					Power exponential model				
	LR_m	LR_m^*	LR_{boot}	LR_{boot}^*	S_r	LR_m	LR_m^*	LR_{boot}	LR_{boot}^*	S_r
0.4	13.0	13.3	13.3	13.3	13.0	12.7	14.1	14.0	14.1	13.7
0.8	23.8	24.5	26.2	26.1	26.3	29.3	31.6	30.1	30.1	28.1
1.2	40.0	40.5	43.3	43.4	43.3	57.3	57.3	57.5	57.4	57.4
1.6	60.3	61.2	63.2	63.1	63.3	73.9	74.0	74.1	74.2	74.2
2.0	77.1	77.3	79.0	79.0	78.7	88.0	88.0	88.3	88.3	88.3
2.4	87.0	87.0	87.1	87.0	87.0	93.4	93.5	93.6	93.6	93.6
2.8	98.1	98.3	98.3	98.4	98.4	97.4	97.5	97.6	97.6	97.5
3.2	98.3	98.4	98.5	98.5	98.4	98.9	99.0	99.1	99.1	99.0

the power exponential model, we have that the LR_{boot} and LR_{boot}^* tests are slightly more powerful than the others.

Figures 1a–b show the relative quantile discrepancies of the test statistics against the corresponding asymptotic quantiles (χ^2) for the t_5 model and Fig. 1c–d for the power exponential model, considering $n = 35$, $p = 3, 5$ and $k = 3$. Relative quantile discrepancy is defined as the ratio of the difference between the exact quantile (estimated by simulation) and the asymptotic quantile by asymptotic quantile. We consider the score test and the corrected tests due to their performances. The closer the curve is to zero ordinate, the better approximated to the reference χ^2 distribution is the test statistic's null distribution. For both models, it is clear that the corrected likelihood ratio test statistic's null distribution is not well approximated by the reference χ^2 distribution in all considered scenarios. In contrast, the null distributions of the score statistic, the corrected modified profile likelihood ratio statistic and bootstrap Bartlett corrected likelihood ratio statistic are well approximated to the reference χ^2 distribution, since their quantile discrepancy curves are very close to zero ordinate, as reflected in the performance of these tests presented in Tables 2, 3, 4 and 5.

5 An illustrative example

In this section, we apply the test method presented in the previous sections to a real dataset. The computer code for computing these statistics can be requested from us by email. The data analyzed refer to the weight of eye lenses of European rabbit in Australia (Oryctolagus Cuniculus), y, in mg, and the age of the animal, x, in days, in a sample containing 71 observations. These data were analyzed by Wei (1998, Example 6.8), who verified the suspicion of two aberrant points under least squares estimation, thus indicating that the dataset supports errors with heavier tails than the normal. Cysneiros et al. (2005) also analyzed this dataset under the Student-t distribution with 4 degrees of freedom. The choice of the degrees of freedom was based on the model that presented the lowest AIC. The residual plot (see Cysneiros et al. 2005, Sect. 3.2) showed that they are not uniformly distributed around zero, giving evidence of
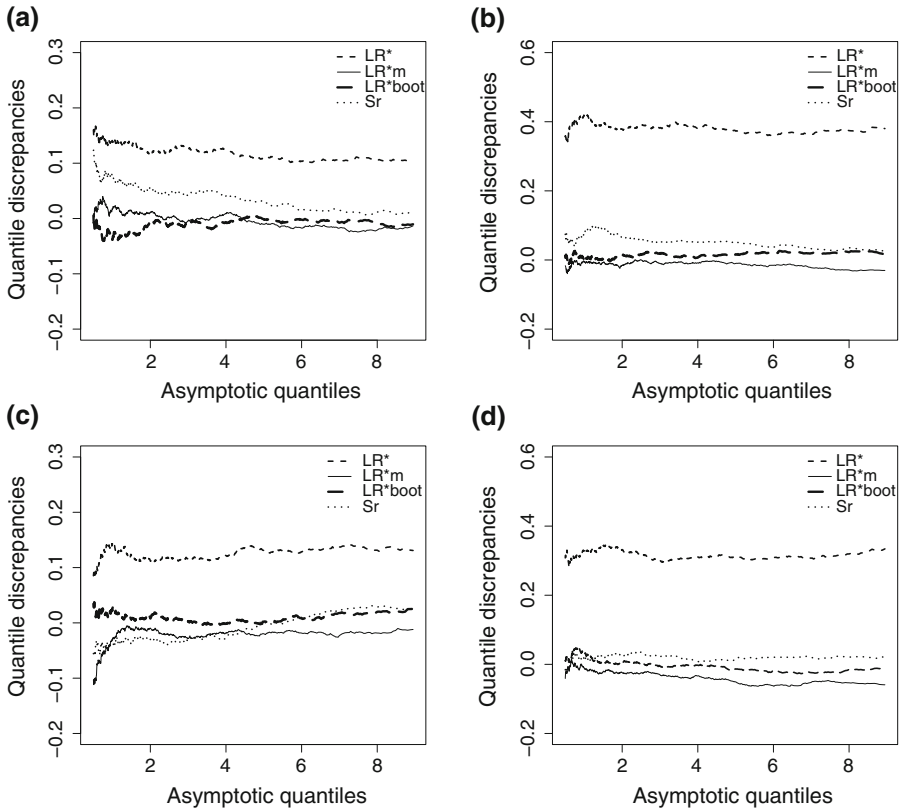
Fig. 1 Quantile relative discrepancies for t_5 and power exponential model with $p = 3$, $k = 3$, (**a**) and (**c**), and $p = 5$, $k = 3$, (**b**) and (**d**)

heteroscedasticity. Motivated by that, we decided to consider a more general model than that adjusted by Cysneiros et al. (2005), introducing a regression structure for modeling dispersion.

We consider the following heteroscedastic model

$$y_\ell = \exp\left(\beta_0 - \frac{\beta_1}{x_\ell + \beta_2}\right) e^{\epsilon_\ell},$$

where $\epsilon_\ell \sim S(0, \sigma^2 \exp\{\delta x_\ell\})$, $\ell = 1, \ldots, 71$. Our main goal now is to test $H_0 : \delta = 0$ (homoscedasticity) against a two-sided alternative (heteroscedasticity).

To test H_0, the observed values of the test statistics are $LR = 8.368$ (*p-value*: 0.004), $LR^* = 7.865$ (*p-value*: 0.005), $LR_m = 8.919$ (*p-value*: 0.003), $LR_m^* = 8.871$ (*p-value*: 0.003), $S_r = 6.766$ (*p-value*: 0.009) and $LR_{boot}^* = 5.777$ (*p-value*: 0.016). The *p-value* for the bootstrap likelihood is 0.026. So, at the 1% nominal level, the bootstrap based tests lead to not rejecting the null hypothesis, that is, the dispersion is constant over the observations (homoscedasticity), while a different decision was reached when we employed all other tests, leading to rejection of the null hypothesis,

that is, heteroscedasticity of the dispersion. Recalling our previous section and the literature, the LR and S_r tests are size distorted when we deal with small or even moderate-sized samples, leading us not to rely on the inference delivered by these tests. Also, recall from our simulation study that the bootstrap based tests performed best in most scenarios and were not affected by the sample size, number of parameters of interest or nuisance parameters, so, the bootstrap based tests should be preferable.

6 Concluding remarks

Symmetric models have received increasing attention in the statistical literature in recent years and much of this attention is due to the fact that such models are less sensitive than the normal model to the presence of outlying observations in the dataset to be modeled. Many works have addressed the class of symmetric models. Cysneiros et al. (2010a) proposed the HSNLM class, whose parametric estimation is performed by numerically maximizing the log-likelihood function, since the maximum likelihood estimators do not have closed form. In this class of models, hypothesis testing of model parameters are usually based on the likelihood ratio test, which is based on first-order asymptotic approximations. Thus, for small and even moderate sample sizes, the likelihood ratio statistic's null distribution is not well approximated from the reference χ^2 distribution and as a consequence, the test based on its statistic shows distorted rejection rates, making it important to develop strategies to yield more accurate inferences when the sample size is not large enough.

In this paper, we presented Bartlett corrections to improve hypothesis tests based on the likelihood ratio and modified profile likelihood ratio statistics in the HSNLM class. Our work extends some results presented by Cordeiro (2004) and Ferrari et al. (2004), who obtained Bartlett correction factors for the likelihood ratio statistic in symmetric nonlinear models and for the modified profile likelihood ratio statistic in normal linear models, respectively. In order to compare the performance of the tests, we also considered the score test and the bootstrap likelihood ratio and bootstrap Bartlett corrected likelihood ratio tests.

Numerical results showed that the usual likelihood ratio test is somewhat oversized and its corrected version, although attenuating this tendency, still has distorted rejection rates that increase as the number of parameters in the model increases. The simulation results also showed that the inference based on the modified profile likelihood ratio test outperforms the usual test and does not suffer influence from the number of parameters in the model when this increases. Moreover, The numerical evidence showed the better performance of the corrected tests and the uncorrected score test. The score test is very simple to compute, since it involves only estimations under the null hypothesis. In particular, the numerical results showed the superior performance of the corrected modified profile likelihood ratio test and also both bootstrap-based tests in small and moderate sample sizes. Thus, we encourage practitioners to use these tests in applications.

Appendix A

In this appendix we obtain the matrix $j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})$ expressed in (3) which denotes the observed information matrix for the nuisance parameters $(\boldsymbol{\beta}^\top, \gamma)^\top$. We can express $j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})$ by

$$j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}}) = -\begin{pmatrix} j^*_{\beta\beta} & \mathbf{0} \\ \mathbf{0} & j^*_{\gamma\gamma} \end{pmatrix},$$

where $j^*_{\beta\beta}$ is a square matrix of order p whose entries are given by

$$j^*_{\beta\beta} = -\frac{\partial^2 l^*(\boldsymbol{y}; \boldsymbol{\theta}^*)}{\partial\beta_j\partial\beta_l} = -\frac{1}{\gamma}\sum_{\ell=1}^{n} t(z_\ell)^{(2)} q_\ell(j,l)_\ell + \frac{1}{\gamma^{1/2}}\sum_{\ell=1}^{n} t(z_\ell)^{(1)} q_\ell^{1/2}(jl)_\ell$$

and

$$j^*_{\gamma\gamma} = -\frac{\partial^2 l^*(\boldsymbol{y}; \boldsymbol{\theta}^*)}{\partial\gamma^2} = -\frac{n}{2\gamma^2} - \frac{3}{4\gamma^2}\sum_{\ell=1}^{n} t(z_\ell)^{(1)} z_\ell - \frac{1}{4\gamma^2}\sum_{\ell=1}^{n} t(z_\ell)^{(2)} z_\ell^2$$

is a scalar, with $\boldsymbol{\theta}^* = (\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top, \gamma)^\top$ $q_\ell = \frac{(\prod_{s=1}^{n} m_s)^{1/n}}{m_\ell}$, $(j,l)_\ell = (\partial\mu_\ell/\partial\beta_j)\,(\partial\mu_\ell/\partial\beta_l)$ and $(jl)_\ell = \partial^2\mu_\ell/\partial\beta_j\partial\beta_l$, $j, l = 1, \ldots, p$.

Appendix B

In this appendix we present the required derivatives of the log-likelihood function in (2) up to the fourth order and their respective moments, considering $m_\ell = \exp(\boldsymbol{\omega}_\ell^\top\boldsymbol{\delta})$, to obtain the constants c and c_m from the Bartlett correction factors. For this, we introduce the following notation: $\lambda_{rs} = E(\partial^2 l/\partial\theta_r\partial\theta_s)$, $\lambda_{rst} = E(\partial^3 l/\partial\theta_r\partial\theta_s\partial\theta_t)$, etc, $a, b, c, \ldots$ index parameters of interest $\boldsymbol{\delta}$, $i, j, l, \ldots$ index nuisance parameters $\boldsymbol{\beta}$ and $r, s, t, \ldots$ index all $p + q + 1$ parameters of the model. Thus, we have

$$\lambda_{ab\gamma} = \frac{\Delta_1}{\gamma}\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab}, \quad \lambda_{acd} = \Delta_1\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{acd},$$

$$\lambda_{abcd} = \Delta_2\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{abcd},$$

$$\lambda_{\gamma\gamma} = (\Delta_4\gamma^2)^{-1}, \quad \lambda_{ija} = \frac{1}{2\gamma}(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})\sum_{\ell=1}^{n} q_{\ell ab},$$

$$\lambda_{abij} = \frac{1}{\gamma}\Delta_6\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab}q_\ell(i,j)_\ell + \frac{1}{\gamma}\Delta_7\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab}(i,j)_\ell,$$

$$\lambda_{ab\gamma\gamma} = \frac{\Delta_8}{\gamma^2} \sum_{\ell=1}^{n} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab} \quad \text{and}$$

$$\lambda_{ai} = \lambda_{a\gamma} = \lambda_{i\gamma} = \lambda_{ai\gamma} = \lambda_{abi} = \lambda_{a\gamma\gamma} = 0,$$

and the derived cumulants required to calculate c and c_m are given by:

$$(\lambda_{\gamma\gamma})_\gamma = \frac{\Delta_3}{\gamma^3}, \quad (\lambda_{ab\gamma})_\gamma = -\frac{\Delta_1}{\gamma^2} \sum_{\ell=1}^{n} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab},$$

$$(\lambda_{ija})_b = \frac{1}{2\gamma} (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{\ell=1}^{n} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab} q_l(i, j)_\ell \text{ and}$$

$$(\lambda_{ac})_d = (\lambda_{ac})_{bd} = (\lambda_{acd})_b = \lambda_{iab}^{(c)} = (\lambda_{iab})_j = \lambda_{ai}^{(bj)} = \lambda_{a\gamma\gamma}^{(b)} = \lambda_{a\gamma\gamma}^{(\gamma)} = 0.$$

Appendix C

Here we obtain in detail the expression c for the Bartlett correction factor for the LR statistic. As defined in Sect. 3.1, the Bartlett correction factor is expressed by $1 + c/k$, where

$$c = \epsilon_{p,k} - \epsilon_p.$$

From Lawley's expansion (1956), we have that $E(LR) = k + \epsilon_{p,k} - \epsilon_p + O(n^{-2})$, where

$$\epsilon_{p,k} = \sum_{\boldsymbol{\theta}^*} (l_{rstu} - l_{rstuvw}), \tag{6}$$

being

$$l_{rstu} = \lambda^{rs} \lambda^{tu} \left\{ \frac{\lambda_{rstu}}{4} - \lambda_{rst}^{(u)} + \lambda_{rt}^{(su)} \right\} \text{ and}$$

$$l_{rstuvw} = \lambda^{rs} \lambda^{tu} \lambda^{vw} \left\{ \lambda_{rtv} \left(\frac{\lambda_{suw}}{6} - \lambda_{sw}^{(u)} \right) \right.$$

$$\left. + \lambda_{rtu} \left(\frac{\lambda_{svw}}{4} - \lambda_{sw}^{(v)} + \lambda_{rt}^{(v)} \lambda_{sw}^{(u)} + \lambda_{rt}^{(u)} \lambda_{sw}^{(v)} \right) \right\},$$

where $-\lambda^{rs}$ is the (r, s) element from the inverse of Fisher's information matrix. The sum in (6) varies in all components of $\boldsymbol{\theta}^*$. The expression ϵ_p is obtained from (6) considering that the sum varies across all nuisance parameters. Thus, we can write c as

$$c = \epsilon_k(\boldsymbol{\delta}) + \epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}) + \epsilon_{p,k}(\boldsymbol{\delta}, \gamma) + \epsilon_{p,q}(\boldsymbol{\beta}, \boldsymbol{\delta}, \gamma),$$

where $\epsilon_k(\boldsymbol{\delta})$ indicates that the sum (6) was performed on all components of $\boldsymbol{\delta}$, analogously to other cases, considering the respective parameters. Replacing $\lambda's$ obtained in Appendix B in the $\epsilon's$ and after of extensive algebra, we have the constant c in (4).

Appendix D

Finally, here we obtain the constant c_m for the Bartlett correction factor for the LR_m statistic, considering Equation (5) of Ferrari et al. (2004). Thus, we have that c_m can be expressed by

$$
\begin{aligned}
c_m = {} & \frac{1}{4}\lambda^{ab}\lambda^{cd}\lambda_{abcd} - \lambda^{ab}\lambda^{cd}(\lambda_{acd})_b + \lambda^{ab}\lambda^{cd}(\lambda_{ac})_{db} - \lambda^{ij}\lambda^{ab}(\lambda_{iab})_j \\
& - \lambda^{\gamma\gamma}\lambda^{ab}(\lambda_{ab\gamma})_\gamma - \left(\frac{1}{4}\lambda^{ab}\lambda^{cd}\lambda^{ef} + \frac{1}{2}\lambda^{ab}\lambda^{cf}\lambda^{de} - \frac{1}{3}\lambda^{ab}\lambda^{cf}\lambda^{de}\right)\lambda_{acd}\lambda_{bef} \\
& + (\lambda^{ab}\lambda^{cd}\lambda^{ef} + \lambda^{ab}\lambda^{cf}\lambda^{de})\lambda_{acd}(\lambda_{be})_f - (\lambda^{ab}\lambda^{cd}\lambda^{ef} + \lambda^{ab}\lambda^{cf}\lambda^{de})(\lambda_{ac})_d(\lambda^{be})_f \\
& - \left(\frac{1}{4}\lambda^{ij}\lambda^{ab}\lambda^{cd} + \frac{1}{2}\lambda^{ij}\lambda^{ad}\lambda^{bc}\right)\lambda_{iab}\lambda_{jcd} + (\lambda^{ij}\lambda^{ab}\lambda^{kl})\lambda_{iab}(\lambda_{jk})_l \\
& - \left(\frac{1}{4}\lambda^{\gamma\gamma}\lambda^{ab}\lambda^{cd} + \frac{1}{2}\lambda^{\gamma\gamma}\lambda^{ad}\lambda^{bc}\right)\lambda_{ab\gamma}\lambda_{cd\gamma} + (\lambda^{\gamma\gamma}\lambda^{ab}\lambda^{\gamma\gamma})\lambda_{ab\gamma}(\lambda_{\gamma\gamma})_\gamma. \quad (7)
\end{aligned}
$$

Replacing the cumulant $\lambda's$ obtained in Appendix B in Equation (7) and after intense algebra, we have

$$
\begin{aligned}
c_m = {} & \frac{1}{4}\Delta_2\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a\lambda^{ab}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_c\lambda^{cd}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d \\
& - \frac{1}{4}\Delta_4\Delta_1^2\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a\lambda^{ab}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_c\lambda^{cd}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d \\
& - \frac{1}{4}\Delta_1^2\sum_{\ell=1}^{n}\sum_{s=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a\lambda^{ab}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_a(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_c\lambda^{cd}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_e\lambda^{ef}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_f \\
& - \frac{1}{6}\Delta_1^2\sum_{\ell=1}^{n}\sum_{s=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a\lambda^{ab}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_b(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_c\lambda^{cf}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_f(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d\lambda^{de}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_e \\
& + \Delta_1\Delta_4\sum_{s=1}^{n}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_a\lambda^{ab}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_b + \Delta_1\Delta_3\Delta_4^2\sum_{\ell=1}^{n}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a\lambda^{ab}(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b \\
& - \frac{1}{2}\Delta_1^2\Delta_4\sum_{\ell=1}^{n}\sum_{s=1}^{n}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_a\lambda^{ad}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_d(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b\lambda^{bc}(\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_c,
\end{aligned}
$$

where $(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_i = \omega_{\ell i} - \bar{\omega}_i$, with $i = a, b, c, d$. In matrix notation, we have

$$
\begin{aligned}
c_m = {} & \frac{1}{4}\Delta_2\text{tr}(\boldsymbol{H_d^2}) - \frac{1}{4}\Delta_1^2\Delta_4[\text{tr}(\boldsymbol{H_d})]^2 + \frac{1}{4}\Delta_1^2\boldsymbol{\iota}^\top\boldsymbol{H_d}\boldsymbol{H}\boldsymbol{H_d}\boldsymbol{\iota} + \frac{1}{6}\Delta_1^2\boldsymbol{\iota}^\top\boldsymbol{H^{(3)}}\boldsymbol{\iota} \\
& - \Delta_1\Delta_4\text{tr}(\boldsymbol{H_d}) - \Delta_1\Delta_3\Delta_4^2\text{tr}(\boldsymbol{H_d}) - \frac{1}{2}\Delta_1^2\Delta_4\boldsymbol{\iota}^\top\boldsymbol{H^{(2)}}\boldsymbol{\iota}.
\end{aligned}
$$

References

Bartlett MS (1937) Properties of sufficiency and statistical tests. Proc R Soc Lond A 160:268–282

Bayer FM, Cribari-Neto F (2013) Bartlett corrections in beta regression models. J Stat Plan Inference 143:531–547

Cook D, Weisberg S (1983) Diagnostics for heteroscedasticity in regression. Biometrika 70:1–10

Cordeiro GM (2004) Corrected likelihood ratio tests in symmetric nonlinear regression models. J Stat Comput Simul 74:609–620

Cordeiro GM, Cribari-Neto F (2014) An introduction to Bartlett correction and bias correction. Springer, New York

Cordeiro GM, Ferrari SLP, Uribe-Opazo MA, Vasconcellos KLP (2000) Corrected maximum-likelihood estimation in a class of symmeytric nonlinear regression models. Stat Probab Lett 46:317–328

Cox DR, Reid N (1987) Parameter orthogonality and approximate conditional inference. J R Stat Soc B 49:1–39

Cox DR, Reid N (1992) A note on the difference between profile and modified profile likelihood. Biometrika 79:408–411

Cribari-Neto F, Queiroz MPF (2014) On testing inference in beta regression. J Stat Comput Simul 84:186–203

Cysneiros FJA, Cordeiro GM, Cysneiros AHMA (2010a) Corrected maximum likelihood estimators in heteroscedastic symmetric nonlinear models. J Stat Comput Simul 80:451–461

Cysneiros AHMA, Ferrari SLP (2006) An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. Stat Prob Lett 76:255–265

Cysneiros FJA, Paula GA, Galea M (2005) Modelos Simétricos Aplicados. 9 Escola de Modelos de Regressão, São Paulo

Cysneiros AHMA, Rodrigues KSP, Cordeiro GM, Ferrari SLP (2010b) Three Bartlett-type corrections for score statistics in symmetric nonlinear regression models. Stat Pap 51:273–284

DiCiccio TJ, Stern SE (1994) Frequentist and Bayesian Bartlett correction of test statistics based on adjusted profile likelihoods. J R Stat Soc Ser B Methodol 56:397–408

Doornik JA (2006) An object-oriented matrix language Ox 5. Timberlake Consultants Press, London

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26

Fang KT, Anderson TW (1990b) Statistical inference in elliptical contoured and related distribuitions. Allerton Press, New York

Fang KT, Kotz S, Ng KW (1990a) Symmetric multivariate and related distributions. Chapman and Hall, London

Ferrari SLP, Cysneiros AHMA, Cribari-Neto F (2004) An improved test for heteroskedasticity using adjusted modified profile likelihood inference. J Stat Plan Inference 124:423–437

Ferrari SLP, Lucambio F, Cribari-Neto F (2005) Improved profile likelihood inference. J Stat Plan Inference 134:373–391

Fujita A, Kojima K, Patriota AG, Sato JR, Severino P, Miyano S (2010) A fast and robust test based on likelihood ratio with Bartlett correction to identify granger causality between gene sets. Bioinformatics 26:2349–2351

Lawley DN (1956) A general method for approximating to the distribution of the likelihood ratio criteria. Biometrika 71:233–244

Lemonte AJ, Cordeiro GM, Moreno G (2012) Bartlett corrections in Birnbaum-Saunders nonlinear regression models. J Stat Comput Simul 82:927–935

Lin JG, Zhu LX, Xie FG (2009) Heteroscasticity diagnostics for t linear regression models. Metrika 70:59–77

Loose LH, Bayer FM, Liberal TL (2016) Bootstrap Bartlett corretion in inflated beta regression. Commun Stat Simul 46(4):2865–2879. doi:10.1080/03610918.2015.1065326

Maior VQS, Cysneiros FJA (2016) SYMARMA: a new dynamic model for temporal data on conditional symmetric distribution. Stat Pap. doi:10.1007/s00362-016-0753-z

Melo TFN, Ferrari SLP, Cribari-Neto F (2009) Improved test inference in mixed linear models. Comp Stat Data Anal 53:2573–2582

Nocedal J, Wright SJ (1999) Numerical optimization. Springer, New York

Pace L, Salvan A (1997) Principles of statistical inference from a Neo-Fisherian perspective. World Scientific, Singapore

Rocke DM (1989) Bootstrap Bartlett adjustment in seemingly unrelated regression. J Am Stat Assoc 84:598–601

Simonoff JS, Tsai CH (1994) Use of modified profile likelihood for improved tests of constancy of variance in regression. Appl Stat 43:357–370

Stein MC, Silva MF, Duczmal LH (2014) Alternatives to the usual likelihood ratio test in mixed linear models. Comp Stat Data Anal 69:184–197

Uribe-Opazo MA, Ferrari SLP, Cordeiro GM (2008) Improved score test in symmetric linear regression model. Commun Stat Theory Methods 37:261–276

Vanegas LH, Rondon LM, Cysneiros FJA (2013) Assessing robustness of inference in symmetrical nonlinear regression models. Commun Stat Theory Methods 42:1692–1711

Wei BL (1998) Exponential family nonlinear models. Springer, Singapore