RT-MAE 9916

BARTLETT-CORRECTED TESTS FOR NORMAL LINEAR MODELS WHEN THE ERROR COVARIANCE MATRIX IS NONSCALAR

by

Elisete C. Q. Aubin
and
Gauss M. Cordeiro

Palavras-Chave: Bartlett correction; heteroskedastic model; likelihood ratio test; maximum likelihood estimate; normal linear model.

Classificação AMS: 62E20, 62F05
(AMS Classification)

- Maio de 1999 -

Bartlett-corrected tests for normal linear models when the error covariance matrix is nonscalar

ELISETE C. Q. AUBIN
Departamento de Estatística, Universidade de São Paulo
Caixa Postal 66281, Agência Cidade de São Paulo
São Paulo SP, 05389-970, Brazil

GAUSS M. CORDEIRO
Departamento de Engenharia Mecânica
CTG - Universidade Federal de Pernambuco
Cidade Universitária, Recife PE, 50740-540, Brazil

Abstract

This paper provides Bartlett corrections to improve likelihood ratio tests for heteroskedastic normal linear models when the error covariance matrix is nonscalar and depends on a set of unknown parameters. The Bartlett corrections are simple enough to be used algebraically to obtain several closed-form expressions in special cases. The corrections have also advantages for numerical purposes because they involve only simple operations on matrices and vectors.

Key words: Bartlett correction; heteroskedastic model; likelihood ratio test; maximum likelihood estimate; normal linear model.

1 Introduction

We consider a normal heteroskedastic linear model

$$y = X\beta + u, \tag{1}$$

where y is an n-dimensional column vector of observed random variables, X is an $n \times p$ matrix of fixed known regressors, β is a p-dimensional column vector of unknown regression coefficients, and u is an n-dimensional column vector of unobserved errors. The random error vector u is distributed normally with zero mean vector and a nonsingular $n \times n$ covariance matrix V^{-1}. It will be more convenient to work with the $n \times n$ precision matrix V. The elements of $V = V(\gamma)$ are known smooth functions of an unknown q-dimensional parameter vector γ. Thus, we have $(p + q)$ parameters for the simultaneous modelling of the mean vector and the covariance structure. The vectors β and γ are unrelated and can vary independently.

The class of models (1) includes many of the important models of autocorrelation and heteroskedasticity discussed in the literature as, for instance, general ARMA models and multiplicative heteroskedastic regression models (Harvey, 1976). Expressions for V or V^{-1} for ARMA models are given in Shaman (1973), Galbraith and Galbraith (1974) and Ljung and Box (1979). For the multiplicative heteroskedastic model, V^{-1} is a diagonal matrix with diagonal entries $\sigma_1^2, \ldots, \sigma_n^2$, where $\sigma_i^2 = \exp(w_i^\top \gamma)$, $w_i^\top$ being a $1 \times q$ vector of exogenous variables. This model is quite useful in many fields, including engineering, economics, and the biological and physical sciences.

A well developed exact statistical theory exists for hypothesis testing in the normal linear model when the errors are independent and homoskedastic. However, there is no general exact analysis for model (1) where the error covariance matrix is nonscalar and depends on a set of unknown parameters. Test statistics having an asymptotic chi-squared distribution are commonly used for testing hypotheses on the components of the vectors β and γ. Although the chi-squared distribution is known to approximate well the null distribution function of such test statistics for large sample sizes, there is no guarantee that it provides a reliable approximation in samples of small to moderate sizes. It has even been shown that asymptotically equivalent tests can deliver conflicting inference when applied to the same data set (Berndt and Savin, 1977; Evans and Savin, 1982). It is thus important to obtain corrected tests with better finite-sample behavior.

The likelihood ratio (LR) statistic w is frequently used to test hypotheses of interest in regression models. Under the null hypothesis H_0, w has an approximate χ^2 distribution with degrees of freedom given by the difference of the dimensions of the parameter spaces under the alternative and null hypotheses. Generally speaking, the main difficulty of testing a null hypothesis using the LR criterion lies not so much in deriving its closed-form expression – when it has one – but in finding its exact distribution, or at least a good approximation for it, when the null hypothesis is true. In an influential paper, Bartlett (1937) proposed an improved LR statistic. His argument goes as follows. Suppose that under the null hypothesis $E(w) = q\{1 + b + O(n^{-3/2})\}$, where b is a constant that can be consistently estimated under H_0. Then, the expected value of the transformed statistic $w^* = w/(1+b)$ is closer to the one from a χ^2_q distribution than is the expected value of w. This became widely known as the *Bartlett correction.* Furthermore, it was shown by Lawley (1956) that all cumulants of w^* agree with those of the reference chi-squared distribution with error of order n^{-2}. Lawley's results appeared to be incompatible with the asymptotic expansion of the null distribution of the LR statistic obtained by Hayakawa (1977). This puzzle was apparently solved by Cordeiro (1987) who showed that a certain coefficient in Hayakawa's expansion vanishes. However, this puzzle was recently revisited by Chesher and Smith (1995) who claimed that Cordeiro's proof is not correct and proved, using other arguments, that such coefficient is always zero.

The purpose of this paper is to show how to obtain Bartlett corrections that can be directly applied to commonly used LR statistics for testing hypotheses on the components of the vectors β and γ in model (1). In particular, Cordeiro's (1993a) results on Bartlett corrections for multiplicative heteroskedastic models are a special case of this paper.

In recent years there has been a renewed interest in Bartlett corrections. Cordeiro (1983, 1987) derived closed-form Bartlett corrections in generalized linear models (Nelder and Wedderburn, 1972) and discussed improved LR tests. Bartlett corrections for models defined by any one-parameter distribution in which the mean is a known function of a linear combination of unknown parameters were obtained by Cordeiro (1985), who generalized his own results of 1983. Cordeiro (1995) presented extensive simulation results on the performance of a Bartlett-corrected deviance in generalized linear models focusing on gamma and log-linear models. Attfield (1995) focused on models that involve systems of equations, and derived Bartlett corrections to the LR statistic in this case. Further Bartlett adjustments for ten multivariate normal testing problems concerning structured covariance matrices were derived by Møller (1986). Many recent papers have focused on deriving closed-form Bartlett corrections for specific regression problems. For example, Moulton, Weissfeld and St. Laurent (1993) obtained Bartlett corrections for logistic regressions; Attfield (1991) and Cordeiro (1993a) showed how to correct LR tests for heteroskedasticity; Wong (1991) obtained a Bartlett correction for testing several slopes in regression models whose independent variables are subject to error; Wang (1994) derived a Bartlett correction for testing the equality of normal variances against an increasing alternative; Cordeiro, Paula and Botter (1994) derived corrections for the class of dispersion models proposed by Jørgensen (1987); and Chesher and Smith (1997) obtained Bartlett corrections for LR specification tests. A correction to the LR statistic in regression models with Student-t errors was obtained by Ferrari and Arellano-Valle (1996), and similar corrections to heteroskedastic linear models and multivariate regression were obtained by Cribari-Neto and Ferrari (1995) and Cribari-Neto and Zarkos (1995), respectively. Furthermore, Bartlett corrections for general models were discussed by a few authors. An algorithm for computing Bartlett corrections in general statistical models was given by Jensen (1993); see also Andrews and Stafford (1993) and Stafford and Andrews (1993). General matrix formulae for computing Bartlett corrections were developed by Cordeiro (1993b).

The remainder of the paper is organized as follows. Section 2 presents a general formula for the expected LR statistic in model (1). This formula has advantage for numerical purposes because it only requires simple operations on matrices and vectors. It is also simple enough to obtain several closed-form Bartlett corrections in a variety of important tests. This formula generalizes Cordeiro's (1993a) equations (3.6) – (3.8). In Section 3, we discuss some special tests of interest in practical applications, such as those tests involving both mean and precision parameters, those ones for testing mean effects and precision effects separately and hypothesis tests for β and γ scalars. Finally, in Section 4, we present some simulation results which show that the Bartlett corrections derived here work well in small samples.

2 General formula for the Bartlett correction

We assume that the parameter space for β is the p-dimensional Euclidean space whereas the parameter space for γ is an open set Γ in a q-dimensional Euclidean space. Further, we require that p and q are small compared to n and that $X^\top VX$ is positive definite for all γ in Γ. Let $\ell = \ell(\theta)$ be the total log-likelihood for $\theta = (\beta^\top, \gamma^\top)^\top$, the $(p+q)$ vector of unknown parameters, given the observable data y. We then have

$$\ell(\theta) = \frac{1}{2}\log|V| - \frac{1}{2}(y - X\beta)^\top V(y - X\beta), \qquad (2)$$

where we have dropped an irrelevant additive constant. We assume that the fuction ℓ is regular (Cox and Hinkley, 1979; Chapter 9) with respect to all β and γ derivatives up to and including those of third order. These regularity conditions are also stated in Rao (1973, p.364) and Serfling (1980, p.144). For every sample size, the elements of V are assumed to possess derivatives up to the second order everywhere in the parameter space Γ. In addition, the derivativs of ℓ must behave nicely as n tends to infinity.

Let $\widehat{\beta}$ and $\widehat{\gamma}$ be the maximum likelihood estimates (MLEs) of β and γ, respectively, and let $\widehat{\theta} = (\widehat{\beta}^\top, \widehat{\gamma}^\top)^\top$. We must assume that the estimate $\widehat{\theta}$ converges to the true parameter θ as $n \to \infty$ and that its asymptotic distribution is multivariate normal with the usual covariance matrix to the correct order. From now on we reserve subscripts r, s, t, u, v, w to denote elements of the β vector and R, S, T, U, V, W for the elements of the γ vector. We define the derivatives $V_R = \partial V/\partial\gamma_R$ and $V^R = \partial V^{-1}/\partial\gamma_R$. The maximum likelihood equations for $\widehat{\beta}$ and $\widehat{\gamma}$ can be written as

$$\widehat{\beta} = (X^\top \widehat{V} X)^{-1} X^\top \widehat{V} y$$

and

$$\mathrm{tr}(\widehat{V}\widehat{V}_R) = (y - X\widehat{\beta})^\top \widehat{V}_R (y - X\widehat{\beta}),$$

for $R = 1, \ldots, q$, where $\widehat{V} = V(\widehat{\gamma})$ and $\widehat{V}_R = V_R(\widehat{\gamma})$. The estimates $\widehat{\beta}$ and $\widehat{\gamma}$ can be calculated numerically using an iterative algorithm (see Section 4).

We adopt the following notation for any components of γ: $V_{RS} = \partial^2 V/\partial\gamma_R\partial\gamma_S$, $V^{RS} = \partial^2 V^{-1}/\partial\gamma_R\partial\gamma_S$, $\overline{V}_R = V^{-1}V_R$, $\overline{V}^R = V^R V = -\overline{V}_R$, $\overline{V}_{RS} = V^{-1}V_{RS}$, $\overline{V}^{RS} = V^{RS}V$, $\overline{V}_{RST} = V^{-1}V_{RST}$, $\overline{V}^{RST} = V^{RST}V$, etc., $m_R = \mathrm{tr}(\overline{V}_R)$, $m^R = \mathrm{tr}(\overline{V}^R) = -m_R$, $m_{RS} = \mathrm{tr}(\overline{V}_{RS})$, $m^{RS} = \mathrm{tr}(\overline{V}^{RS})$, $m_{R,S} = \mathrm{tr}(\overline{V}_R\overline{V}_S)$, $m^{R,S} = \mathrm{tr}(\overline{V}^R\overline{V}^S)$, $m_{RS,T} = \mathrm{tr}(\overline{V}_{RS}\overline{V}_T)$, $m^{RS,T} = \mathrm{tr}(\overline{V}^{RS}\overline{V}^T)$, $m_{R,S,T} = \mathrm{tr}(\overline{V}_R\overline{V}_S\overline{V}_T)$, $m^{R,S,T} = \mathrm{tr}(\overline{V}^R\overline{V}^S\overline{V}^T)$, $m_{R,S,T,U} = \mathrm{tr}(\overline{V}_R\overline{V}_S\overline{V}_T\overline{V}_U)$, $m^{R,S,T,U} = \mathrm{tr}(\overline{V}^R\overline{V}^S\overline{V}^T\overline{V}^U)$ and so on. The m's defined above satisfy certain equations which facilitate their calculations (see, Cordeiro and Klein, 1994). For example, $m_{RS} = 2m^{R,S} - m^{RS}$, $m^{RS} = 2m_{R,S} - m_{RS}$, $m_{RS,T} = m^{RS,T} - m^{R,S,T} - m^{R,T,S}$ and $m_{R,S,T} = -m^{R,S,T}$.

Furthermore, we use the standard notation for the moments of the total log-likelihood derivatives with respect to both components of β and γ, all assumed to be $O(n)$ (see, for example, McCullagh, 1984, 1987):

$$\kappa_{ij} = E\left[\frac{\partial^2 \ell}{\partial\theta_i\partial\theta_j}\right], \quad \kappa_{i,j} = E\left[\left(\frac{\partial \ell}{\partial\theta_i}\right)\left(\frac{\partial \ell}{\partial\theta_j}\right)\right],$$

$$\kappa_{ijl} = E\left[\frac{\partial^3 \ell}{\partial\theta_i\partial\theta_j\partial\theta_l}\right], \quad \kappa_{ij,l} = E\left[\left(\frac{\partial^2 \ell}{\partial\theta_i\partial\theta_j}\right)\left(\frac{\partial \ell}{\partial\theta_l}\right)\right],$$

$$\kappa_{ijlm} = E\left[\frac{\partial^4 \ell}{\partial\theta_i\partial\theta_j\partial\theta_l\partial\theta_m}\right], \quad \kappa_{ij,lm} = E\left[\left(\frac{\partial^2 \ell}{\partial\theta_i\partial\theta_j}\right)\left(\frac{\partial^2 \ell}{\partial\theta_l\partial\theta_m}\right)\right] - \kappa_{ij}\kappa_{lm},$$

$$\kappa_{i,j,lm} = E\left[\left(\frac{\partial \ell}{\partial \theta_i}\right)\left(\frac{\partial \ell}{\partial \theta_j}\right)\left(\frac{\partial^2 \ell}{\partial \theta_l \partial \theta_m}\right)\right] - \kappa_{i,j}\kappa_{lm} \ ,$$

and

$$\kappa_{i,j,l,m} = E\left[\left(\frac{\partial \ell}{\partial \theta_i}\right)\left(\frac{\partial \ell}{\partial \theta_j}\right)\left(\frac{\partial \ell}{\partial \theta_l}\right)\left(\frac{\partial \ell}{\partial \theta_m}\right)\right] - \kappa_{i,j}\kappa_{l,m} - \kappa_{i,l}\kappa_{j,m} - \kappa_{i,m}\kappa_{j,l} \ .$$

The derivatives of moments are represented by $\kappa_{ij}^{(l)} = \partial \kappa_{ij}/\partial \theta_l$ and $\kappa_{ij}^{(lm)} = \partial^2 \kappa_{ij}/\partial \theta_l \partial \theta_m$. These κ's satisfy standard regularity equations given by Lawley (1956).

The total Fisher information matrix of order $(p+q)$ for θ is $K = \{\kappa_{ij}\}$ and let $K^{-1} = \{-\kappa^{ij}\}$ be its inverse. All κ's with subscripts are assumed to be $O(n)$. Then, all κ's with superscripts are $O(n^{-1})$. Differentiating (2) and taking expectations we can find the moments of the derivatives of $\ell(\theta)$ with respect to the elements of β and γ. We can easily show that $E(-\partial^2 \ell/\partial \beta \partial \gamma^\top) = 0$, i.e., the parameters β and γ are globally orthogonal (Cox and Reid, 1987). The partition $\theta^\top = (\beta^\top, \gamma^\top)$ induces a corresponding block diagonal total information matrix $K = \text{diag}\{K_{\beta,\beta}, K_{\gamma,\gamma}\}$ with submatrices $K_{\beta,\beta} = E(-\partial^2 \ell/\partial \beta \partial \beta^\top)$ for β and $K_{\gamma,\gamma} = E(-\partial^2 \ell/\partial \gamma \partial \gamma^\top)$ for γ. The information for β is $K_{\beta,\beta} = X^\top V X$ whose (r,s)th typical element is $\kappa_{r,s} = -\kappa_{rs} = x_r^\top V x_s$, where x_s is the sth column of the matrix X. We also have $K_{\gamma,\gamma} = \{\kappa_{R,S}\}$ where $\kappa_{R,S} = -\kappa_{RS} = \frac{1}{2} m_{R,S} = \frac{1}{2}\text{tr}(\overline{V}_R \overline{V}_S)$ is the (R,S)th typical element of the information $K_{\gamma,\gamma}$ for γ.

We now need the following results:

$$\text{(i)} \quad V^R = V^{-1} V_R V^{-1} \ \text{ and } \ V_R = -V V^R V;$$

$$\text{(ii)} \quad \frac{\partial \log |V|}{\partial \gamma_R} = \text{tr}(\overline{V}_R) = m_R;$$

$$\text{(iii)} \quad E\{(y-\mu)^\top A (y-\mu)\} = \text{tr}(V^{-1} A),$$

where $E(y) = \mu$ and $\text{Cov}(y) = V^{-1}$, for any positive definite matrix A.

Differentiating (2) and making use of (i) – (ii), we can find the following log likelihood derivatives

$$\frac{\partial^2 \ell}{\partial \beta_r \partial \gamma_S} = x_r^\top V_S (y - X\beta), \quad \frac{\partial^3 \ell}{\partial \beta_r \partial \beta_s \partial \gamma_T} = -x_r^\top V_T x_s,$$

$$\frac{\partial^3 \ell}{\partial\beta_r\partial\gamma_S\partial\gamma_T} = x_r^\top V_{ST}(y - X\beta), \quad \frac{\partial^4 \ell}{\partial\beta_r\partial\beta_s\partial\beta_t\partial\gamma_U} = 0,$$

$$\frac{\partial^4 \ell}{\partial\beta_r\partial\beta_s\partial\gamma_T\partial\gamma_U} = -x_r^\top V_{TU} x_s, \quad \frac{\partial^4 \ell}{\partial\beta_r\partial\gamma_S\partial\gamma_T\partial\gamma_U} = x_r^\top V_{STU}(y - X\beta),$$

$$\frac{\partial \ell}{\partial \gamma_R} = \frac{1}{2}\left\{m_R - \frac{1}{2}(y - X\beta)^\top V_R (y - X\beta)\right\},$$

$$\frac{\partial^2 \ell}{\partial\gamma_R\partial\gamma_S} = -\frac{1}{2}\left\{m_{R,S} - m_{RS} + (y - X\beta)^\top V_{RS}(y - X\beta)\right\},$$

$$\begin{aligned}\frac{\partial^3 \ell}{\partial\gamma_R\partial\gamma_S\partial\gamma_T} = \frac{1}{2}\{ & m_{RST} - m_{RS,T} - m_{RT,S} - m_{ST,R} + m_{R,S,T} \\ & + m_{R,T,S} - (y - X\beta)^\top V_{RST}(y - X\beta)\}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^4 \ell}{\partial\gamma_R\partial\gamma_S\partial\gamma_T\partial\gamma_U} = \frac{1}{2}\{ & m_{RSTU} - m_{RST,U} - m_{RSU,T} - m_{RTU,S} - m_{STU,R} \\ & - m_{RS,TU} - m_{RT,SU} - m_{RU,ST} + m_{RS,U,T} + m_{RS,T,U} \\ & + m_{RT,U,S} + m_{RT,S,U} + m_{ST,R,U} + m_{ST,U,R} + m_{TU,S,R} \\ & + m_{TU,R,S} + m_{SU,R,T} + m_{SU,T,R} + m_{RU,T,S} + m_{RU,S,T} \\ & - m_{S,R,T,U} - m_{S,R,U,T} - m_{R,T,S,U} - m_{R,U,S,T} - m_{R,S,U,T} \\ & - m_{R,S,T,U} + (y - X\beta)^\top V_{RSTU}(y - X\beta)\}.\end{aligned}$$

From the above log likelihood derivatives and using (iii) we obtain the following moments

$$\kappa_{RS} = -\frac{1}{2} m_{R,S},$$

$$\kappa_{Rst} = -x_s^\top V_R x_t,$$

$$\kappa_{RST} = -\frac{1}{2}(m_{RS,T} + m_{ST,R} + m_{RT,S} - m_{R,S,T} - m_{R,T,S}),$$

$$\begin{aligned}\kappa_{RSTU} = -\frac{1}{2}\Big(& m_{RST,U} + m_{RSU,T} + m_{RTU,S} + m_{STU,R} + m_{RS,TU} + m_{RT,SU} + m_{RU,ST} \\ & - m_{RS,T,U} - m_{RS,U,T} - m_{RT,S,U} - m_{RT,U,S} - m_{ST,R,U} - m_{ST,U,R} - m_{SU,R,T} \\ & - m_{SU,T,R} - m_{RU,T,S} - m_{RU,S,T} - m_{TU,S,R} - m_{TU,R,S} + m_{R,S,T,U} + m_{R,S,U,T} \\ & + m_{R,U,S,T} + m_{R,U,T,S} + m_{R,T,S,U} + m_{R,T,U,S}\Big),\end{aligned}$$

$$\kappa_{rs}^{(T)} = -x_r^\top V_T x_s \ ,$$

$$\kappa_{RS}^{(T)} = -\frac{1}{2}(m_{ST,R} + m_{RT,S} - m_{R,S,T} - m_{R,T,S}) \ ,$$

$$\kappa_{rs}^{(TU)} = -x_r^\top V_{TU} x_s \ ,$$

$$\begin{aligned} \kappa_{RS}^{(TU)} = & -\frac{1}{2}\Big(m_{STU,R} + m_{RTU,S} + m_{RU,ST} + m_{RT,SU} - m_{ST,R,U} - m_{ST,U,R} - m_{RT,U,S} \\ & -m_{RT,S,U} - m_{SU,T,R} - m_{SU,R,T} - m_{RU,S,T} - m_{RU,T,S} - m_{TU,R,S} - m_{TU,S,R} \\ & +m_{R,S,T,U} + m_{R,S,U,T} + m_{R,U,S,T} + m_{R,U,T,S} + m_{R,T,S,U} + m_{R,T,U,S}\Big) \end{aligned}$$

and

$$\begin{aligned} \kappa_{RST}^{(U)} = & -\frac{1}{2}\Big(m_{RSU,T} + m_{RTU,S} + m_{STU,R} + m_{RS,TU} + m_{RT,SU} + m_{RU,ST} - m_{RS,T,U} \\ & -m_{RS,U,T} - m_{RT,S,U} - m_{RT,U,S} - m_{ST,R,U} - m_{ST,U,R} - m_{SU,T,R} - m_{SU,R,T} \\ & -m_{TU,R,S} - m_{TU,S,R} - m_{RU,S,T} - m_{RU,T,S} + m_{R,S,T,U} + m_{R,S,U,T} + m_{R,T,S,U} \\ & +m_{R,T,U,S} + m_{R,U,S,T} + m_{R,U,T,S}\Big) \ . \end{aligned}$$

From Lawley's (1956) expansion we can write the expected LR statistic to $O(n^{-1})$ as $2E\{\ell(\widehat{\beta},\widehat{\gamma}) - \ell(\beta,\gamma)\} = p + q + \varepsilon_{p+q}$, where $l(\beta,\gamma)$ is the log-likelihood at the true parameter point and ε_{p+q} is a term of order n^{-1} evaluated at the true parameter point given by

$$\varepsilon_{p+q} = \sum_{\beta,\gamma}(l_{i_1i_2i_3i_4} - l_{i_1i_2i_3i_4i_5i_6}), \tag{3}$$

where all indices $i_1, \ldots, i_6$ vary over both vectors β and γ and $\sum_{\beta,\gamma}$ denotes the summation over all the combinations of the $p+q$ parameters $\beta_1, \ldots, \beta_p$, $\gamma_1, \ldots, \gamma_q$. The notation on the right hand side of (3) follows Cordeiro (1987), the l's being obtained from

$$l_{i_1i_2i_3i_4} = \kappa^{i_1i_2}\kappa^{i_3i_4}\left(\frac{\kappa_{i_1i_2i_3i_4}}{4} - \kappa_{i_1i_2i_3}^{(i_4)} + \kappa_{i_1i_3}^{(i_2i_4)}\right) \tag{4}$$

and

$$\begin{aligned} l_{i_1i_2i_3i_4i_5i_6} = & \kappa^{i_1i_2}\kappa^{i_3i_4}\kappa^{i_5i_6}\Big\{\kappa_{i_1i_3i_5}\left(\frac{\kappa_{i_2i_4i_6}}{6} - \kappa_{i_2i_6}^{(i_4)}\right) \\ & +\kappa_{i_1i_3i_4}\left(\frac{\kappa_{i_2i_5i_6}}{4} - \kappa_{i_2i_6}^{(i_5)}\right) + \kappa_{i_1i_3}^{(i_5)}\kappa_{i_2i_6}^{(i_4)} + \kappa_{i_1i_3}^{(i_4)}\kappa_{i_2i_6}^{(i_5)}\Big\}. \end{aligned} \tag{5}$$

The proof of (3) – (5) given in Lawley (1956) contains many references to the difficulty of the required symbolic manipulations and as he pointed out, it involves exceedingly complicated and laborious algebra. Lawley's formula involves certain products of higher-order arrays called tensors but has no simple closed-form expression, since it is not explicitly written in terms of the unknown parameters.

The important simplification for deriving a general formula for ε_{p+q} from (3) in matrix notation is the block diagonality of the information matrix K. In fact, several of the mixed cumulants vanish due to the orthogonality between β and γ. Further, the number of terms necessary to compute (3) is greatly reduced because of this orthogonality.

Henceforth, any matrix with (i,j) element a_{ij} will be represented by $A = \{a_{ij}\}$ and quantities that can be expressed solely in terms of the matrix V and its first two derivatives with respect to γ will be denoted with suffix γ, whereas those that also depend on the model matrix X are used with double suffices γ, β. In order to find the general formula for Bartlett corrections in model (1) we need the following matrices and vectors. Let $T_\gamma^{(R)} = \{t_{\gamma ST}^{(R)}\}$ and $U_\gamma^{(R)} = \{u_{\gamma ST}^{(R)}\}$ be $q \times q$ matrices referring to the Rth component of γ with typical (S,T) elements given by $t_{\gamma ST}^{(R)} = m_{R,ST}$ and $u_{\gamma ST}^{(R)} = m_{R,S,T}$, respectively. We also define the $q \times q$ matrices $S_\gamma = \{s_{\gamma RS}\}$, $A_\gamma = \{a_{\gamma RS}\}$, $P_{\gamma,\beta} = \{p_{\gamma,\beta RS}\}$ and $M_{\gamma,\beta} = \{m_{\gamma,\beta RS}\}$ whose typical (R,S) elements are $s_{\gamma RS} = \mathrm{tr}(K_{\gamma,\gamma}^{-1}U_\gamma^{(R)})\mathrm{tr}(K_{\gamma,\gamma}^{-1}U_\gamma^{(S)})$, $a_{\gamma RS} = 6\mathrm{tr}(K_{\gamma,\gamma}^{-1}T_\gamma^{(R)}K_{\gamma,\gamma}^{-1}T_\gamma^{(S)}) - 3\mathrm{tr}(K_{\gamma,\gamma}^{-1}T_\gamma^{(R)})\mathrm{tr}(K_{\gamma,\gamma}^{-1}\{T_\gamma^{(S)} - 4U_\gamma^{(S)}\}) - 4\mathrm{tr}(K_{\gamma,\gamma}^{-1}U_\gamma^{(R)T}K_{\gamma,\gamma}^{-1}\{U_\gamma^{(S)} + U_\gamma^{(S)T}\})$, $p_{\gamma,\beta RS} = \mathrm{tr}(X^T V_R X K_{\beta,\beta}^{-1} X^T V_S X K_{\beta,\beta}^{-1})$ and $m_{\gamma,\beta RS} = \mathrm{tr}(K_{\beta,\beta}^{-1}X^T V_{RS} X)$, respectively. Further, for any two Rth and Sth components of γ, we define the following $q \times q$ matrices $D_\gamma^{(RS)} = \{d_{\gamma TU}^{(RS)}\}$ where their (T,U) elements are $d_{\gamma TU}^{(RS)} = 2m_{RT,SU} - m_{RS,TU} + 4m_{RS,T,U} - 4m_{R,S,T,U} - 2m_{R,T,S,U}$. From these matrices we can easily construct the $q \times q$ matrix $N_\gamma = \{n_{\gamma RS}\}$ whose (R,S) element is given by $n_{\gamma RS} = \mathrm{tr}(K_{\gamma,\gamma}^{-1}D_\gamma^{(SR)})$. Finally, we define the $q \times 1$ vectors $\tau_{\gamma,\beta}$ and $\rho_{\gamma,\beta}$ whose Rth components are $\mathrm{tr}(K_{\beta,\beta}^{-1}X^T V_R X)$ and $2\mathrm{tr}(K_{\beta,\beta}^{-1}X^T V_R X - K_{\gamma,\gamma}^{-1}\{T_\gamma^{(R)} - 2U_\gamma^{(R)}\})$, respectively.

On inserting the moments κ's given before in (4) – (5) and then summing over the sample after evaluating the sums over the parameters, it can be shown from the matrices and vectors defined above that

$$\varepsilon_{p+q} = B_\gamma + C_{\gamma,\beta}, \tag{6}$$

where

$$B_\gamma = \frac{1}{8}\text{vec}(K_{\gamma,\gamma}^{-1})^\top \text{vec}(N_\gamma) + \frac{1}{48}\text{tr}\{K_{\gamma,\gamma}^{-1}(12S_\gamma - A_\gamma)\} \tag{7}$$

and

$$C_{\gamma,\beta} = \frac{1}{2}\text{vec}(K_{\gamma,\gamma}^{-1})\text{vec}(N_{\gamma,\beta}) - \frac{1}{2}\text{tr}(K_{\gamma,\gamma}^{-1}P_{\gamma,\beta}) + \frac{1}{8}\tau_{\gamma,\beta}^\top K_{\gamma,\gamma}^{-1}\rho_{\gamma,\beta}. \tag{8}$$

The details involved in the derivation of the equations (6) – (8) are tedious and are omitted here to save space but they follow from similar algebraic developments of Cordeiro (1983, 1987, 1993a), Cordeiro, Paula and Botter (1994), Botter and Cordeiro (1997) and Aubin and Cordeiro (1999). They can be obtained from the authors upon request. Clearly, B_γ depends only on the vector γ through the first two partial derivatives of V with respect to this vector, whereas $C_{\gamma,\beta}$ is a function of X, β and γ. Both expressions (7) and (8) involve simple operations on matrices and vectors. Expressing ε_{p+q} in matrix formulae has great advantages to obtain closed-form Bartlett corrections for several special tests of practical use such as those that will be discussed in Section 3. Also, the quantities B_γ and $C_{\gamma,\beta}$ can be made computationally atractive using a computer algebra system such as MATHEMATICA or MAPLE, or using a programming language with support for matrix operations such as GAUSS, Ox or S-PLUS. Although these quantities are easy to compute under the null hypothesis because they involve only simple operations on matrices and vectors, they are not easy to interpret. The fundamental difficulty is that the individual terms in equations (7) and (8) are not invariant under reparametrization and therefore their interpretation depend on the coordinate system chosen. The entire expression for ε_{p+q} is of course invariant under reparametrization. The main advantage of formulae (6) – (8) over Lawley's equation is that we avoid computations involving all possible products of higher-order arrays. In these formulae the matrices required are given in a more readily computable form by exploiting special structures of the cumulants involved. Moreover, the matrix notation provides some insights into the nature of the Bartlett correction, especially in cases where $K_{\beta,\beta}$ and $K_{\gamma,\gamma}$ have closed-form inverse. We have checked formulae (6) – (8) in some specific situations and they work properly. Some partial checks of ε_{p+q} are provided in Section 3.

3 Some special cases

3.1 Testing both mean and precision parameters

We first consider the hypothesis where both the parameter vector of interest and the nuisance parameter vector may be regarded as being composed of some components of β and γ. Partitioning the parameters as $\beta^\top = (\beta_1^\top, \beta_2^\top)$ and $\gamma^\top = (\gamma_1^\top, \gamma_2^\top)$, where $\beta_1 = (\beta_1, \ldots, \beta_{p_1})^\top$, $\beta_2 = (\beta_{p_1+1}, \ldots, \beta_p)^\top$, $\gamma_1 = (\gamma_1, \ldots, \gamma_{q_1})^\top$ and $\gamma_2 = (\gamma_{q_1+1}, \ldots, \gamma_q)^\top$ with $p_1 \le p$ and $q_1 \le q$, we are interested in testing the null hypothesis $H_1 : \beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$ versus A_1: violation of at least one equality, where $\beta_1^{(0)}$ and $\gamma_1^{(0)}$ are specified vectors of dimensions p_1 and q_1, respectively. Following the partition induced by H_1, let $X = (X_1, X_2)$ be the corresponding partitioned model matrix, where X_1 and X_2 are respectively $n \times p_1$ and $n \times p_2$ matrices of full ranks. Let $\widehat{\beta}$ and $\widehat{\gamma}$ be the unrestricted MLEs of β and γ and $\widetilde{\beta}_2$ and $\widetilde{\gamma}_2$ be the restricted MLEs of β_2 and γ_2 under H_1.

From now on, functions evaluated at the unrestricted MLEs will be denoted by the addition of a circumflex and those evaluated at the restricted estimates by the addition of a tilde. The LR statistic for testing H_1 is simply

$$\omega_1 = 2\{\ell(\widehat{\beta}, \widehat{\gamma}) - \ell(\beta_1^{(0)}, \widetilde{\beta}_2, \gamma_1^{(0)}, \widetilde{\gamma}_2)\}$$

which is, under H_1, asymptotically distributed as $\chi^2_{p_1+q_1}$. The key to the evaluation of the Bartlett correction to improve the test of H_1 is to write the LR statistic ω_1 as the difference of two LR statistics for testing hypotheses without nuisance parameters. More specifically we have

$$\begin{aligned} E(\omega_1) &= 2E\{\ell(\widehat{\beta},\widehat{\gamma}) - \ell(\beta_1^{(0)},\beta_2,\gamma_1^{(0)},\gamma_2)\} \\ &\quad -2E\{\ell(\beta_1^{(0)},\widetilde{\beta}_2,\gamma_1^{(0)},\widetilde{\gamma}_2) - \ell(\beta_1^{(0)},\beta_2,\gamma_1^{(0)},\gamma_2)\}. \end{aligned}$$

Here β_2 and γ_2 represent the true values of these parameters. Using (6) we find under H_1

$$E(\omega_1) = p_1 + q_1 + B_\gamma + C_{\gamma,\beta} - B_{\gamma_2} - C_{\gamma_2,\beta_2},$$

where B_{γ_2} and C_{γ_2,β_2} are obtained directly from equations (7) and (8) with $K_{\gamma,\gamma}$ and $K_{\beta,\beta}$ being substituted by K_{γ_2,γ_2} and K_{β_2,β_2}, respectively. All terms in $E(\omega_1)$ are determined subject to $\beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$. Clearly, the Bartlett correction $c_1 = E(\omega_1)/(p_1+q_1)$ for improving $H_1 : \beta_1 = \beta_1^{(0)}, \gamma_1 = \gamma_1^{(0)}$ follows as

$$c_1 = 1 + (B_\gamma + C_{\gamma,\beta} - B_{\gamma_2} - C_{\gamma_2,\beta_2})/(p_1+q_1), \tag{9}$$

where the vector γ should be evaluated at $(\gamma_1^{(0)\top}, \widetilde{\gamma}_2^\top)^\top$. Under H_1, the improved statistic $\omega_1^* = \widetilde{c}_1^{-1}\omega_1$ is distributed to order n^{-1} as $\chi^2_{p_1+q_1}$. Therefore, the improved test compares ω_1^* with the upper point of the $\chi^2_{p_1+q_1}$ distribution. The importance of equation (9) in applications is that it involves only simple matrix operations. Several formulae for Bartlett corrections in special cases can be obtained by exploring models whose information matrices $K_{\beta,\beta}$, K_{β_2,β_2}, $K_{\gamma,\gamma}$ and K_{γ_2,γ_2} have closed-form inverses. Evidently, in the simplest case $p_1 = p$ and $q_1 = q$ of little practical interest since H_1 becomes simple, the Bartlett correction reduces to $c_1 = 1 + (p+q)^{-1}(B_\gamma + C_{\gamma,\beta})$ with γ being evaluated under the null hypothesis.

Expression (9) can be reduced considerably for models with special structures for the precision matrix (9). An important special case of (9) refers to a homoskedastic model where the $n \times n$ covariance matrix V^{-1} reduces to γI_n, γ a scalar parameter. In this case, $K_{\beta,\beta} = \gamma^{-1} X^\top X$, $V_\gamma = -\gamma^{-2} I_n$, $V_{\gamma\gamma} = 2\gamma^{-3} I_n$, $K_{\gamma,\gamma} = n/2\gamma^2$ and after some algebra we obtain $P_{\gamma,\beta} = p/\gamma^2$, $M_{\gamma,\beta} = 2p/\gamma^2$, $\tau_{\gamma,\beta} = -p/\gamma$, $\rho_{\gamma,\beta} = -2p/\gamma$, $D_\gamma^{(\gamma\gamma)} = 6/\gamma^4$, $U_\gamma^{(\gamma)} = -n/\gamma^3$, $T_\gamma^{(\gamma)} = -2n/\gamma^3$, $N_\gamma = 12/\gamma^2$, $A_\gamma = 112/\gamma^2$ and $S_\gamma = 4/\gamma^2$. Finally, we find from equations (7) and (8)

$$B_\gamma = \frac{1}{3n}, \quad C_{\gamma,\beta} = \frac{p(p+2)}{2n}. \tag{10}$$

Note that $C_{\gamma,\beta}$ in (10) depends on the model matrix only through its rank p. Expressions (10) provide a partial check of equations (6) – (8) since, in this case, the Bartlett correction to improve the test of $H_1 : \beta = \beta^{(0)}$, $\gamma = \gamma^{(0)}$ comes directly from the exact expected value of the LR statistic given by $2E\{\ell(\widehat{\beta}, \widehat{\gamma}) - \ell(\beta, \gamma)\} = \log n - E(\log \chi^2_{n-p})$ if terms of order n^{-2} are neglected.

Unfortunately, when β is a scalar parameter and V is an arbitrary $n \times n$ precision matrix depending on the q-dimensional parameter vector γ, there is no substantial reduction in the expressions for β_γ and $C_{\gamma,\beta}$.

3.2 Testing mean effects

We are now interested in testing a subset of parameters only in β. In this situation, the null hypothesis is $H_2 : \beta_1 = \beta_1^{(0)}$ to be tested against $A_2 : \beta_1 \neq \beta_1^{(0)}$, where $\beta_1^{(0)}$ is a specified vector of dimension p_1 and $\beta_2(p_2 \times 1)$ and $\gamma(q \times 1)$ are the vectors of nuisance parameters. The LR statistic for testing H_2 is given by $\omega_2 = 2\{\ell(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\gamma}) - \ell(\beta_1^{(0)}, \widetilde{\beta}_2, \widetilde{\gamma})\}$, which is, under H_2, distributed to first order as $\chi^2_{p_1}$.

We can easily show that

$$E(\omega_2) = p + q + B_\gamma + C_{\gamma,\beta} - (p_2 + q + B_\gamma + C_{\gamma,\beta_2}) = p_1 + C_{\gamma,\beta} - C_{\gamma,\beta_2},$$

where C_{γ,β_2} comes directly from (8) with $K_{\beta_2,\beta_2} = X_2^\top V X_2$ in place of $K_{\beta,\beta} = X^\top V X$. The Bartlett correction associated with ω_2 is $c_2 = 1 + (C_{\gamma,\beta} - C_{\gamma,\beta_2})/p_1$, where $C_{\gamma,\beta}$ and C_{γ,β_2}

are both evaluated at $(\beta_1^{(0)}, \widetilde{\beta}_2^{\top})^{\top}$ and $\widetilde{\gamma}$. Under H_2, the improved statistic $\omega_2^* = \widetilde{c}_2^{-1}\omega_2$ is distributed to order n^{-1} as $\chi^2_{p_1}$. An important application corresponds to the test of homogeneity of means assuming that the model (1) is homoskedastic. We are interested in testing $H_2 : \beta_1 = \cdots = \beta_{p_1} = 0$ with $p_1 = p - 1$ and $x_{lp} = 1$ for $l = 1, \ldots, n$ against $A_2 : H_2$ is not true. The Bartlett correction reduces to $c_2 = 1 + (C_{\gamma,\beta} - C_{\gamma,\beta_p})/(p-1)$, where β_p is a scalar parameter. If the covariance matrix V^{-1} is γI_n, $C_{\gamma,\beta}$ follows from (10) and C_{γ,β_p} is also obtained from (10) by making $p = 1$. The Bartlett correction becomes $c_2 = 1 + (p^2 + 2p - 3)/\{2n(p-1)\}$ which coincides with Cordeiro's (1993a) expression (4.3), thus providing another check of equations (6) – (8).

If the null hypothesis specifies the whole vector β, i.e. $H_2 : \beta = \beta^{(0)}$, the Bartlett correction c_2 for improving the LR test of H_2 is given by $c_2 = 1 + C_{\gamma,\beta}/p$. Moreover, if the model is homoskedastic, $C_{\gamma,\beta}$ comes from (10) and the Bartlett correction reduces to a linear function $c_2 = 1 + (p+2)/2n$ of the dimension of the vector β.

3.3 Testing precision effects

We now consider the LR statistic for testing $H_3 : \gamma_1 = \gamma_1^{(0)}$ against $A : \gamma_1 \neq \gamma_1^{(0)}$, where $\gamma_1^{(0)}$ is a specific vector of dimension q_1 and β and γ_2 are vectors of nuisance parameters. It is given by

$$\omega_3 = 2\{\ell(\widehat{\beta}, \widehat{\gamma}_1, \widehat{\gamma}_2) - \ell(\widetilde{\beta}, \gamma_1^{(0)}, \widetilde{\gamma}_2)\},$$

and, under H_3, the distribution of ω_3 is generally of order n^{-1} away from $\chi^2_{q_1}$. We obtain

$$E(\omega_3) = q_1 + B_\gamma + C_{\gamma,\beta} - B_{\gamma_2} - C_{\gamma_2,\beta}.$$

The terms B_{γ_2} and $C_{\gamma_2,\beta}$ come from equations (7) and (8) by substituting $K_{\gamma,\gamma}$ by K_{γ_2,γ_2} subject to $\gamma_1 = \gamma_1^{(0)}$. The Bartlett correction determined by $c_3 = E(\omega_3)/q_1$ renders the n^{-1} term in $E(\omega_3)$ equal to zero and the error of the $\chi^2_{q_1}$ approximation to the distribution of $\omega_3^* = \widetilde{c}_3^{-1}\omega_3$ becomes of order n^{-2}, the nuisance parameters γ_2 and β being evaluated at

$\widetilde{\gamma}_2$ and $\widetilde{\beta}$. If the null hypothesis specifies all components of the vector γ, i.e., $H_3 : \gamma = \gamma^{(0)}$, the Bartlett correction for improving the test of $H_3 : \gamma = \gamma^{(0)}$ based on the LR statistic $\omega_3 = 2\{\ell(\widehat{\beta}, \widehat{\gamma}) - \ell(\widetilde{\beta}, \gamma^{(0)}\}$ is given by $c_3 = 1 + (B_\gamma + C_{\gamma,\beta})/q$ with $\gamma = \gamma^{(0)}$. If the model is homoskedastic, the correction reduces to $c_3 = 1 + \{3p(p+2) + 2\}/6n$. This result is the classical Bartlett correction for testing the variance in a homoskedastic normal linear model.

3.4 Both β and γ scalars

When both β and γ are scalar parameters, i.e. $p = q = 1$, the model (1) is homoskedastic with $n \times n$ covariance matrix $V^{-1} = \gamma I_n$ and β is the common mean, the information matrix for $\theta = (\beta, \gamma)^\top$ reduces to $K = \text{diag}\left\{\frac{n}{\gamma}, \frac{n}{2\gamma^2}\right\}$. For testing the mean $H_2 : \beta = \beta^{(0)}$ against $A_2 : \beta \neq \beta^{(0)}$, the LR statistic is given by $w_2 = n \log\left\{\frac{\sum(y_i - \beta^{(0)})^2}{\sum(y_i - \overline{y})^2}\right\}$, where $\overline{y}$ is the mean of the observations. The Bartlett correction is simply $c_2 = 1 + C_{\beta,\gamma}$, where $C_{\beta,\gamma}$ comes from the last expression of Section 3.2 with $p = 1$. We have $c_2 = 1 + 3/(2n)$ and the improved test of the mean compares $\omega_2^* = \left(1 + \frac{3}{2n}\right)^{-1} \omega_2$ with the χ_1^2 distribution.

We can also consider the test of the common variance $H_3 : \gamma = \gamma^{(0)}$ against $A_3 : \gamma \neq \gamma^{(0)}$ in which the LR statistic is given by

$$\omega_3 = n\left\{\log\left(\frac{\gamma^{(0)}}{\widehat{\gamma}}\right) + \frac{\widehat{\gamma} - \gamma^{(0)}}{\gamma^{(0)}}\right\},$$

where $\widehat{\gamma} = \sum(y_i - \overline{y})^2/n$. The Bartlett correction follows directly from the result of Section 3.3 with $p = 1$. Thus, $c_3 = 1 + 11/(6n)$ and the improved statistic $\omega_3^* = \left(1 + \frac{11}{6n}\right)^{-1} \omega_3$ has, under H_3, a χ_1^2 distribution to order n^{-1}. The results presented here are in agreement with the classical Bartlett corrections to improve the LR tests of the mean and variance of a normal distribution which can be obtained directly by Taylor series expansion.

4 Simulation results

We now present some simulation results for two LR statistics and their corrected versions for the normal heteroskedastic linear model (1) in which the components of the vector of errors u follow the stationary AR(1) equation $u_i = \rho u_{i-1} + \varepsilon_i$, $|\rho| < 1$, with $\varepsilon_i \sim NID(0, \sigma^2)$, for $i = 1, \ldots, n$. The covariance matrix V^{-1} of the model has simple form given by

$$V = \frac{\sigma^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{pmatrix}$$

with inverse

$$V = \sigma^2 \begin{pmatrix} 1 & -\rho & 0 & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

The linear structure of the model is $X\beta = \beta_0 1_n + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$, where 1_n is an $n \times 1$ vector of ones and the covariates x_1, x_2 and x_3 are chosen, respectively, as random draws from the following distributions: $U(0,1)$, $N(\hat{0},1)$ and $F(3,5)$. Their values were held constant throughout the simulations with equal sample sizes.

We wish to test the null hypotheses $H_2 : \beta_1 = 2, \beta_2 = 1$ and $H_3 : \rho = 0.6$ against the hypothesis of violation of at least one equality. For the simulations the nuisance parameters β_0 and β_3 were fixed at $\beta_0 = 4$, $\beta_3 = 3$ and the variance σ^2 was taken as 1,4 and 16. For testing H_2, ρ was set equal to 0.4 and for testing H_3, β_1 and β_2 were taken as 3 and 2, respectively. The number of observations was set at 20, 30 and 40. The simulations run-size was 10,000 in each case and calculations were performed using the GLIM environment.

In each simulation, we generate u and calculate the MLEs of β and γ iteratively using Fisher scoring method given by

$$\beta^{(m+1)} = \beta^{(m)} + (X^\top V^{(m)} X)^{-1} X^\top V^{(m)} (X\beta^{(m)} + u),$$

$$\begin{pmatrix} \rho^{(m+1)} \\ \sigma^{2(m+1)} \end{pmatrix} = \begin{pmatrix} \rho^{(m)} \\ \sigma^{2(m)} \end{pmatrix} + K_{\gamma,\gamma}^{(m)^{-1}} U_{\gamma}^{(m)},$$

where U_γ and $K_{\gamma,\gamma}$ are the 2×1 score vector and the 2×2 information matrix for $\gamma = (\rho \ \sigma^2)$ given by

$$U_\gamma = \frac{1}{4}\begin{pmatrix} \frac{-4\rho}{1-\rho^2} - u^\top U_\rho u \\ \frac{-2}{\sigma^2}\left(n - 2 + \frac{2}{1-\rho^2}\right) - u^\top V_{\sigma^2} u \end{pmatrix}$$

and

$$K_{\gamma,\gamma} = \begin{pmatrix} \frac{n(1-\rho^2)+3\rho^2-1}{(1-\rho^2)^2} & \frac{-\rho}{\sigma^2(1-\rho^2)} \\ \frac{-\rho}{\sigma^2(1-\rho^2)} & \frac{n}{2\sigma^4} \end{pmatrix},$$

respectively. Then we obtain the LR statistics ω_2 and ω_3 for testing H_2 and H_3 and the modified statistics $\omega_2^* = \widetilde{c_2}^{-1}\omega_2$ and $\omega_3^* = \widetilde{c_3}^{-1}\omega_3$, where the Bartlett corrections c_2 and c_3 are evaluated under H_2 and H_3, respectively. Further, we investigate the rejection rates of ω_2 and ω_2^* and ω_3 and ω_3^* at the nominal 10% and 5% levels of the references χ_2^2 and χ_1^2 distributions, respectively. The simulated rejection rates of these statistics, i.e., the percentage of times that they exceed the appropriate upper points of the chi-squared distributions are given in Tables 1 and 2 (entries are percentages) for the tests of H_2 and H_3, respectively.

Table 1: Rejection rates of ω_2 and ω_2^* for the hypothesis $H_2 : \beta_1 = 2$, $\beta_2 = 1$.

n	nominal level (α%)	$\sigma^2 = 1$ ω_2	$\sigma^2 = 1$ ω_2^*	$\sigma^2 = 4$ ω_2	$\sigma^2 = 4$ ω_2^*	$\sigma^2 = 16$ ω_2	$\sigma^2 = 16$ ω_2^*
20	10	13.22	11.83	14.15	11.77	14.83	12.04
	5	6.88	5.51	7.84	5.62	8.35	5.58
30	10	11.87	10.74	12.21	10.83	12.97	10.94
	5	6.51	6.17	7.35	5.89	7.76	6.19
40	10	11.79	10.89	11.92	10.69	12.49	10.73
	5	6.40	5.33	6.81	5.47	7.12	5.76

Table 2: Rejection rates of ω_3 and ω_3^* for the hypothesis $H_3 : \rho = 0.6$.

n	nominal level (α%)	$\sigma^2 = 1$ ω_3	$\sigma^2 = 1$ ω_3^*	$\sigma^2 = 4$ ω_3	$\sigma^2 = 4$ ω_3^*	$\sigma^2 = 16$ ω_3	$\sigma^2 = 16$ ω_3^*
20	10	13.68	11.56	14.22	12.14	15.41	12.43
	5	6.94	5.85	7.19	5.92	8.47	5.93
30	10	12.87	11.15	13.16	12.25	13.73	11.82
	5	5.88	5.41	6.38	5.60	7.84	5.79
40	10	12.44	10.87	12.87	11.29	13.28	11.88
	5	5.52	5.40	5.93	5.38	6.45	5.52

The figures in these tables convey important information. First, it is clear that the usual LR statistics tend to reject the null hypothesis more often than the expected based on the nominal sizes. In fact, for all 18 cases reported, the rejection rates of the unmodified statistics ω_2 and ω_3 are greater than the corresponding nominal levels. The asymptotic chi-squared distribution usually delivers a very poor approximation to the null distributions of the unmodified statistics ω_2 and ω_3 for small values of n and large values of σ^2. Second, the empirical sizes of the tests based on the modified statistics ω_2^* and ω_3^* are closer to the nominal levels than the empirical sizes of the corresponding unmodified statistics ω_2 and ω_3. Thus, there is strong evidence that the chi-squared distribution provides a better approximation to the distributions of ω_2^* and ω_3^* than to the distributions of ω_2 and ω_3, respectively. This means that the Bartlett corrections c_2 and c_3 are very effective in pushing the rejection rates of the modified statistics ω_2^* and ω_3^* toward to the nominal levels. Third, the asymptotic chi-squared approximation for all statistics works better for large values of n and small values of σ^2, i.e., when the variability of the normal distribution is small, in agreement with the so-called "small-dispersion asymptotic result". Overall, the simulation results presented in this section suggest that the first order asymptotics usually employed with asymptotic chi-squared LR criteria can deliver inaccurate inferences for normal heteroskedastic linear models with samples of small to moderate sizes and when the variability of the data is large. Bartlett corrections based on second order asymptotic theory can then be used to obtain tests with more reliable finite-sample size behaviour.

References

Andrews, D.F. and Stafford, J.E. (1993). Tools for the symbolic computation of asymptotic expansions. *Journal of the Royal Statistical Society B*, **55**, 613–627.

Attfield, C.L.F. (1991). A Bartlett adjustment to the likelihood ratio test for homoskedasticity in the linear model. *Economics Letters*, **37**, 119–123.

Attfield, C.L.F. (1995). A Bartlett adjustment to the likelihood ratio test for a system of equations. *Journal of Econometrics*, **66**, 207–223.

Aubin, E.C.Q. and Cordeiro, G.M. (1999). Some adjusted likelihood ratio tests for heteroscedastic regression models. *Braz. J. Probab. Statist.* (forthcomming).

Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A*, **160**, 268–282.

Berndt, E. and Savin, N.E. (1997). Conflict among criteria for testing hypotheses in the multivariate regression model. *Econometrica*, **45**, 1263–1277.

Botter, D.A. and Cordeiro, G.M. (1997). Bartlett corrections for generalized linear models with dispersion covariates. *Comm. Statist. Theor. Meth.*, **26**, 279–307.

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

Chesher, A. and Smith, R. (1997). Likelihood ratio specification tests. *Econometrica*, **65**, 627–646.

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

Cordeiro, G.M. (1985). The null expected deviance for an extended class of generalized linear models. *Lecture Notes in Statistics*, **32**, 27–34.

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

Cordeiro, G.M. (1993a). Bartlett corrections and bias correction for two heteroscedastic regression models. *Communications in Statistics - Theory and Methods*, **22**, 169–188.

Cordeiro, G.M. (1993b). General matrix formulae for computing Bartlett corrections. *Statistics and Probability Letters*, **16**, 11–18.

Cordeiro, G.M. (1995). Performance of a Bartlett-type modification for the deviance. *Journal of Statistical Computation and Simulation*, **51**, 385–403.

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

Cordeiro, G.M. and Klein, R. (1994). Bias correction in ARMA models. *Statistics and Probability Letters*, **19**, 169–176.

Cox, D.R. and Hinkley, D.V. (1979). *Theoretical Statistics.* London, Chapman and Hall.

Cox, D.R. and Reid, N. (1987). Approximations to noncentral distributions. *Can. J. Statist.*, **15**, 105–114.

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

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

Evans, E.B.A. and Savin, N.E. (1982). Conflict among the criteria revisited: the W, LR and LM tests. *Econometrica*, **50**, 737–748.

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

Galbraith, R.F. and Galbraith, J.I. (1974). On the inverses of some patterned matrices arising in the theory of stationary time series. *Journal of Applied Statistics*, **11**, 63–71.

Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics A*, **29**, 359–378.

Harvey, A. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica*, 461–465.

Jensen, J.L. (1993). A historical sketch and some new results on the improved likelihood ratio statistic. *Scandinavian Journal of Statistics*, **20**, 1–15.

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

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

Ljung, G.M. and Box, G.E.P. (1979). The likelihood function of stationary autorregressive-moving average models. *Biometrika*, **66**, 265–270.

McCullagh, P. (1984). Tensor notation and cumulants of polynomials. *Biometrika*, **71**, 461–476.

McCullagh, P. (1987). *Tensor methods in statistics.* London, Chapman and Hall. 285p.

Møller, J. (1986). Bartlett adjustments for structured covariances. *Scandinavian Journal of Statistics*, **13**, 1–15.

Moulton, L.H., Weissfeld, L.A. and St. Laurent, R.T. (1993). Bartlett corrections factors in logistic regression models. *Computational Statistics and Data Analysis*, **15**, 1–11.

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

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications* (2nd ed.). New York, Wiley. 625p.

Serfling, R.J. (1980). *Approximations Theorems of Mathematical Statistics.* New York, Wiley.

Shaman, P. (1973). On the inverse of the covariance matrix for an autorregressive-moving average process. *Biometrika*, **60**, 193–196.

Stafford, J.E. and Andrews, D.F. (1993). A symbolic algorithm for studying adjustments to the profile likelihood. *Biometrika*, **80**, 715–730.

Wang, Y. (1994). A Bartlett-type adjustment for the likelihood ratio statistic with an ordered alternative. *Statistics and Probability Letters*, **20**, 347–352.

Wong, M.Y. (1991). Bartlett adjustment to the likelihood ratio statistic for testing several slopes. *Biometrika*, **78**, 221–224.

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

9901 - **MONTOYA-DELGADO, L.E., IRONY, T.Z., WHITTLE, M.R., PEREIRA, C.A.B.** Unconditional Exact Test of Hardy-Weinberg Law. 1999. 19p.(RT-MAE-9901)

9902 - **MINKOVA, L.** Modelling of financial markets in high level of inflation. 1999. 8p. (RT-MAE-9902)

9903 - **GUERBEROFF, G.R.** Remark on the critical temperature of the hopfield model. 1999. 10p. (RT-MAE-9903)

9904 - **BAKEVA, V., KOLEV, N.** The optimal blocking time of the unreliable $Geo/G_D/1$ Queuing System. 1999. 12p. (RT-MAE-9904)

9905 - **KOLEV, N., MINKOVA, L.** Quotas on runs of successes and failures in a multi-state markov chain. 1999. 14p. (RT-MAE-9905)

9906 - **KOLEV, N., MINKOVA, L.** Run and frequency quotas in a multi-state Markov chain. 1999. 11p. (RT-MAE-9906)

9907 - **CHIANN, C., MORETTIN, P.A.** Estimation of time varying linear systems. 1999. 39p. (RT-MAE-9907)

9908 - **FONTES, L.R.G., NEVES, E.J., SIDORAVICIUS, V.** Limit velocity for a driven particle in a random medium with mass aggregation. 1999. 14p. (RT-MAE-9908)

9909 - **PEREIRA, C.A.B., STERN, J.M.** Evidence and credibility: Full bayesian significance test for precise hypotheses, 1999. 13p. (RT-MAE-9909)

9910 - **GALEA, M., PAULA, G.A., URIBE-OPAZO, M.** Assessing local influence in univariate elliptical linear regression models, 1999, 16p. (RT-MAE-9910)

9911 - **CARDOSO NETO, J., PAULA, G.A.** Wald one-sided test using generalized estimating equations approach, 1999, 18p. (RT-MAE-9911)

9912 - **CORGINI, M., SANKOVICH, D.P., TANAKA, N.I.** The imperfect boson gas: some mathematical approaches, 1999. 25p. (RT-MAE-9912)

9913 - **BOLFARINE, H., CASTRO, M.** Ancova models with measurement errors, 1999. 13p. (RT-MAE-9913)

9914 - **MACHADO, F.P., MENSHIKOV, M.V., POPOV, S.YU.** Recurrence and transience of multi type branching random walks, 1999. 20p. (RT-MAE-9914)

9915 - **FERNÁNDEZ, R., GALVES, A.** Identifying features in the presence of competing evidence. The case of first-language acquisition, 1999. 11p. (RT-MAE-9915)

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

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05315-970 - São Paulo, Brasil