Ann. Inst. Statist. Math.
Vol. 48, No. 1, 111–125 (1996)

A NOTE ON THE SIMPLE STRUCTURAL REGRESSION MODEL

R. B. ARELLANO-VALLE AND H. BOLFARINE

Departamento de Estatística, IME, Universidade de São Paulo, Caixa Postal 20570, CEP 05389-970, São Paulo, SP, Brasil

(Received August 15, 1994; revised June 8, 1995)

Abstract. In this paper we investigate some aspects like estimation and hypothesis testing in the simple structural regression model with measurement errors. Use is made of orthogonal parametrizations obtained in the literature. Emphasis is placed on some properties of the maximum likelihood estimators and also on the distribution of the likelihood ratio statistics.

Key words and phrases: Bartlett correction factors, information matrix, likelihood ratio statistics, maximum likelihood estimator, orthogonal parametrization, structural normal model.

1. Introduction

The classical simple regression model with measurement errors is defined by the equations

$$\begin{cases} Y_k = y_k + e_k, \\ X_k = x_k + u_k, \\ y_k = \alpha + \beta x_k, \end{cases} \tag{1.1}$$

where e_k and u_k are independent and normally distributed with zero means and variances σ_e^2 and σ_u^2, respectively, which we denote by

$$\begin{pmatrix} e_k \\ u_k \end{pmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_e^2 & 0 \\ 0 & \sigma_u^2 \end{pmatrix}\right),$$

$k = 1, \dots, n$, where N_2 denotes the bivariate normal distribution. If the quantity x_k is considered to be a fixed quantity then, the functional regression model follows. On the other hand, if the quantity x_k is considered to be a random quantity, then the structural regression model follows. In this paper, we consider $x_k \sim N(\mu_x, \sigma_x^2)$, with x_k independent of (e_k, u_k), $k = 1, \dots, n$, a typically made assumption. The main idea behind the equations (1.1) is that $(y_1, x_1), \dots, (y_n, x_n)$ are not observed directly and the estimation has to be based on $(Y_1, X_1), \dots, (Y_n, X_n)$, which are observed. Examples of practical situations where the x_k are measured with error are reported in Fuller (1987). An interesting situation is the case where x_k is the amount of nitrogen in the soil and y_k is the yield of a certain cereal. In this case, the observed X_k values are determined by laboratory analysis and are only estimates of the unobserved x_k values.

As is well known, there are problems with the estimation of the parameters in both cases. In the functional case, β is not consistently estimated. In the structural case, some nonidentifiability problems arise. See, for example, Fuller (1987) and Kendall and Stuart (1979), where extensive bibliographies are provided. A Bayesian treatment for the problem can be found in Zellner (1971). Therefore, in order to make the estimation problem feasible, some additional assumptions have to be considered. In the structural model, a typically made assumption considers that the reliability ratio (Fuller (1987)) $k_x = \sigma_x^2/(\sigma_x^2 + \sigma_u^2)$, or equivalently, $\lambda_x = \sigma_x^2/\sigma_u^2$ is known. Fuller (1987) reports several situations particularly in Sociology, Psychology and Survey Sampling where k_x is so well estimated that it may be taken to be known. Bolfarine and Cordani (1993) derived an orthogonal parametrization in this case and investigated the performance of confidence intervals for β. Another common assumption is to consider that the ratio of the two variances $\lambda_e = \sigma_e^2/\sigma_u^2$ is known. This case has been investigated by Wong (1989) where an orthogonal parametrization is derived and Bartlett correction factors are provided for the likelihood ratio statistic by using the approach of Lawley (1956).

In this paper, a unified approach is developed for both (λ_x known and λ_e known) cases. By studying the distribution of the maximum likelihood estimators of the orthogonal parameters, we investigate the distribution of the likelihood ratio criteria in both cases. The approach also makes it possible to compute directly the expected value of the likelihood ratio criteria to order n^{-2}. As shown, the correction factors obtained are exactly the same in both cases and coincide with the one obtained by Wong (1989).

Section 2 presents a general matrix representation for the model and the orthogonal parametrization (in the sense of Cox and Reid (1987)) under both assumptions. Section 3 discusses maximum likelihood estimation and some properties of the estimators derived are studied in Section 4. Section 5 investigates Bartlett correction factors for the likelihood ratio statistics.

2. Orthogonal parametrizations

Note that we may rewrite model (1.1) as

$$\boldsymbol{Z}_k = \boldsymbol{g}_k + \epsilon_k, \tag{2.1}$$

where

$$\boldsymbol{Z}_k = \begin{pmatrix} Y_k \\ X_k \end{pmatrix}, \quad \boldsymbol{g}_k = \boldsymbol{g}(x_k) = \begin{pmatrix} \alpha + \beta x_k \\ x_k \end{pmatrix}, \quad \text{and} \quad \epsilon_k = \begin{pmatrix} e_k \\ u_k \end{pmatrix},$$

$k = 1, \ldots, n$. Thus, from (2.1), we have that $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$, are independent and identically distributed with

$$\boldsymbol{Z}_k \sim N_2(\mu; \boldsymbol{\Sigma}),$$

where

$$\mu = E[\boldsymbol{Z}_k] = \begin{pmatrix}\mu_Y\\ \mu_X\end{pmatrix} = \begin{pmatrix}\alpha+\beta\mu_x\\ \mu_x\end{pmatrix}, \tag{2.2}$$

and

$$\boldsymbol{\Sigma} = \mathrm{Cov}[\boldsymbol{Z}_k] = \begin{cases} \begin{pmatrix}\lambda_x\beta^2\sigma_u^2+\sigma_e^2 & \lambda_x\beta\sigma_u^2\\ \lambda_x\beta\sigma_u^2 & (\lambda_x+1)\sigma_u^2\end{pmatrix}, & \text{if } \lambda_x \text{ is known,}\\ \begin{pmatrix}\beta^2\sigma_x^2+\lambda\sigma_u^2 & \beta\sigma_x^2\\ \beta\sigma_x^2 & \sigma_x^2+\sigma_u^2\end{pmatrix}, & \text{if } \lambda_e \text{ is known.}\end{cases} \tag{2.3}$$

Further, it can be shown that

$$|\boldsymbol{\Sigma}| = \begin{cases} [\lambda_x\beta^2\sigma_u^2+(\lambda_x+1)\sigma_e^2]\sigma_u^2, & \text{if } \lambda_x \text{ is known,}\\ [\lambda_e\sigma_u^2+(\beta^2+\lambda_e)\sigma_x^2]\sigma_u^2, & \text{if } \lambda_e \text{ is known.}\end{cases} \tag{2.4}$$

Let

$$\theta = \begin{cases} (\alpha,\mu_x,\sigma_e^2,\sigma_u^2,\beta), & \text{if } \lambda_x \text{ is known,}\\ (\alpha,\mu_x,\sigma_x^2,\sigma_u^2,\beta), & \text{if } \lambda_e \text{ is known}\end{cases} \tag{2.5}$$

and $l = l(\theta)$, the log likelihood function which may be written as

$$l \propto -\frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{k=1}^{n}(\boldsymbol{Z}_k-\mu)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}_k-\mu), \tag{2.6}$$

where $\mu(\theta)$ and $\boldsymbol{\Sigma}(\theta)$ are as given in (2.2) and (2.3), respectively. Let $\boldsymbol{K}(\theta) = [\kappa_{i,j}]$ denote the expected information matrix. Thus, from (2.6), after some algebraic manipulations, it can be shown that

$$\kappa_{i,j} = E\left[\frac{\partial l}{\partial\theta_i}\frac{\partial l}{\partial\theta_j}\right] = n\left[\frac{\partial\mu}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial\mu}{\partial\theta_j} + \frac{n}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right)\right], \tag{2.7}$$

where θ_i denotes the i-th component of θ, as defined in (2.5). From (2.7) it follows that

$$\kappa_{i,j} = \begin{cases} n\dfrac{\partial\mu'}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\dfrac{\partial\mu}{\partial\theta_j}, & \text{if } i=1,2,5,\ j=1,2,\\[2ex] \dfrac{n}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\dfrac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\dfrac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right), & \text{if } i=3,4,5 \text{ and } j=3,4,\\[2ex] 0, & \text{if } i=1,2 \text{ and } j=3,4,\end{cases}$$

and

$$\kappa_{5,5} = \kappa_{\beta,\beta} = n\frac{\partial\mu'}{\partial\beta}\boldsymbol{\Sigma}^{-1}\frac{\partial\mu}{\partial\beta} + \frac{n}{2}\operatorname{tr}\left\{\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\beta}\right)^2\right\},$$

where tr($\boldsymbol{A}$) denotes the trace of the matrix $\boldsymbol{A}$. It follows that $\kappa_{\beta,j} \neq 0$, whatever be θ_j and, furthermore, $\kappa_{i,j} = \kappa_{j,i}$, $i, j = 1, \dots, 5$. This fact makes it hard to obtain large sample inference for β, particularly correction factors for testing statistics. One way of aliviating this difficulty is to consider an orthogonal parametrization (as considered in Cox and Reid (1987)), by transforming θ into $\phi = (\phi_1, \phi_2, \phi_3, \phi_4, \beta)'$ so that $\theta_i = \theta_i(\phi)$, $i = 1, 2, 3, 4$, are the solutions to the differential equations:

$$\sum_{i=1}^{4} \kappa_{i,j} \frac{\partial \theta_i}{\partial \beta} = -\kappa_{\beta,j}, \tag{2.8}$$

$j = 1, 2, 3, 4$. Typically, solving a system like the one in (2.8) is not simple. Moreover, when solvable, such equations may not always be easily interpretable. In the case when λ_e is known, a solution is given in Wong (1989) and when λ_x is known, a solution is given in Bolfarine and Cordani (1993). We note that the problem of obtaining the orthogonal parametrization can be simplified by first making the location parameters in μ orthogonal to scale parameters in $\boldsymbol{\Sigma}$, in the sense of Cox and Reid (1987). This is easily accomplished by taking $\phi_0 = \alpha + \beta\mu_x$ and $\phi_1 = \mu_x$. The problem now is to make β orthogonal to the other parameters which appear in $\boldsymbol{\Sigma}$.

The solution presented in Wong (1989) and Bolfarine and Cordani (1993) may be written as

$$\phi_1 = \alpha + \beta\mu_x, \qquad \phi_2 = \mu_x, \qquad \phi_4 = \sigma_u^2, \tag{2.9}$$

and

$$\phi_3 = \begin{cases} \lambda_x \beta^2 \sigma_u^2 + (\lambda_x + 1)\sigma_e^2, & \text{if } \lambda_x \text{ is known}, \\ (\beta^2 + \lambda_e)\sigma_x^2 + \lambda_e \sigma_u^2, & \text{if } \lambda_e \text{ is known}. \end{cases} \tag{2.10}$$

Considering the above parametrization, we have that

$$\mu = \mu(\phi_L) = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

and

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\phi_S) = \begin{cases} (\lambda_x + 1)^{-1} \begin{pmatrix} \phi_3 + (\lambda_x\beta)^2\phi_4 & (\lambda_x + 1)\lambda_x\beta\phi_4 \\ (\lambda_x + 1)\lambda_x\beta\phi_4 & (\lambda_x + 1)^2\phi_4 \end{pmatrix}, \\ \qquad\qquad \text{if } \lambda_x \text{ is known}, \\ (\beta^2 + \lambda_e)^{-1} \begin{pmatrix} \beta^2\phi_3 + \lambda_e^2\phi_4 & \beta(\phi_3 - \lambda_e\phi_4) \\ \beta(\phi_3 - \lambda_e\phi_4) & \phi_3 + \beta^2\phi_4 \end{pmatrix}, \\ \qquad\qquad \text{if } \lambda_e \text{ is known}, \end{cases} \tag{2.11}$$

where $\phi_L = (\phi_1, \phi_2)'$ (the location parameters) and $\phi_S = (\phi_3, \phi_4, \beta)'$ (the scale parameters). Note that in this new parametrization

$$|\boldsymbol{\Sigma}| = \phi_3\phi_4.$$

We call attention to the fact that the choice of the scale parameters are not as obvious as the location parameters. However, the choice of the new parameters becomes obvious and clear from (2.4). Moreover, when λ_e is known and taken to be equal to one (without loss of generality), it can be shown that

$$\operatorname{tr}(\mathbf{\Sigma}) = \phi_3 + \phi_4,$$

so that ϕ_3 and ϕ_4 are the characteristic roots of $\mathbf{\Sigma}$. In the sequel, we present some properties of the matrix $\mathbf{\Sigma}$ which will make it easier to derive the cumulants of the log likelihood function $l = l(\phi)$.

Let

$$\alpha_3 = \begin{cases} (\lambda_x + 1)^{-1/2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)^{-1/2}\begin{pmatrix} \beta \\ 1 \end{pmatrix}, & \text{if } \lambda_e \text{ is known} \end{cases}$$

and

$$\alpha_4 = \begin{cases} (\lambda_x + 1)^{-1/2}\begin{pmatrix} \lambda_x\beta \\ \lambda_x \end{pmatrix}, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)^{-1/2}\begin{pmatrix} \lambda_e \\ -\beta \end{pmatrix}, & \text{if } \lambda_e \text{ is known,} \end{cases}$$

and note that

$$\frac{\partial \mathbf{\Sigma}}{\partial \phi_i} = \alpha_i \alpha_i',$$

$i = 3, 4$, and

$$\mathbf{\Sigma} = \phi_3 \frac{\partial \mathbf{\Sigma}}{\partial \phi_3} + \phi_4 \frac{\partial \mathbf{\Sigma}}{\partial \phi_4} = \phi_3 \alpha_3 \alpha_4' + \phi_4 \alpha_4 \alpha_4'.$$

Similarly,

$$\mathbf{\Sigma}^{-1} = \phi_3 \mathbf{\Sigma}^{-1} \frac{\partial \mathbf{\Sigma}}{\partial \phi_3} \mathbf{\Sigma}^{-1} + \phi_4 \mathbf{\Sigma}^{-1} \frac{\partial \mathbf{\Sigma}}{\partial \phi_4} \mathbf{\Sigma}^{-1} = \phi_3^{-1} \bar{\alpha}_3 \bar{\alpha}_3' + \phi_4^{-1} \bar{\alpha}_4 \bar{\alpha}_4'$$

and

$$\mathbf{\Sigma}^{-1} \frac{\partial \mathbf{\Sigma}}{\partial \phi_i} = \phi_i^{-1} \bar{\alpha}_i \alpha_i' \tag{2.12}$$

where

$$\bar{\alpha}_i = \phi_i \mathbf{\Sigma}^{-1} \alpha_i,$$

$i = 3, 4$, are such that $\bar{\alpha}_i' \alpha_i = 1$, $\bar{\alpha}_i' \alpha_j = 0$, $||\bar{\alpha}_i|| = ||\alpha_j||$, $i \neq j = 3, 4$; that is,

$$\bar{\alpha}_3 = \begin{cases} (\lambda_x + 1)^{-1/2}\begin{pmatrix} \lambda_x + 1 \\ -\lambda_x\beta \end{pmatrix}, & \text{if } \lambda_x \text{ is known} \\ (\beta^2 + \lambda_e)^{-1/2}\begin{pmatrix} \beta \\ \lambda_e \end{pmatrix}, & \text{if } \lambda_e \text{ is known,} \end{cases} \tag{2.13}$$

and

$$\bar{\alpha}_4 = \begin{cases} (\lambda_x + 1)^{-1/2}\begin{pmatrix}0\\1\end{pmatrix}, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)^{-1/2}\begin{pmatrix}1\\-\beta\end{pmatrix}, & \text{if } \lambda_e \text{ is known.} \end{cases} \tag{2.14}$$

Note that if $\bar{\boldsymbol{A}} = (\bar{\alpha}_3, \bar{\alpha}_4)$ and $\boldsymbol{\Phi} = \operatorname{diag}(\phi_3, \phi_4)$ then $\bar{\boldsymbol{A}}(\boldsymbol{Z}_i - \mu) \overset{\text{iid}}{\sim} N_2(\mathbf{0}, \boldsymbol{\Phi})$, $i = 1, \ldots, n$. Furthermore, from the above results, it is easy to see that

$$\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \beta} = \left(\frac{\phi_3\phi_4}{\sigma_\beta^2}\right)^{1/2}(\phi_3^{-1}\bar{\alpha}_3\alpha_4' + \phi_4^{-1}\bar{\alpha}_4\alpha_3'),$$

from where it follows that

$$\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \beta}\right)^2 = \sigma_\beta^{-2}\boldsymbol{I}_2, \tag{2.15}$$

where $\boldsymbol{I}_2$ is the two dimensional identity matrix and $\sigma_\beta = \sqrt{\sigma_\beta^2}$, with

$$\sigma_\beta = \begin{cases} \dfrac{1}{\lambda_x}\left(\dfrac{\phi_3}{\phi_4}\right)^{1/2}, & \text{if } \lambda_x \text{ is known,} \\ \left(\dfrac{\beta^2 + \lambda_e}{\phi_3 - \lambda_e\phi_4}\right)(\phi_3\phi_4)^{1/2}, & \text{if } \lambda_e \text{ is known.} \end{cases} \tag{2.16}$$

From (2.12) it follows that

$$\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_i}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_j}\right) = (\phi_i\phi_j)^{-1}\bar{\alpha}_i'\bar{\alpha}_j = \begin{cases} \phi_i^{-2}, & i = j \\ 0, & i \neq j, \end{cases} \tag{2.17}$$

$i, j = 3, 4$, and from (2.15)

$$\operatorname{tr}\left\{\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \beta}\right)^2\right\} = \operatorname{tr}(\sigma_\beta^{-2}\boldsymbol{I}_2) = 2\sigma_\beta^{-2}. \tag{2.18}$$

Now let $\boldsymbol{K} = \boldsymbol{K}(\phi) = [\kappa_{i,j}]$ the information matrix under the orthogonal parametrization. Thus, from (2.7) with θ replaced by ϕ, we have that

$$\kappa_{i,j} = \begin{cases} n\dfrac{\partial \mu'}{\partial \phi_i}\boldsymbol{\Sigma}^{-1}\dfrac{\partial \mu}{\partial \phi_j}, & i, j = 1, 2, \\ \dfrac{n}{2}\operatorname{tr}\left\{\left(\boldsymbol{\Sigma}^{-1}\dfrac{\partial \boldsymbol{\Sigma}}{\partial \phi_i}\right)^2\right\}, & i = 3, 4, 5, \\ 0, & i = 1, 2,\ j = 3, 4, 5, \end{cases}$$

where $\mu = \mu(\phi_L)$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\phi_S)$ are as defined above and $\kappa_{i,j} = \kappa_{j,i}$, $i, j = 1, \ldots, 5$. Thus, $\boldsymbol{K} = \operatorname{diag}(\boldsymbol{K}_L, \boldsymbol{K}_S)$, that is, $\boldsymbol{K}$ is a block diagonal matrix with

$$\boldsymbol{K}_L = n\boldsymbol{\Sigma}^{-1}$$

and, using (2.17) and (2.18),

$$\boldsymbol{K}_S = n\,\mathrm{diag}\left(\frac{1}{2\phi_3^2}, \frac{1}{2\phi_4^2}, \frac{1}{\sigma_\beta^2}\right),$$

where σ_β is given in (2.16). Note that in both cases (λ_x and λ_e known), σ_β^2 can be the asymptotic variance of $\sqrt{n}\hat{\beta}$, where $\hat{\beta}$ is the likelihood estimator of β (Section 3) and may be written as

$$\sigma_\beta^2 = \beta^2 \frac{\sigma_{XX}\sigma_{YY.X}}{\sigma_{YX}^2} = \beta^2\left(\frac{1-\rho_{YX}^2}{\rho_{YX}^2}\right),$$

where $\sigma_{YX} = \mathrm{Cov}[Y_i, X_i]$, $\sigma_{XX} = \mathrm{Var}[X_i]$, $\sigma_{YY.X} = \mathrm{Var}[Y_i|X_i] = \sigma_{YY} - \sigma_{XX}^{-1}\sigma_{YX}^2 = \sigma_{YY}(1-\rho_{YX}^2)$ and

$$\rho_{YX} = \frac{\sigma_{YX}}{(\sigma_{XX}\sigma_{YY})^{1/2}},$$

which denotes the correlation between Y_k and X_k, $k = 1, \ldots, n$.

3. Maximum likelihood estimators

The loglikelihood function $l = l(\phi)$ with respect to the orthogonal parameters given in (2.9) and (2.10) may be written as in (2.6), with μ replaced by $\mu(\phi_L)$ and $\boldsymbol{\Sigma}$ replaced by $\boldsymbol{\Sigma}(\phi_S)$. Hence, the maximum likelihood estimator $(\hat{\phi}_L, \hat{\phi}_S)$ of (ϕ_L, ϕ_S) is obtained by solving the equations

$$\left.\frac{\partial l}{\partial \phi_i}\right|_{\phi=\hat{\phi}} = \sum_{k=1}^{n} \frac{\partial \mu'}{\partial \phi_i}\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}_k - \mu)|_{\phi=\hat{\phi}} = 0,$$

$i = 1, 2,$

$$\left.\frac{\partial l}{\partial \phi_i}\right|_{\phi=\hat{\phi}} = \left\{ -\frac{n}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_i}\right) + \frac{1}{2}\sum_{k=1}^{n}(\boldsymbol{Z}_k - \mu)'\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \phi_i}\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}_k - \mu)\right\}\Bigg|_{\phi=\hat{\phi}} = 0,$$

$i = 3, 4,$ and

$$\left.\frac{\partial l}{\partial \beta}\right|_{\phi=\hat{\phi}} = \frac{1}{2}\sum_{k=1}^{n}(\boldsymbol{Z}_k - \mu)'\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \beta}\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}_k - \mu)|_{\phi=\hat{\phi}} = 0.$$

The first equation leads to $\hat{\mu} = \mu(\hat{\phi}_L) = \bar{\boldsymbol{Z}}$, from which we get

$$\hat{\phi}_1 = \bar{Y}, \quad \text{and} \quad \hat{\phi}_2 = \bar{X},$$

since $\mu = (\phi_1, \phi_2)'$ and $\bar{\boldsymbol{Z}} = \sum_{k=1}^{n} \boldsymbol{Z}_k/n = (\bar{Y}, \bar{X})'$. Using these estimators, it follows from the above equations that

$$\hat{\phi}_i = \bar{\alpha}_i'(\hat{\beta})\boldsymbol{S}\bar{\alpha}_i(\hat{\beta}),$$

$i = 3, 4$, and

$$\bar{\alpha}_3'(\hat{\beta})\boldsymbol{S}\bar{\alpha}_4(\hat{\beta}) = 0,$$

respectively, where $\bar{\alpha}_i(\hat{\beta})$, $i = 3, 4$, is as defined in (2.13) and (2.14), respectively, with β replaced by $\hat{\beta}$ and

$$\boldsymbol{S} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{Z}_i - \bar{\boldsymbol{Z}})(\boldsymbol{Z}_i - \bar{\boldsymbol{Z}})' = \begin{pmatrix} S_{YY} & S_{YX} \\ S_{YX} & S_{XX} \end{pmatrix},$$

where $S_{XX} = \sum_{i=1}^{n}(X_i - \bar{X})^2/n$, $S_{YY} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2/n$ and $S_{YX} = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})/n$. From the above equations, it follows that

(i) when λ_x is known (Bolfarine and Cordani (1993)):

$$\hat{\phi}_3 = (\lambda_x + 1)S_{YY} - 2(\lambda_x\hat{\beta})S_{YX} + (\lambda_x\hat{\beta})^2\hat{\phi}_4,$$
$$\hat{\phi}_4 = \frac{S_{XX}}{\lambda_x + 1},$$
$$\hat{\beta} = \left(\frac{\lambda_x + 1}{\lambda_x}\right)\frac{S_{YX}}{S_{XX}}.$$

Replacing $\hat{\phi}_4$ and $\hat{\beta}$ in $\hat{\phi}_3$, we have that

$$\hat{\phi}_3 = (\lambda_x + 1)S_{YY.X},$$

where

$$S_{YY.X} = S_{YY} - S_{XX}^{-1}S_{YX}^2 = S_{YY}(1 - r_{YX}^2) \quad \text{and} \quad r_{YX} = \frac{S_{YX}}{(S_{YY}S_{XX})^{1/2}};$$

(ii) when λ_e is known (Wong (1989)):

$$\hat{\phi}_3 = \frac{\hat{\beta}^2 S_{YY} + 2\lambda_e\hat{\beta}S_{YX} + \lambda_e^2 S_{XX}}{\hat{\beta}^2 + \lambda_e},$$
$$\hat{\phi}_4 = \frac{S_{YY} - 2\hat{\beta}S_{YX} + \hat{\beta}^2 S_{XX}}{\hat{\beta}^2 + \lambda_e},$$
$$\hat{\beta} = \left(\frac{S_{YY} - \lambda_e S_{XX}}{2S_{YX}}\right) + (\text{sign}(S_{YX}))\left\{\left(\frac{S_{YY} - \lambda_e S_{XX}}{2S_{YX}}\right)^2 + \lambda_e\right\}^{1/2}.$$

It can also be shown that the maximum likelihood estimator of $\phi_S = (\phi_3, \phi_4, \beta)'$ is given by the solution of the equation

$$\boldsymbol{\Sigma}(\hat{\phi}_S) = \boldsymbol{S},$$

where $\Sigma(\hat{\phi}_S)$ is as given in (2.11), with ϕ_S replaced by $\hat{\phi}_S$. Some estimators may also be given alternative expressions as, for example,

$$\hat{\phi}_3 = \begin{cases} (\lambda_x + 1)S_{YY} - (\lambda_x\hat{\beta})^2\hat{\phi}_4, & \text{if } \lambda_x \text{ is known,} \\ \lambda_e S_{YY} - \hat{\beta}S_{YX}, & \text{if } \lambda_e \text{ is known.} \end{cases}$$

Note also that, in both cases,

$$\hat{\phi}_3\hat{\phi}_4 = |\hat{\Sigma}| = S_{YY.X}S_{XX}. \tag{3.1}$$

4. Some properties of the maximum likelihood estimators

In this section we study some properties of $\hat{\phi} = (\phi_L', \phi_S')'$, for λ_x and λ_e known. Under model (1.1), it follows that

$$\hat{\mu} = \bar{Z} \sim N_2\left(\mu, \frac{1}{n}\Sigma\right),$$

and

$$\hat{\Sigma} = S \sim W_2\left(\frac{1}{n}\Sigma, n-1\right), \tag{4.1}$$

are independent, where

$$\mu = \begin{pmatrix} \mu_Y \\ \mu_X \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_{YY} & \sigma_{YX} \\ \sigma_{YX} & \sigma_{XX} \end{pmatrix},$$

with $\bar{Z}$ and S as before. Here, $W_k(A, m)$ denotes the k-variate Wishart distribution with dispersion matrix A and m degrees of freedom (Muirhead (1982)). From these results, it follows that $\hat{\phi}_L = (\hat{\phi}_1, \hat{\phi}_2)'$ and $\hat{\phi}_S = (\hat{\phi}_2, \hat{\phi}_4, \hat{\beta})'$ are independent and

$$\hat{\phi}_L = \bar{Z} \sim N_2\left(\phi_L, \frac{1}{n}\Sigma\right).$$

Thus, confidence regions or hypothesis testing for $H_0 : \phi_L = \gamma_0$ can be performed by considering the variable

$$\begin{aligned} F &= \left(\frac{n-2}{2}\right)(\hat{\phi}_L - \phi_L)'(\Sigma^{-1}(\hat{\phi}_S))(\hat{\phi}_L - \phi_L) \\ &= \left(\frac{n-2}{2}\right)(\bar{Z} - \mu)'S^{-1}(\bar{Z} - \mu) \sim F_{2,n-2}, \end{aligned}$$

that is, F is ditributed according to the Fisher F distribution with 2 and $n-2$ degrees of freedom. Furthermore, confidence intervals or hypothesis testing for functions of the form $a'\phi_L$ where a is a known vector of constants, we can use

$$\sqrt{n}\frac{(a'\hat{\phi}_L - a'\phi_L)}{(a'Sa)^{1/2}} \sim t(0, 1; n-1),$$

where $t(0,1;n-1)$ denotes the central t distribution with $n-1$ degrees of freedom. On the other hand, the exact marginal distribution of the components of the vector $\hat{\phi}_S$ is particularly difficult to obtain when λ_e is known. In particular, if $\lambda_e = 1$, then $\hat{\phi}_3$ and $\hat{\phi}_4$ are the proper values of the Wishart matrix $\mathbf{\Sigma}(\hat{\phi}_S)$. Thus, in this situation, the distribution function of $(\hat{\phi}_3, \hat{\phi}_4)$ can be represented in terms of an infinite series (Muirhead (1982)). Similarly, the distribution function of $\hat{\beta}$ can be given as a convergent infinite series of incomplete beta functions (Anderson and Sawa (1982) derived the distribution of $\hat{\beta}$ in the functional case. This distribution corresponds to the conditional distribution of $\hat{\beta}$ given $\boldsymbol{x} = (x_1, \ldots, x_n)'$ in the structural case). For this reason, inference on β when λ_e is known typically is based on large samples, since as $n \to \infty$,

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} N(0, \sigma_\beta^2).$$

Moreover, using properties of the Wishart distribution we can study the exact distribution of certain functions of $\hat{\phi}_S$ in both cases (λ_x or λ_e known). Thus, considering the Wishart distribution in (4.1) it follows, in both cases, that

(i) $S_{YY.X} = S_{YY} - \frac{S_{YX}^2}{S_{XX}} = S_{YY}(1 - r_{YX}^2)$ is independent of (S_{YX}, S_{XX});

(ii) $\frac{nS_{YY.X}}{\sigma_{YY.X}} \sim \chi_{n-2}^2$, where as seen before, $\sigma_{YY.X} = \sigma_{YY} - \sigma_{XX}^{-1}\sigma_{YX}^2 = \sigma_{YY}(1-\rho_{YX}^2)$;

(iii) $S_{YX} \mid S_{XX} \sim N(\frac{\sigma_{YX}}{\sigma_{XX}}S_{XX}, \frac{\sigma_{YY.X}}{n}S_{XX})$;

(iv) $\frac{nS_{YY}}{\sigma_{YY}} \sim \chi_{n-1}^2$ and $\frac{nS_{XX}}{\sigma_{XX}} \sim \chi_{n-1}^2$;

(v) $\frac{S_{YX}}{S_{XX}} \sim t(\frac{\sigma_{YX}}{\sigma_{XX}}, \frac{\sigma_{YY.X}}{(n-1)\sigma_{XX}}; n-1)$

and

(vi) $(\frac{(n-2)S_{XX}}{S_{YY.X}})^{1/2}(\frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}}) \sim t(0,1;n-2)$.

Results (i) to (ii) follows directly from Theorem 3.2.10 in Muirhead (1982). Notice also that (iii) follows from the fact that $Y_i \mid X_i \sim N(\mu_Y + \sigma_{YX}\sigma_{XX}^{-1}(X_i - \mu_x), \sigma_{YY.X})$, $i = 1, \ldots, n$. To prove (v) and (vi) notice first that (iii) implies that

$$\left(\frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}}\right) \Bigg| S_{XX} \sim N\left(0, \frac{\sigma_{YY.X}}{nS_{XX}}\right). \tag{4.2}$$

Let

$$U_{XX} = \frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}} \quad \text{and} \quad V_{XX} = \frac{nS_{XX}}{(n-1)\sigma_{XX}}.$$

Thus, from (4.2) it follows that

$$(U_{XX} \mid S_{XX}) \stackrel{d}{=} (U_{XX} \mid V_{XX}) \sim N\left(0, \frac{\sigma_{YY.X}}{(n-1)\sigma_{XX}} V_{XX}^{-1}\right), \tag{4.3}$$

where $X \stackrel{d}{=} Y$ means that X and Y have the same distribution. Moreover according to (iv), it follows that

$$V_{XX} \sim \frac{1}{n-1}\chi_{n-1}^2. \tag{4.4}$$

From (4.3) and (4.4) it follows that the unconditional distribution of U_{XX} is given by

$$U_{XX} = \frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}} \sim t\left(0, \frac{\sigma_{YY.X}}{(n-1)\sigma_{XX}}; n-1\right),$$

since, as is well known, the t distribution is a mixture of a normal distribution with a chisquare distribution (Muirhead (1982), Arellano-Valle *et al.* (1995)). This proves (v). Now, notice that (4.2) implies that $S_{XX}^{-1/2}U_{XX}$ is independent of S_{XX}, from where it follows that

$$S_{XX}^{-1/2}U_{XX} = S_{XX}^{-1/2}\left(\frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}}\right) \sim N\left(0, \frac{\sigma_{YY.X}}{n}\right). \tag{4.5}$$

Thus, (vi) follows from (i), (ii) and (4.5).

Moreover, (vi) allows making inference on the ratio (function of ϕ_S)

$$\frac{\sigma_{YX}}{\sigma_{XX}} = \begin{cases} \left(\dfrac{\lambda_x}{\lambda_x + 1}\right)\beta, & \text{if } \lambda_x \text{ is known,} \\ \left(\dfrac{\phi_3 - \lambda_e\phi_4}{\phi_3 + \beta^2\phi_4}\right)\beta, & \text{if } \lambda_e \text{ is known.} \end{cases}$$

For example, for testing $H_0 : \beta = 0$ (independence between X and Y), we can use the fact that

$$\frac{\sqrt{n-2}\hat{\beta}}{\hat{\sigma}_\beta} = \frac{\sqrt{n-2}r_{YX}}{\sqrt{1-r_{YX}^2}} = \frac{\sqrt{n-2}S_{YX}}{\sqrt{S_{XX}S_{YY.X}}} \sim t(0, 1; n-2),$$

where $\sigma_\beta = \sigma_\beta(\phi_S)$ is as defined in (2.16). From results given above, we also have, for λ_x and λ_e known, that

$$E[\hat{\phi}_3\hat{\phi}_4] = E[S_{YY.X}]E[S_{XX}] = \left(\frac{n-2}{n}\right)\left(\frac{n-1}{n}\right)\phi_3\phi_4, \tag{4.6}$$

since, as considered in (3.1),

$$S_{YY.X}S_{XX} = |\boldsymbol{S}| = |\boldsymbol{\Sigma}(\hat{\phi}_S)| = \hat{\phi}_3\hat{\phi}_4$$

and

$$\sigma_{YY.X}\sigma_{XX} = |\boldsymbol{\Sigma}| = \phi_3\phi_4.$$

Moreover,

$$\begin{aligned} E\left[\log\left(\frac{\hat{\phi}_3\hat{\phi}_4}{\phi_3\phi_4}\right)\right] &= E\left[\log\left(\frac{S_{YY.X}S_{XX}}{\sigma_{YY.X}\sigma_{XX}}\right)\right] \\ &= E\left[\log\left(\frac{S_{YY.X}}{\sigma_{YY.X}}\right)\right] + E\left[\log\left(\frac{S_{XX}}{\sigma_{XX}}\right)\right] \\ &= \psi\left(\frac{n-2}{2}\right) + \psi\left(\frac{n-1}{2}\right) - 2\log\frac{n}{2}, \end{aligned} \tag{4.7}$$

which follows from the fact that if $V \sim Ga(\nu, \delta)$, the gamma distribution with parameters ν and δ, then $E[\log V] = \psi(\nu) - \log \delta$, with $\psi(\cdot)$ being the digamma function (Abramowitz and Stegun (1965)). In particular, if $\lambda_e = 1$ (or known), we can write

$$\hat{\phi}_3 + \hat{\phi}_4 = \mathrm{tr}\{\boldsymbol{\Sigma}(\hat{\phi}_E)\} = \mathrm{tr}(\boldsymbol{S}) = S_{YY} + S_{XX},$$

from which it follows that

$$E[\hat{\phi}_3 + \hat{\phi}_4] = \left(\frac{n-1}{n}\right)(\phi_3 + \phi_4).$$

Finally, considering λ_x known, we have that

$$\phi_3 = (\lambda_x + 1)\sigma_{YY.X}, \qquad \phi_4 = \frac{\sigma_{XX}}{\lambda_x + 1}, \qquad \beta = \left(\frac{\lambda_x + 1}{\lambda_x}\right)\frac{\sigma_{YX}}{\sigma_{XX}},$$

so that the maximum likelihood estimators of the above parameters are given by

$$\hat{\phi}_3 = (\lambda_x + 1)S_{YY.X}, \qquad \hat{\phi}_4 = \frac{S_{XX}}{\lambda_x + 1}, \qquad \hat{\beta} = \frac{S_{YX}}{S_{XX}} = \left(\frac{\lambda_x + 1}{\lambda_x}\right)\frac{S_{YX}}{\lambda_x \hat{\phi}_4}.$$

Considering the above relations and the previous results, when λ_x is known, we have that

(i) $\hat{\phi}_3$, $\hat{\phi}_4$ and $\hat{\phi}_4^{1/2}(\hat{\beta} - \beta)$ are independent;

(ii) $\frac{n\hat{\phi}_3}{\phi_3} \sim \chi^2_{n-2}$ and $\frac{n\hat{\phi}_4}{\phi_4} \sim \chi^2_{n-1}$;

(iii) $\hat{\phi}_4^{1/2}(\hat{\beta} - \beta) \sim N(0, \frac{\sigma^2_\beta \phi_4}{n})$, where σ^2_β is as given in (2.16).

(iv) $\hat{\beta} \sim t(\beta, \frac{\sigma^2_\beta}{n-1}; n-1)$; and

(v) $\frac{\sqrt{n-2}(\hat{\beta}-\beta)}{\hat{\sigma}_\beta} \sim t(0, 1; n-2)$.

Notice from (v) that,

$$E[\hat{\beta}] = \beta, \quad n > 2 \quad \text{and} \quad \mathrm{Var}[\hat{\beta}] = \frac{\sigma_\beta}{n-3}, \quad n > 3.$$

Moreover, from (v) it follows that an exact $(1-\alpha)100\%$ confidence interval for β is given by

$$\left(\hat{\beta} \mp t_{n-2,\alpha/2}\frac{\hat{\sigma}_\beta}{\sqrt{n-2}}\right),$$

where $t_{n-2,\alpha/2}$ is the upper $1-\alpha/2$ point of a t distribution with $n-2$ degrees of freedom which can also be used as an exact α level test for $H_0 : \beta = \beta_0$.

5. The likelihood ratio statistics

Let $\tilde{\phi} = (\tilde{\phi}_L', \tilde{\phi}_S')'$ the maximum likelihood estimator of $\phi = (\phi_L', \phi_S')'$ under the null hypothesis $H_0 : \beta = \beta_0$. It is easy to see that $\tilde{\phi}_L = \hat{\phi}_L = \bar{\boldsymbol{Z}}$ and $\tilde{\phi}_S = (\tilde{\phi}_3, \tilde{\phi}_4, \tilde{\beta})'$ follows from the equations

$$\tilde{\beta} = \beta_0$$

and

$$\tilde{\phi}_i = \bar{\alpha}_i'(\beta_0)\boldsymbol{S}\bar{\alpha}_i(\beta_0),$$

where $\bar{\alpha}_i(\beta_0)$, $i = 3, 4$, are as defined in (2.13) and (2.14), with β replaced by β_0. In the model with λ_x known, it follows that $\tilde{\phi}_4 = \hat{\phi}_4$. Under $H_0 : \beta = \beta_0$, we have that

$$n\boldsymbol{S} \sim W_2(\boldsymbol{\Sigma}_0, n-1),$$

where $\boldsymbol{\Sigma}_0$ is the same as $\boldsymbol{\Sigma}$ (defined in (2.11)), but evaluated at $(\phi_3, \phi_4, \beta_0)$. This implies that

$$\frac{n\tilde{\phi}_i}{\phi_i} \sim \chi^2_{n-1}, \tag{5.1}$$

$i = 3, 4$. However, $\tilde{\phi}_3$ and $\tilde{\phi}_4$ are independent only in the model with λ_x known. The likelihood ratio statistics for testing $H_0 : \beta = \beta_0$ against $H_1 : \beta \neq \beta_0$ is given by

$$G = 2(L(\hat{\phi}) - L(\tilde{\phi})) = n \log \left\{ \frac{|\boldsymbol{\Sigma}(\tilde{\phi}_S)|}{|\boldsymbol{\Sigma}(\hat{\phi}_S)|} \right\} = n \log \left\{ \frac{\tilde{\phi}_3 \tilde{\phi}_4}{\hat{\phi}_3 \hat{\phi}_4} \right\}. \tag{5.2}$$

Under $H_0 : \beta = \beta_0$, the statistics G has asymptotic chisquare distribution with one degree of freedom, denoted by $G \sim A\chi^2_1$, and can be represented in terms of β_0 and the elements of $\boldsymbol{S}$ as

$$G(\beta_0) = \begin{cases} n \log \left\{ \dfrac{(\lambda_x + 1)S_{YY} - 2\beta_0 \lambda_x S_{YX} + (\beta_0 \lambda_x)^2 S_{XX}}{(\lambda_x + 1) S_{YY.X}} \right\}, & \text{if } \lambda_x \text{ is known,} \\ n \log \left\{ \dfrac{(\beta_0^2 S_{YY} + 2\lambda_e \beta_0 S_{YX} + \lambda_e S_{XX})(S_{YY} - 2\beta_0 S_{YX} + \beta_0^2 S_{XX})}{(\beta_o^2 + \lambda_e)^2 S_{YY.X} S_{XX}} \right\}, & \text{if } \lambda_e \text{ is known,} \end{cases} \tag{5.3}$$

where, as before, $S_{YY.X} = S_{YY}(1 - r^2_{YX})$. Notice that in the case where λ_x is known,

$$G = n \log \left(\frac{\tilde{\phi}_3}{\hat{\phi}_3} \right).$$

As has been extensively discussed in the literature (Cordeiro (1983), Wong (1989)), the approximation of the distribution of the statistics G to the chisquare distribution can be improved by using Bartlett correction factors. For the case when λ_e is known, the correction factor has been derived by Wong (1989), by using the approach developed by Lawley (1956), which typically is difficult to implement since it depends on the fourth order cumulants of the likelihood ratio statistics. We propose now an alternative approach of deriving the correction factor for both cases (λ_x known and λ_e known) by computing directly the expected value of the likelihood ratio statistics by using some results derived in the previous section. Letting $E_0[G]$ denote the expected value of G under the null hypothesis $H_0 : \beta = \beta_0$, it follows from (5.1) and (4.7) that

$$
\begin{aligned}
E_0[G] &= n\left\{E_0\left[\log\left\{\frac{\tilde{\phi}_3\tilde{\phi}_4}{\phi_3\phi_4}\right\}\right] - E_0\left[\log\left\{\frac{\hat{\phi}_3\hat{\phi}_4}{\phi_3\phi_4}\right\}\right]\right\}\\
&= n\left\{\left(\psi\left(\frac{n-1}{2}\right) + \psi\left(\frac{n-1}{2}\right) - 2\log\frac{n}{2}\right)\right.\\
&\qquad \left. - \left(\psi\left(\frac{n-2}{2}\right) - \psi\left(\frac{n-1}{2}\right) - 2\log\frac{n}{2}\right)\right\}\\
&= n\left\{\psi\left(\frac{n-1}{2}\right) - \psi\left(\frac{n-2}{2}\right)\right\},
\end{aligned}
$$

where $\psi(m)$ is the digamma function evaluated at m. Using the fact that (Abramowitz and Stegun (1965))

$$
\begin{aligned}
\psi(m) &= \psi(m-1) + \frac{1}{m-1},\\
\psi(m) &= \log m - \frac{1}{2m} - \frac{1}{12m^2} + \frac{1}{120m^4} + \cdots\\
&= \log m - \frac{1}{2m} - \frac{1}{12m^2} + O(m^{-4}),
\end{aligned}
$$

we have that

$$
\begin{aligned}
(5.4)\quad E_0[G] &= n\left\{\psi\left(\frac{n+1}{2}\right) - \psi\left(\frac{n}{2}\right) - 2\left(\frac{1}{n-1} - \frac{1}{n-2}\right)\right\}\\
&= n\left\{\left[\log\frac{n+1}{2} - \log\frac{n}{2}\right] - \left(\frac{1}{n+1} - \frac{1}{n}\right)\right.\\
&\qquad \left. - \frac{1}{3}\left(\frac{1}{(n+1)^2} - \frac{1}{n^2}\right) + O(n^{-4}) - 2\left(\frac{1}{n-1} - \frac{1}{n-2}\right)\right\}\\
&= n\left\{\log\left(1 + \frac{1}{n}\right) - \frac{1}{n}\left[\left(1 + \frac{1}{n}\right)^{-1} - 1\right]\right.\\
&\qquad - \frac{1}{3n^2}\left[\left(1 + \frac{1}{n}\right)^{-2} - 1\right] + O(n^{-4})\\
&\qquad \left. - \frac{2}{n}\left[\left(1 - \frac{1}{n}\right)^{-1} - \frac{1}{2}\left(\frac{1}{2} - \frac{1}{n}\right)^{-1}\right]\right\}.
\end{aligned}
$$

Considering the expansions

$$\log\left(1+\frac{1}{n}\right)=\frac{1}{n}-\frac{1}{2n^2}+O(n^{-3}),$$

$$\left(1+\frac{1}{n}\right)^{-k}=1-\frac{k}{n}+O(n^{-2}),\qquad k=1,2,$$

$$\left(1-\frac{1}{n}\right)^{-1}=1+\frac{1}{n}+O(n^{-2}),\qquad \left(\frac{1}{2}-\frac{1}{n}\right)^{-1}=2+\frac{4}{n}+O(n^{-2}),$$

it follows from (5.4) that

$$E_0[G]=1+\frac{5}{2n}+O(n^{-2}),$$

so that for both cases (λ_x and λ_e known) the corrected likelihood ratio statistics is given by

$$G^*=G/(1+5/2n), \tag{5.5}$$

with G as given in (5.3). Notice from (5.5) that the correction factor derived is exactly the one derived by Wong (1989) for the case λ_e known, and that it is the same for the case λ_x known.

Acknowledgements

Helpfull suggestions from a referee improved the presentation. The authors acknowledge partial finantial support from CNPq-Brasil.

References

Abramowitz, M. and Stegun, I. A. (eds.) (1965). *Handbook of Mathematical Functions*, Dover, New York.

Anderson, T. W. and Sawa, T. (1982). Exact and approximate distribution of maximum likelihood estimation of a slope coefficient, *J. Roy. Statist. Soc. Ser. B*, **44**, 52–62.

Arellano-Valle, R. B., Bolfarine, H. and Iglesias, P. L. (1995). A predictivistic interpretation of the t distribution, *Test*, **3** (2), 224–236.

Bolfarine, H. and Cordani, L. K. (1993). Estimation of a structural linear regression model with a known reliability ratio, *Ann. Inst. Statist. Math.*, **3**, 531–540.

Cordeiro, G. M. (1983). Improved likelihood ratio statistics for generalized linear models, *J. Roy. Statist. Soc. Ser. B*, **45**, 404–413.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion), *J. Roy. Statist. Soc. Ser. B*, **49**, 1–39.

Fuller, W. A. (1987). *Measurement Error Models*, Wiley, New York.

Kendall, M. and Stuart, A. (1979). *The Advanced Theory of Statistics*, Vol. 2, Griffin, London.

Lawley, D. N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria, *Biometrika*, **43**, 295–303.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.

Wong, M. Y. (1989). Likelihood estimation of a simple regression model when both variables have error, *Biometrika*, **76**, 141–148.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.