RT-MAE 9425

ON SCORE TESTS IN STRUCTURAL REGRESSION MODELS

by

R. B. Arellano-Vale
and
H. Bolfarine

Palavras Chaves: Orthogonal parametrizations; score statistics; structural normal and elliptical models.

Classificação AMS : 62F05, 62J05.
(AMS Classification)

- Dezembro de 1994 -

ON SCORE TESTS IN STRUCTURAL REGRESSION MODELS

R.B. Arellano-Valle and H. Bolfarine
Departamento de Estatística -IME
Universidade de São Paulo - Caixa Postal 20570
CEP 01452-990 - São Paulo - SP - BRASIL

Summary

In this paper we investigate the distribution of the score statistics for testing hypothesis about the slope parameter in a simple structural regression model. It is shown that for two of the most commom ways of making the model identifiable, the distribution of the score statistics under the null hypothesis can be found exactly as an increasing function of an F statistics, providing thus exact test statistics for testing hypothesis about the slope parameter. This property seems not to be shared by the likelihood ratio statistics. Use is made of orthogonal parametrizations obtained in the literature. Generalizations to an elliptical structural model are also investigated.

Key Words: orthogonal parametrizations; score statistics; structural normal and elliptical models

1. Introduction

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

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

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

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

$k = 1, \ldots, n$, 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), $k = 1, \ldots, n$, a typically made assumption. An extension of the above normality assumptions is also contemplated. As is well known, there are problems with the estimation of the parameters in both cases. In the funcional 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 assumptions 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 type corrections for the likelihood ratio statistics for hypothesis involving β are proposed. For the case where λ_x is assumed known, Bartlett type corrections are considered in Arellano-Valle and Bolfarine (1994). These results seem to indicate that the likelihood ratio statistics, being λ_x or λ_e known, does not have treatable exact distribution. Thus, hypothesis testing for the slope parameter using the likelihood ratio statistics has to be performed aproximately.

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 score statistics under the null hypothesis in both cases. We show that the score statistics can be written as an increasing function of an F statistics meaning that, by using the score statistics, exact tests are obtained for testing hypothesis about the slope parameter, a property which seems not to hold for the likelihood ratio statistics. This result, to the best of the authors knowledge, seems to have been unnoticed in the literature. Using orthogonal parametrizations is crucial in the derivations of the results.

Section 2 presents a general matrix representation for the model and the orthogonal parametrization under both assumptions and normality. Section 3 discusses some properties of the maximum likelihood estimators. Section 4 investigates the distribution of the score statistics under normality. Section 5 discusses possible extensions to an elliptical structural model.

2. Orthogonal parametrizations

Note that we may rewrite model (1.1) as

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

where

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

$k = 1, \ldots, n$. Thus, from (2.1), we have that $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ are independent and identically distributed with $\mathbf{Z}_k \sim N_2(\mu; \Sigma)$, where

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

and

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

Further, it can be shown that

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

Let

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

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

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

where $\mu(\theta)$ e $\Sigma(\theta)$ are as given in (2.2) and (2.3), respectively.

Let $\mathbf{K}(\theta) = [\kappa_{i,j}]$ denote the expected information matrix. Then, after some algebraic manipulations, it can be shown that

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

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

$$\kappa_{i,j} = \begin{cases} n\frac{\partial\mu'}{\partial\theta_i}\Sigma^{-1}\frac{\partial\mu}{\partial\theta_j}, & \text{if } i = 1,2,5,\ j = 1,2, \\ \frac{n}{2}\mathrm{tr}(\Sigma^{-1}\frac{\partial\Sigma}{\partial\theta_i}\Sigma^{-1}\frac{\partial\Sigma}{\partial\theta_j}), & \text{if } i = 3,4,5 \text{ and } j = 3,4, \\ 0, & \text{if } i = 1,2 \text{ and } j = 3,4, \end{cases}$$

and

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

where $\mathrm{tr}(\mathbf{A})$ denotes the trace of the matrix $\mathbf{A}$. It follows that $\kappa_{\beta,j} \neq 0$, whatever be θ_j. 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 transformation of θ, that is, 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 (Cox and Reid, 1987):

$$(2.7)\qquad \sum_{i=1}^{4}\kappa_{i,j}\frac{\partial\theta_i}{\partial\beta} = -\kappa_{\beta,j},$$

$j = 1,2,3,4$. Typically, solving a systhem like the one in (2.7) is not simple. Moreover, when solvable, such equations may not be always 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 μ orthogonal to Σ. 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 Σ.

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

$$\phi_1 = \alpha + \beta\mu_x, \quad \phi_2 = \mu_x, \quad \phi_4 = \sigma_u^2, \tag{2.8}$$

$$\phi_3 = \begin{cases} \lambda_x\beta^2\sigma_u^2 + (\lambda_x + 1)\sigma_e^2, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)\sigma_x^2 + \lambda_e\sigma_u^2, & \text{if } \lambda_e \text{ is known.} \end{cases} \tag{2.9}$$

Considering the above parametrization, we have that $\mu = \mu(\phi_L) = (\phi_1, \phi_2)'$ and

$$\Sigma = \Sigma(\phi_S) = \begin{cases} (\lambda_x + 1)^{-1} \begin{pmatrix} \phi_3 + (\lambda_x\beta)^2\phi_4 & (\lambda_x + 1)\lambda_x\beta\phi_4 \\ (\lambda_x + 1)\lambda_x\beta\phi_4 & (\lambda_x + 1)^2\phi_4 \end{pmatrix}, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)^{-1} \begin{pmatrix} \beta^2\phi_3 + \lambda_e^2\phi_4 & \beta(\phi_3 - \lambda_e\phi_4) \\ \beta(\phi_3 - \lambda_e\phi_4) & \phi_3 + \beta^2\phi_4 \end{pmatrix}, & \text{if } \lambda_e \text{ is known,} \end{cases} \tag{2.10}$$

where $\phi_L = (\phi_1, \phi_2)'$ (the location parameters) and $\phi_S = (\phi_3, \phi_4, \beta)'$ (the scale parameters). Note that $|\Sigma| = \phi_3\phi_4$. 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 news 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 $\text{tr}(\Sigma) = \phi_3 + \phi_4$, so that ϕ_3 and ϕ_4 are the caracteristic roots of Σ. In the sequel, we present some properties of the matrix Σ wich will make it easier to derive the cumulants of the derivatives of the log likelihood function $l = l(\phi)$.

Let

$$\mathbf{A} = \begin{pmatrix} \alpha_3' \\ \alpha_4' \end{pmatrix} = \begin{cases} (\lambda_x + 1)^{-1/2} \begin{pmatrix} 1 & 0 \\ \lambda_x\beta & \lambda_x \end{pmatrix}, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)^{-1/2} \begin{pmatrix} \beta & 1 \\ \lambda_e & -\beta \end{pmatrix}, & \text{if } \lambda_e \text{ is known,} \end{cases}$$

and note that

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

$i = 3, 4$, and

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

Similarly, if

$$\bar{\mathbf{A}} = \begin{pmatrix} \bar{\alpha}_3' \\ \bar{\alpha}_4' \end{pmatrix} = \begin{cases} (\lambda_x + 1)^{-1/2} \begin{pmatrix} \lambda_x + 1 & -\lambda_x\beta \\ 0 & 1 \end{pmatrix}, & \text{if } \lambda_x \text{ is known,} \\ (\beta^2 + \lambda_e)^{-1/2} \begin{pmatrix} \beta & \lambda_e \\ 1 & -\beta \end{pmatrix}, & \text{if } \lambda_e \text{ is known,} \end{cases} \tag{2.11}$$

then,

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

and

(2.12)
$$\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_i} = \phi_i^{-1}\bar{\alpha}_i\alpha_i',$$

since $\bar{\alpha}_i = \phi_i\Sigma^{-1}\alpha_i$, $i = 3, 4$. Note that $\mathbf{A}\bar{\mathbf{A}}' = \mathbf{I}$, $\bar{\mathbf{A}}\Sigma\bar{\mathbf{A}}' = \Phi$ and $\mathbf{A}\Sigma^{-1}\mathbf{A}' = \Phi^{-1}$, where $\Phi = \text{diag}(\phi_3, \phi_4)'$, that is, $\bar{\mathbf{A}}'(\mathbf{Z}_i - \mu) \overset{iid}{\sim} N_2(0, \phi)$, $i = 1, \ldots, n$. Furthermore, it is easy to see that

$$\frac{\partial\Sigma}{\partial\beta} = \phi_3\frac{\partial^2\Sigma}{\partial\phi_3\partial\beta} + \phi_4\frac{\partial^2\Sigma}{\partial\phi_4\partial\beta} = (\frac{\phi_3\phi_4}{\sigma_\beta^2})^{1/2}(\alpha_3\alpha_4' + \alpha_4\alpha_3')$$

and

(2.13)
$$\Sigma^{-1}\frac{\partial\Sigma}{\partial\beta} = (\frac{\phi_3\phi_4}{\sigma_\beta^2})^{1/2}(\phi_3^{-1}\bar{\alpha}_3\alpha_4' + \phi_4^{-1}\bar{\alpha}_4\alpha_3'),$$

where

(2.14)
$$\sigma_\beta^2 = \begin{cases} \frac{\phi_3}{\lambda_x^2\phi_4}, & \text{if } \lambda_x \text{ is known,} \\ (\frac{\beta^2+\lambda_e}{\phi_3-\lambda_e\phi_4})^2\phi_3\phi_4, & \text{if } \lambda_e \text{ is known.} \end{cases}$$

From (2.12) and (2.13) it follows that

(2.15)
$$\text{tr}(\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_i}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j}) = \begin{cases} \phi_i^{-2}, & i = j \\ 0, & i \neq j, \end{cases}$$

$i, j = 3, 4$, and

(2.16)
$$tr\{(\Sigma^{-1}\frac{\partial\Sigma}{\partial\beta})^2\} = \text{tr}(\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\beta^2}) = 2\sigma_\beta^{-2}.$$

Let now $\mathbf{K} = \mathbf{K}(\phi) = [\kappa_{i,j}]$ the information matrix under the orthogonal parametrization. Then,

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

where $\mu = \mu(\phi_L)$ and $\Sigma = \Sigma(\phi_S)$ are as defined above. Thus, $\mathbf{K} = \text{diag}(\mathbf{K}_L, \mathbf{K}_S)$, that is, $\mathbf{K}$ is a block diagonal matrix with $\mathbf{K}_L = n\Sigma^{-1}$ and, using (2.15) and (2.16), $\mathbf{K}_S = n\text{diag}(1/2\phi_3^2, 1/2\phi_4^2, 1/\sigma_\beta^2)$, where σ_β^2 is given in (2.14). Note that $\sigma_\beta^2 = \beta^2\sigma_{XX}\sigma_{YY.X}/\sigma_{YX}^2 = \beta^2(1-\rho_{YX}^2)/\rho_{YX}^2$, where $\sigma_{YX} = Cov[Y_i, X_i]$, $\sigma_{XX} = Var[X_i]$, $\sigma_{YY.X} = Var[Y_i|X_i]$ and $\rho_{YX} = \sigma_{YX}/(\sigma_{XX}\sigma_{YY})^{1/2}$, which denotes the correlation between Y_i and X_i, $i = 1, \ldots, n$.

3. Maximum likelihood estimators and properties

The log likelhood function $l = l(\phi)$ with respect to the orthogonal transformations ϕ given in (2.8) and (2.9) may be written as in (2.6), with μ replaced by $\mu(\phi_L)$ and Σ replaced by $\Sigma(\phi_S)$. The maximum likelihood estimator $(\hat{\phi}_L, \hat{\phi}_S)$ of (ϕ_L, ϕ_S) is obtained by solving the equation $\hat{\mu} = \mu(\hat{\phi}_L) = \bar{\mathbf{Z}}$, from where we get $\hat{\phi}_1 = \bar{Y}$, and $\hat{\phi}_2 = \bar{X}$, since $\mu = (\phi_1, \phi_2)'$ and $\bar{\mathbf{Z}} = \sum_{k=1}^{n} \mathbf{Z}_k/n = (\bar{Y}, \bar{X})'$. Using the above estimators, it follows that the maximum likelihood estimators of β, ϕ_3 and ϕ_4 are obtained by solving $\hat{\phi}_i = \bar{a}_i'(\hat{\beta})\mathbf{S}_{ZZ}\bar{a}_i(\hat{\beta})$, $i = 3, 4$, and $\bar{a}_3'(\hat{\beta})\mathbf{S}_{ZZ}\bar{a}_4(\hat{\beta}) = 0$, where $\bar{a}_i(\hat{\beta})$ is as defined in (2.12), with β replaced by $\hat{\beta}$ and

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

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 (Bolfarine and Cordani, 1993) when λ_x is known that

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

$$\hat{\phi}_4 = \frac{S_{XX}}{\lambda_x + 1},$$

$$\hat{\beta} = (\frac{\lambda_x + 1}{\lambda_x})\frac{S_{YX}}{S_{XX}},$$

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

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

where $S_{YY.X} = S_{YY} - S_{XX}^{-1}S_{YX}^2 = S_{YY}(1 - r_{YX}^2)$, and $r_{YX} = S_{YX}/(S_{YY}S_{XX})^{1/2}$. When λ_e is known, it follows (Wong, 1989) that

$$\hat{\phi}_3 = \frac{\hat{\beta}^2 S_{YY} + 2\lambda_e\hat{\beta}S_{YX} + \lambda_e^2 S_{XX}}{\hat{\beta}^2 + \lambda_e},$$

$$\hat{\phi}_4 = \frac{S_{YY} - 2\hat{\beta}S_{YX} + \hat{\beta}^2 S_{XX}}{\hat{\beta}^2 + \lambda_e},$$

$$\hat{\beta} = (\frac{S_{YY} - \lambda_e S_{XX}}{2S_{YX}}) + \{(\frac{S_{YY} - \lambda_e S_{XX}}{2S_{YX}})^2 + \lambda_e\}^{1/2}.$$

Under model (1.1), it follows that $\hat{\mu} = \bar{\mathbf{Z}} \sim N_2(\mu, \frac{1}{n}\Sigma)$, and

$$\hat{\Sigma} = \mathbf{S}_{ZZ} \sim W_2(\frac{1}{n}\Sigma, n - 1), \tag{3.1}$$

are independent, with Σ as in (2.10). Here, $W_k(\mathbf{A}, m)$ denotes the k-variate Wishart distribution with dispersion matrix $\mathbf{A}$ 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 $\hat{\phi}_L = \bar{\mathbf{Z}} \sim N_2(\phi_L, \Sigma/n)$. Thus, confidence regions or hypothesis testing for ϕ_L or functions of the form $\mathbf{a}'\phi_L$ where $\mathbf{a}$ is known are easily obtained.

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. For example, when $\lambda_e = 1$, $\hat{\phi}_3$ and $\hat{\phi}_4$ are the proper values of the Wishart matrix $\Sigma(\hat{\phi}_S)$. Thus, in this situation, the distribution function of $(\hat{\phi}_3, \hat{\phi}_4)$ can be represented in terms of infinite series (Muirhead, 1982). Similarly, the distribution function of $\hat{\beta}$ can be given in a form of 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 $\mathbf{x} = (x_1, \ldots, x_n)'$ 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 distributions of certain functions of $\hat{\phi}_S$ in both cases (λ_x or λ_e known). Thus, considering the Wishart distribution in (3.1) it follows that (Muirhead, 1982)

(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 $\sigma_{YY.X} = \sigma_{YY} - \frac{\sigma_{YX}^2}{\sigma_{XX}} = \sigma_{YY}(1 - \rho_{YX}^2)$;

(iii) $(\frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}})|S_{XX} \sim N(0, \frac{\sigma_{YY.X}}{nS_{XX}})$;

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

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

From the above results, it follows that

(3.2)
$$\left(\frac{(n-2)S_{XX}}{S_{YY.X}}\right)^{1/2}\left(\frac{S_{YX}}{S_{XX}} - \frac{\sigma_{YX}}{\sigma_{XX}}\right) \sim t(0, 1; n-2).$$

Notice that (v) follows from (iii) and (iv), since, as is well known, by mixing a normal with a chisquared distribution we get a t-distribution.

4. The score 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 $\tilde{\phi}_L = \hat{\phi}_L = \bar{\mathbf{Z}}$ and $\tilde{\phi}_S = (\tilde{\phi}_3, \tilde{\phi}_4, \tilde{\beta})'$ follows from the equations

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

and

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

where $\bar{\alpha}_i(\beta_0)$, $i = 3, 4$, are as defined in (2.12), 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

$$n\mathbf{S}_{ZZ} \sim W_2(\Sigma_0, n-1),$$

where Σ_0 is the same as Σ (defined in (2.10)), but evaluated at $(\phi_3, \phi_4, \beta_0)$. This implies that

$$\frac{n\tilde{\phi}_i}{\phi_i} \sim \chi^2_{n-1}, \tag{4.1}$$

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

$$S = \frac{1}{\tilde{\kappa}_{\beta,\beta}}\left(\frac{\partial l(\phi)}{\partial \beta}|_{\phi=\tilde{\phi}}\right)^2 = n\frac{(\bar{\alpha}_3'(\beta_0)\mathrm{S}_{ZZ}\bar{\alpha}_4(\beta_0))^2}{\tilde{\phi}_3\tilde{\phi}_4},$$

Let $\tilde{\Phi} = [\tilde{\phi}_{ij}]$, the mat*rix defined by

$$\tilde{\Phi} = \bar{\mathrm{A}}_0\mathrm{S}_{ZZ}\bar{\mathrm{A}}_0',$$

where $\bar{\mathrm{A}}_0' = \bar{\mathrm{A}}'(\beta_0)$, that is,

$$\tilde{\phi}_{ij} = \bar{\alpha}_i'(\beta_0)\mathrm{S}_{ZZ}\bar{\alpha}_j(\beta_0),$$

$i, j = 3, 4$, with $\tilde{\phi}_{ii} = \tilde{\phi}_i$. Clearly, the score statistics S can be written as

$$S = n\frac{\tilde{\phi}_{34}^2}{\tilde{\phi}_3\tilde{\phi}_4},$$

so that the distribution of S under H_0 is determined by the distribution of $\tilde{\Phi}$. Moreover, under $H_0 : \beta = \beta_0$, it follows that

$$\tilde{\Phi} \sim W_2(\frac{1}{n}\Phi_0; n-1),$$

where W_2 denotes the Wishart distribution and

$$\Phi_0 = \bar{\mathrm{A}}_0\Sigma_0\bar{\mathrm{A}}_0' = \mathrm{diag}(\phi_3, \phi_4),$$

where, as before, diag means diagonal matrix. Using (3.2) with $\Sigma_{ZZ} = [S_{ij}]$ and $\Sigma = [\sigma_{ij}]$ replaced by $\tilde{\Phi} = [\tilde{\phi}_{ij}]$ and $\Phi_0 = [\phi_{ij}]$, respectively, we have that

$$t^2 = \{\frac{(n-2)\tilde{\phi}_4}{\tilde{\phi}_{3.4}}\}^{1/2}(\frac{\tilde{\phi}_{34}}{\tilde{\phi}_4} - \frac{\phi_{34}}{\phi_4}) \sim t(0, 1; n-2), \tag{4.2}$$

where

$$\tilde{\phi}_{3.4} = \tilde{\phi}_3 - \tilde{\phi}_{34}^2\tilde{\phi}_4^{-1} = \tilde{\phi}_3(1 - S/n). \tag{4.3}$$

From (4.2), (4.3) and from the fact that $\phi_{34} = 0$, it follows that

$$S = n(\frac{t^2}{n-2})(1 + \frac{t^2}{n-2})^{-1},$$

so that, under H_0,

$$n^{-1}S \sim Beta(\frac{1}{2}, \frac{(n-2)}{2}),$$

for both λ_z and λ_e known.

5. An elliptical structural model

As a possible extension of the normal model considered in the previous sections, we consider $\mathbf{Z} = (\mathbf{Z}_1', \ldots, \mathbf{Z}_n')'$, where $\mathbf{Z}_k = (Y_k, X_k)'$ with

$$\mathbf{Z} \sim El_{2n}(\mathbf{1}_n \otimes \mu, \mathbf{I}_n \otimes \Sigma), \tag{5.1}$$

where, as in Fang et al. (1990), $El_{2n}(\mathbf{a}, \mathbf{A})$ denotes a $2n$-dimensional elliptical variate with parameters $\mathbf{a}$ and $\mathbf{B}$, $\mathbf{1}_n$ is an n-dimensional vector of ones, $\mathbf{I}_n$ the n-dimensional identity matrix and μ and Σ are as in (2.2) and (2.3). Assuming that $\mathbf{Z}$ has a density given by

$$p(\mathbf{z}|\theta) = |\Sigma|^{-n/2} f(\sum_{k=1}^{n}(\mathbf{z}_k - \mu)'\Sigma^{-1}(\mathbf{z} - \mu)), \tag{5.2}$$

where $f(\mathbf{t}'\mathbf{t})$ is a $2n$-dimensional spherical density, the density of $\mathbf{T} = (\mathbf{T}_1', \ldots, \mathbf{T}_n')' \sim El_{2n}(0, \mathbf{I}_{2n})$, where $\mathbf{T}_k = \Sigma^{-1/2}(\mathbf{Z}_k - \mu)$, $k = 1, \ldots, n$. Note that (5.1) implies that $\mathbf{Z}_k \sim El_2(\mu, \Sigma)$ and provided they exists, $E[\mathbf{Z}_k] = \mu$ and $Var[\mathbf{Z}_k] = \alpha_f \Sigma$, with $Cov[\mathbf{Z}_k, \mathbf{Z}_l] = 0$, $k \neq l$, where $\alpha_f = -2\psi'(0)$, being $\psi(.)$ the characteristic function of $\mathbf{T}_k$ defined above. However, $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ are not independent unless $\mathbf{Z}$ is normally distributed. Moreover, given f, (5.2) implies that the density $\mathbf{Z}$ is characterized by $\mu = \mu(\theta)$ and $\Sigma = \Sigma(\theta)$ only. Thus, as the normal model defined in Section 2, the elliptical model defined above is not identifiable. However, as in the normal model, the assumptions λ_z or λ_e known make the elliptical model identifiable. Under these assumptions, the orthogonal transformations considered in Section 2 for the normal model also works in this more general set up. Thus, working with the orthogonal transformation given in in (2.8) and (2.9), and using some well known facts about elliptical distributions such as $||\mathbf{T}||$ and $\mathbf{T}/||\mathbf{T}||$ are independent and $\mathbf{T}/||\mathbf{T}|| \stackrel{d}{=} \mathbf{e}/||\mathbf{e}||$, with $\mathbf{e} \sim N_{2n}(\mu, \mathbf{I}_{2n})$ ("$\stackrel{d}{=}$" meaning distributed as) we can show that the information matrix $\mathbf{K} = [\kappa_{ij}]$, is such that

$$\kappa_{ij} = n\{\frac{4a_f(1,2)}{p}\frac{\partial\mu'}{\partial\phi_i}\Sigma^{-1}\frac{\partial\mu}{\partial\phi_j} + \frac{2a_f(2,2)}{p(p+2)}\mathrm{tr}(\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_i}\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j})$$
$$+\frac{n}{4}(\frac{4a_f(2,2)}{p(p+2)} - 1)\mathrm{tr}(\Sigma^{-1}\frac{\partial\Sigma}{\phi_i})\mathrm{tr}(\Sigma^{-1}\frac{\partial\Sigma}{\partial\phi_j})\},$$

where $p = 2n$ is the dimension of $\mathbf{T}$, $a_f(i,j) = E[||\mathbf{T}||^{2i}W_f^j(||\mathbf{T}||^2)]$, with $W_f(u) = \partial\log f(u)/\partial u = f'(u)/f(u)$. We call attention to the fact that W_f and a_f are independent of ϕ since the distribution of $\mathbf{T}$ is independent of (μ, Σ) and thus of ϕ. In particular, it can be shown that

$$\kappa_{\beta,\beta} = \frac{4a_f(2,2)}{p(p+2)}(\frac{\sigma_\beta^2}{n})^{-1}, \tag{5.3}$$

where $\sigma_\beta^2 = \sigma_\beta^2(\phi)$ is as in (2.14).

As in Anderson et al. (1986), it follows that maximum likelihood estimators of β and ϕ are obtained by solving the equations $\hat{\mu} = \hat{\mu}_N = \bar{\mathbf{Z}} = \sum_{k=1}^n \mathbf{Z}_k/n$ and $\hat{\Sigma} = \hat{\Sigma}_N = \mathbf{S}_{ZZ} = p\hat{\Sigma}_N/u_f$, where $\mathbf{S}_{ZZ} = \sum_{k=1}^n (\mathbf{Z}_k - \bar{\mathbf{Z}})(\mathbf{Z}_k - \bar{\mathbf{Z}})'/n$, $\hat{\mu}_N$ and $\hat{\Sigma}_N$ are the maximum likelihood estimators under the normal model given in Section 3 and u_f is the maximum of the function $u^{p/2}f(u)$, $u > 0$ and $p = 2n$ is the dimension of $\mathbf{Z}$. Thus, if $\hat{\beta}_N$ and $\hat{\phi}_{i,N}$ $i = 1, \ldots, 4$ denotes the maximum likelihood estimators of β and ϕ_i, $i = 1, \ldots, 4$, under the normal model, it follows that the maximum likelihood estimators under the elliptical model defined by (5.1) are such that

$$\hat{\beta} = \hat{\beta}_N, \quad \hat{\phi}_i = \hat{\phi}_{i,N}, \quad i = 1, 2, \quad \text{and} \quad \hat{\phi}_i = \frac{p}{u_f}\hat{\phi}_{i,N}, \quad i = 3, 4.$$

The derivations of maximum likelihood estimators under the hypothesis $H : \beta = \beta_0$ is similar. Additionally, using some results in Fang et al. (1990), it can be shown that several statistical properties which hold under the normal model also hold under elliptical models. In particular, $\hat{\beta} \stackrel{d}{=} \hat{\beta}_N$ since $\hat{\beta}(a\mathbf{z}) = \hat{\beta}(\mathbf{z})$, for any real a. We show next that this is the case with the score statistics which, except for a constant, is the same as under the normal model.

Using the fact that

$$\frac{\partial \log|\Sigma|}{\partial \beta} = \text{tr}(\Sigma \frac{\partial \Sigma}{\partial \beta}) = 0,$$

it follow from (5.2) with $l = \log p(\mathbf{z}|\phi)$ that

$$\frac{\partial l}{\partial \beta} = -W_f\left(n\text{tr}(\Sigma^{-1}\mathbf{S}_{ZZ}) + n(\bar{\mathbf{Z}} - \mu)'\Sigma^{-1}(\bar{\mathbf{Z}} - \mu)\right)\left\{n\text{tr}(\mathbf{B}\mathbf{S}_{ZZ}) + n(\bar{\mathbf{Z}} - \mu)'\mathbf{B}(\bar{\mathbf{Z}} - \mu)\right\},$$

where

$$\mathbf{B} = \Sigma^{-1}\frac{\partial \Sigma}{\partial \beta}\Sigma^{-1} = \frac{1}{\sigma_\beta\sqrt{\phi_3\phi_4}}(\tilde{\alpha}_3\tilde{\alpha}_4' + \tilde{\alpha}_4\tilde{\alpha}_3').$$

Using (5.3) and the fact that $\tilde{\mu} = \tilde{\mu}_N = \bar{\mathbf{Z}}$ and $\tilde{\Sigma} = (p/u_f)\tilde{\Sigma}_N$ it follows that

$$S = \frac{1}{\tilde{\kappa}_{\beta,\beta}}\left(\frac{\partial l(\phi)}{\partial \beta}\Big|_{\phi=\tilde{\phi}}\right)^2 = \frac{p(p+2)}{a_f(2,2)}\left(\frac{f'(u_f)}{f(u_f)}\right)^2 S_N,$$

where $p = 2n$ and S_N is the score statistics under the normal model derived in Section 4. In the special case of the structural elliptical t-model, that is, $\mathbf{Z} \sim t_k(\mathbf{1}_n \otimes \mu, \mathbf{I}_n \otimes \Sigma; \nu)$ we have that

$$f(u) = \frac{\Gamma[\frac{\nu+p}{2}]}{\Gamma[\frac{\nu}{2}]\pi^{p/2}}\nu^{\nu/2}\{\nu + u\}^{-(\nu+p)/2}, \quad u \geq 0,$$

it follows that $u_f = p = 2n$, for $\nu > 0$ (including $\nu = \infty$),

$$W_f(u) = \frac{f'(u)}{f(u)} = -(\frac{\nu+p}{2})(\nu+u)^{-1} \quad \text{and} \quad a_f(2,2) = \frac{p(p+2)}{4}\frac{\nu+p}{\nu+p+2},$$

so that

$$S = \frac{\nu + 2n + 2}{\nu + 2n} S_N.$$

Reference

Anderson, T.W., Fang, K.T. and Hsu, H. (1986). Maximum likelihood estimates and likelihood ratio criteria for multivariate elliptical contoured distributions. *The Canadian Journal of Statistics*, 44,55-59.

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

Arellano-Valle, R.B. and Bolfarine, H. (1994). A note on the simple structural regression model. (submitted for publication).

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

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

Fang, K.T., Kotz, S. and Ng, K.W. (1990). *Symmetrical multivariate and related distributions.* Chapman and Hall. London.

Fuller, W.A. (1987). *Measurement error models*, Wiley, New York.

Kendall, M. and Stuart, (1979). *The advanced theory of Statistics*, Vol. 2, Griffin, London.

Muirhead, R. (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.

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLIC...

9401 - PAULA, G. A., SEN P.K., Tests of ordered Hypotheses in Linkage in Heredity, 12p.

9402 - PAULA, G. A., SEN P. K. One-Sided Tests in Generalized Linear Models with Parallel Regression Lines, 18p.

9403 - PAULA, G. A., .Influence Diagnostics in Proper Dispersion Models,22 p.

9404 - SALINAS-TORRES, V.H. & PEREIRA, C.A.B. and TIWARI, R.C. - Convergence of Dirichlet Measures Arising in Context of Bayesian Analysis of Competing Risk Models, 17p.

9405 - CRIBARI-NETO, F. & FERRARI, S.L.P. - Second Order Asymptotics for Score Tests in Generalized Linear Models, 25p.

9406 - ZUAZOLA, P.L.I., PEREIRA, C.A.B. & WECHSLER, S. - Predictivistic Statistical Inference in Finite Populations, 10p.

9407 - BOLFARINE, H. & ROJAS, M.G. - Maximum Likelihood Estimation of Simultaneous Pairwise Linear Structural Relationships, 19p.

9408 - FERRARI, P.A. & KIPNIS, C. - Second Class Particles in the Rarefaction Fan, 9p.

9409 - ANDRÉ, C.D.S., NARULA, S.C. & PERES, C.A. - Asymptotic Properties of MSAE Estimates in Multistage Model. Preliminary Results, 19p.

9410 - ARELLANO-VALLE, R.B. & BOLFARINE, H. - On Some Characterizations of the t-Distribution, 12p.

9411 - NEVES, E.J., A Discrete Variational Problem Related to Ising Droplets at Low Temperatures, 16p.

9412 - FERRARI, S.L.P. & CORDEIRO, G.M. - Corrected Score Tests for Exponential Family Nonlinear Models, 14p.

9413 - ANDRADE, D.F. & SINGER, J.M. - Profile Analysis for Randomized Complete Block Experiments, 8p.

9414 - BOLFARINE, H. & ROJAS, M.G. - On Structural Comparative Calibration, 13p.

9415 - AUBIN, E.C.Q. & CORDEIRO, G.M. Bias in Linear Regression Models with Unknown Covariance Matrix, 16p.

9416 - ELIAN, S.N. Previsão do Vetor de Elementos não observados numa população finita, 17p.

9417 - ARELLANO, R.B. & BOLFARINE, H. A Note on the Simple Structural Regression Model, 14p.

9418 - CORDEIRO, G.M. & FERRARI, S.L.P. Bartlett-Type Corrections for some score tests in Proper Dispersion Models, 16p.

9419 - FERRARI,P.A. & FONTES, L.R.G. Poissonian Approximation for the Tagged Particle Asymmetric Simple Exclusion, 16p.

9420 - BOLFARINE, H. & ROJAS M.G. On Structural Comparative Calibration Under a t-Model, 20p.

9421 - PAULA, G.A. Influence and Residuals in Restricted Generalized Linear Models, 26p.

9422 - FONTES, L.R.G. & NEVES, E.J. Phase Uniqueness and Correlation Length in Diluted-Field Ising Models, 13p.

9423 - BENJAMINI, I., FERRARI, P.A. & LANDIM, C. Asymmetric Conservative Processes with Random Rates, 28p.

9424 - ARELLANO-VALE, R.B., BOLFARINE, H., VILCA-LABRA, F. Ultrastructural Elliptical Models, 16p.

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

- Departamento de Estatística
IME-USP
Caixa Postal 20.570
01452-990 - São Paulo, Brasil