ESTIMATION OF PARAMETERS IN AUTOREGRESSIVE MODELS

Özlem Türker
Department of Statistics, METU
Ankara, Turkey
e-mail: ozlemt@metu.edu.tr

ABSTRACT

Autoregressive models have many applications in business and economics. In this paper, we consider two regressive models

$$Y_{i,t} = \mu_i + \delta_i X_{i,t} + e_{i,t} \quad (i = 1, 2 ; t = 1, 2, ..., n_i)$$

where the random errors $e_{i,t}$ are autocorrelated, i.e.,

$$e_{i,t} = \phi_i e_{i,t-1} + a_{i,t} \ , \quad |\phi_i| < 1;$$

$a_{i,t}$ are iid random errors. The autoregression coefficicents ϕ_i (i = 1, 2) may or may not be equal. The problem is to estimate μ_i , δ_i and ϕ_i and $\sigma_i^2 = V(a_{i,t})$; the variances σ_i^2 (i = 1, 2) may or may not be equal.

Traditionally, the random errors a_t have been assumed to be normal $N(0, \sigma^2)$. There is now a realization that non-normal distributions are more prelavent in practice. We consider non-normal distributions and derive efficient estimators by using the methodology of modified likelihood. We also give a test for H_0: $\delta_1 = \delta_2$. Both situations are considered when X_t $(1 \le t \le n)$ are fixed design points and when they change with Y_t $(1 \le t \le n)$.

1. Introduction

The main problem in a simple regression model with autocorrelated errors is the estimation of the coefficients. Most of the literature on this topic has hinged on the assumption of normality; see, for example, Anderson (1949), Cochrane and Orcutt (1949), Durbin (1960), Tiao and Tan (1966), Gallant and Goebel (1967), Beach and Mackinnon (1978), Kramer (1980), Magee et al. (1987), Velu and Gregory (1987), Dielman and Pfaffenberger (1989), Maller (1989), Cogger (1990), Weiss (1990), Schäffler (1991) and Nagaraj et al. (1992). However, there is now a realization that non-normal distributions are more prevalent in practice (Pearson 1932; Elveback et al. 1970; Tukey 1970; Tse 1991) than the normal. The methodology of maximum likelihood (ML) does indeed extend in principle to non-normal distributions but there are difficulties. Due to enormous mathematical complexity with the maximum likelihood methodology, extensions to nonnormal distributions have not been possible. Therefore, least squares (LS) methodology has been widely used in this context. Recently, however, Tiku et al. (1999) and Akkaya and Tiku (2001) worked out these extensions by using the modified likelihood methodology (MML). They showed, in particular, that the LS estimators have low efficiency and their relative efficiency (as compared to the MML estimators) decreases as the sample size n increases. This is indeed a disconcerting feature of the LS estimatiors. Tiku et al. (1999) and Akkaya and Tiku (2001) results, however, are available for a single source of information. However, in practice, several independent sources of information may be available so that several models have to be considered simultaneously. In such systems of linear regression equations, the main interest is in testing whether the parameter vector is the same for all equations or not. In this paper, two linear regression models are considered in detail and estimation and hypothesis tests are developed.

2. Autoregressive Models

Consider two simple regressive models

$$y_{i,t} = \mu_i^{/} + \delta_i x_{i,t} + e_{i,t}$$

$$e_{i,t} = \phi_i e_{i,t-1} + a_{i,t} \qquad (1 \le t \le n_i, i = 1, 2, |\phi_i| < 1) \qquad (1)$$

where $a_{i,t}$ ($1 \le t \le n_i$; $i = 1, 2$) are individually and jointly iid random errors having equal or different variances σ_1^2 and σ_2^2, respectively. Here $y_{i,t}$ ($1 \le t \le n_i$) are the observed value of a random variable Y_i ($i = 1, 2$) at time t and $x_{i,t}$ ($1 \le t \le n_i$) (i) are design variables and pre-determined as in controlled experiments in agricultural, biological and engineering sciences, or (ii) change with $y_{i,t}$ as in business and economics. In case (ii), the full likelihood is $L = L_{Y|X} L_X$ but is difficult to handle mathematically. Instead, we work with $L_{Y|X}$ and regard $x_{i,t}$ fixed as is usually done in practice.

Model (1) can alternatively be written as

$$y_{i,t} - \phi_i y_{i,t-1} = \mu_i + \delta_i (x_{i,t} - \phi_i x_{i,t-1}) + a_{i,t} \qquad (|\phi_i| < 1, \ i = 1, 2) \qquad (2)$$

which is nonlinear because of the parameter $\delta_i\phi_i$ and is mathematically more complex than the simple linear model, i.e., model (2) with $\phi_i = 0$ ($i = 1, 2$). The autoregressive model has many applications. For example, in predicting future stock prices, the effect of an innovation might persist for some time. Numerous other applications of the model are in agricultural, biological and biomedical sciences besides business and economics; see, for example, Anderson (1949), Durbin (1960), Tiao and Tan (1966), Beach and Mackinnon (1978), Cogger (1990), Weiss (1990) and Schäffler (1991).

Although our technique can be used for any location-scale distribution of the type $(1/\sigma)\, f((y - \mu) / \sigma)$, for illustration, we consider the family of Generalized Logistic

$$GL(b,\sigma): f(a) = \frac{b}{\sigma}\frac{e^{-a/\sigma}}{\left(1+e^{-a/\sigma}\right)^{b+1}} \qquad , -\infty < a < \infty .$$

This family of distributions is very flexible: (i) for $b < 1$, it represents negatively skewed distributions, (ii) for $b = 1$, it is the well-known logistic distribution and is symmetric and close to a normal distribution, and (iii) for $b > 1$, it represents positively skewed distributions. We assume that b is known, and take $n_1 = n_2 = n$ for simplicity in presentation. In practice b is determined by using a Q-Q plot.

The initial values (y_0 , x_0) are generated by using the Vinod-Shenton Models A or B. Since their Model B is more general, we will use this model.

The conditional (given y_0 and x_0) likelihood function is

$$L = \left(\frac{b}{\sigma}\right)^{2n} \exp\left\{-\left(\sum_{t=1}^{n} z_{1t} + \sum_{t=1}^{n} z_{2t}\right)\right\}\left[\prod_{t=1}^{n}\left(1+e^{-z_{1t}}\right)^{b+1}\prod_{t=2}^{n}\left(1+e^{-z_{2t}}\right)^{b+1}\right]^{-1}$$

where $z_{ij} = \{(y_{ij} - \phi_i y_{ij-1}) - \mu_i - \delta_i(x_{ij} - \phi_i x_{ij-1})\}/\sigma_i$.

The ML estimators are the solutions of the following equations:

$$\frac{\partial \ln L}{\partial \mu_i} = \frac{n}{\sigma_i} - \frac{b+1}{\sigma_i}\sum_{j=1}^{n} g_i\left(z_{ij}\right) = 0 \qquad (i = 1, 2) \tag{3}$$

$$\frac{\partial \ln L}{\partial \delta_i} = \frac{1}{\sigma_i}\sum_{j=1}^{n}\left(x_{i[j]} - \phi_i x_{i[j]-1}\right) - \frac{b+1}{\sigma_i}\sum_{j=1}^{n}\left(x_{i[j]} - \phi_i x_{i[j]-1}\right) g_i\left(z_{ij}\right) = 0 \quad (i = 1, 2) \tag{4}$$

$$\frac{\partial \ln L}{\partial \phi_i} = \frac{1}{\sigma_i}\sum_{j=1}^{n}\left(y_{i[j]-1} - \delta_i x_{i[j]-1}\right) - \frac{b+1}{\sigma_i}\sum_{j=1}^{n}\left(y_{i[j]-1} - \delta_i x_{i[j]-1}\right) g_i\left(z_{ij}\right) = 0 \quad (i = 1, 2) \tag{5}$$

$$\frac{\partial \ln L}{\partial \sigma_i} = -\frac{n}{\sigma_i} + \frac{1}{\sigma_i}\sum_{j=1}^{n} z_{i(j)} - \frac{b+1}{\sigma_i}\sum_{j=1}^{n} z_{i(j)} g_i\left(z_{ij}\right) = 0 \qquad (i = 1, 2) \tag{6}$$

where $g_i(z_{ij}) = e^{-z_{ij}} \big/ \left(1+e^{-z_{ij}}\right)$.

Solving these equations is very problematic indeed; see, for example, Pearson and Hartley (1972, p.87-9), Barnett (1966), Lee et al. (1980), Tiku et al. (1986), Puthenpura and Sinha (1986), Tiku and Suresh (1992) and Vaughan (1992, 2002).

3. Modified Likelihood

To obtain efficient and robust, and explicit, estimators, we express the equations (3)-(6) in terms of the order statistics $y_{(i)}$ ($1 \le i \le n$) and the corresponding concomitants $x_{[i]}$; see Akkaya and Tiku (2001), Tiku et al. (1999). The linearized equations are called modified likelihood equations and are given by

$$\frac{\partial \ln L}{\partial \mu_i} \cong \frac{\partial \ln L^*}{\partial \mu_i} = \frac{n}{\sigma} - \frac{b+1}{\sigma}\sum_{j=1}^{n}\left(\alpha_{ij} - \beta_{ij} z_{i(j)}\right) = 0 \qquad (i = 1, 2) \tag{7}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \delta_i} \cong \frac{\partial \ln L^*}{\partial \delta_i} &= \frac{1}{\sigma}\sum_{j=1}^{n}\left(x_{i[j]} - \phi_i x_{i[j]-1}\right) \\ &\quad - \frac{b+1}{\sigma}\sum_{j=1}^{n}\left(x_{i[j]} - \phi_i x_{i[j]-1}\right)\left(\alpha_{ij} - \beta_{ij} z_{i(j)}\right) = 0 \end{aligned} \qquad (i = 1, 2) \tag{8}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \phi_i} \cong \frac{\partial \ln L^*}{\partial \phi_i} &= \frac{1}{\sigma}\sum_{j=1}^{n}\left(y_{i[j]-1} - \delta_i x_{i[j]-1}\right) \\ &\quad - \frac{b+1}{\sigma}\sum_{j=1}^{n}\left(y_{i[j]-1} - \delta_i x_{i[j]-1}\right)\left(\alpha_{ij} - \beta_{ij} z_{i(j)}\right) = 0 \end{aligned} \qquad (i = 1, 2) \tag{9}$$

and

$$\frac{\partial \ln L}{\partial \sigma_i} \cong \frac{\partial \ln L^*}{\partial \sigma_i} = -\frac{n}{\sigma_i} + \frac{1}{\sigma_i}\sum_{j=1}^{n} z_{i(j)} - \frac{b+1}{\sigma_i}\sum_{j=1}^{n} z_{i(j)}\left(\alpha_{ij} - \beta_{ij} z_{i(j)}\right) = 0 \quad (i = 1, 2). \tag{10}$$

The solutions of these equations are the MML estimators:

$$\hat{\delta}_i = G_i - H_i\, \hat{\sigma} \tag{11}$$

$$\hat{\phi}_i = K_i - D_i \hat{\sigma} \qquad (i=1, 2) \tag{12}$$

$$\hat{\sigma}_i = \frac{-B_i + \sqrt{B_i^2 + 4nC_i}}{2n} \tag{13}$$

and

$$\hat{\mu}_i = \bar{v}_{i[.]} - \hat{\delta}_i\, \bar{u}_{i[.]} - (\Delta_i/m_i)\hat{\sigma}_i\ , \tag{14}$$

where $m_i = \sum_{j=1}^{n}\beta_{ij}$, $\Delta_{ij} = \alpha_{ij} - 1/(b + 1)$, $\Delta_i = \sum_{j=1}^{n}\Delta_{ij}$ $(i = 1, 2)$,

$$v_{i[j]} = y_{i[j]} - \phi_i y_{i[j]-1}, \qquad \bar{v}_{i[.]} = \frac{1}{m_i}\sum_{j=1}^{n}\beta_{ij} v_{i[j]}\ ,$$

$$u_{i[j]} = x_{i[j]} - \phi_i\, x_{i[j]-1}, \qquad \bar{u}_{i[.]} = \frac{1}{m_i}\sum_{j=1}^{n}\beta_{ij} u_{i[j]}\ ;$$

$$G_i = \frac{\sum_{j=1}^{n} \beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)v_{i[j]}}{\sum_{j=1}^{n} \beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} \quad , \qquad H_i = \frac{\sum_{j=1}^{n} \Delta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)}{\sum_{j=1}^{n} \beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} \ ,$$

$$K_i = \frac{\left[\begin{array}{l} \sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]} - \hat{\delta}_i x_{i[j]}\right)\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right) \\ -\frac{1}{m_i}\sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)\sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]} - \hat{\delta}_i x_{i[j]}\right) \end{array}\right]}{\sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)^2 - \frac{1}{m_i}\left\{\sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)\right\}^2},$$

$$D_i = \frac{\sum_{j=1}^{n} \left\{\Delta_{ij} - (\Delta_i / m_i)\beta_{ij}\right\}\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)}{\sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)^2 - \frac{1}{m_i}\left\{\sum_{j=1}^{n} \beta_{ij}\left(y_{i[j]-1} - \hat{\delta}_i x_{i[j]-1}\right)\right\}^2},$$

$$B_i = (b+1)\left[\sum_{j=1}^{n} \Delta_{ij}\left\{\left(v_{i[j]} - \bar{v}_{i[.]}\right) - G_i\left(u_{i[j]} - \bar{u}_{i[.]}\right)\right\}\right] \ ,$$

$$C_i = (b+1)\left[\sum_{j=1}^{n} \beta_{ij}\left(v_{i[j]} - \bar{v}_{i[.]}\right)^2 - G_i\sum_{j=1}^{n} \beta_{ij} v_{i[j]}\left(u_{i[j]} - \bar{u}_{i[.]}\right)\right].$$

The estimators $\hat{\sigma}_i$ ($i = 1, 2$) are always real and positive since $\beta_{ij} > 0$.

Computations: Write $\gamma_i = -\delta_i \phi_i$ ($i = 1, 2$) and obtain the initial values from the equations

$$\begin{pmatrix} \hat{\delta}_{i0} \\ \hat{\phi}_{i0} \\ \hat{\gamma}_{i0} \end{pmatrix} = \begin{pmatrix} \sum y_{ij-1}x_{ij} & \sum x_{ij}^2 & \sum x_{ij-1}x_{ij} \\ \sum y_{ij}^2 & \sum y_{ij-1}x_{ij} & \sum y_{ij-1}x_{ij-1} \\ \sum y_{ij-1}x_{ij-1} & \sum x_{ij-1}x_{ij} & \sum x_{ij-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum x_{ij}y_{ij} \\ \sum y_{ij-1}y_{ij} \\ \sum x_{ij-1}y_{ij} \end{pmatrix}, \tag{15}$$

($i = 1, 2$); each sum is carried over $j = 1, 2, ..., n$. The estimators given by (15) are essentially the LS estimators obtained by minimizing $\sum_{j=1}^{n} (\mu_i + \varepsilon_{ij})^2$, $i = 1, 2$.

Using these initial values, calculate the MML estimates $\hat{\delta}_i$ ($i = 1, 2$) and $\hat{\sigma}_i$ from (11) and (13). Then calculate the MML estimates $\hat{\phi}_i$ ($i = 1, 2$) from (12). Carry out a second iteration by replacing $\hat{\delta}_{i0}$, $\hat{\phi}_{i0}$ and $\hat{\gamma}_{i0}$ ($i = 1, 2$) by $\hat{\delta}_i$, $\hat{\phi}_i$ and $-\hat{\delta}_i\, \hat{\phi}_i$ ($i = 1, 2$), respectively.

Final MML estimates $\hat{\delta}_i$, $\hat{\phi}_i$ (i = 1, 2) and $\hat{\sigma}_i$ are obtained at the end of the third iteration. Then compute the MML estimates $\hat{\mu}_i$ (i = 1, 2) from (14).

This is, in fact, the procedure suggested by Akkaya and Tiku (2001) and is similar to those adopted by earlier authors, e.g., Durbin (1960), Tan and Lin (1993), and Tiku et al. (1999).

4. Asymptotic Properties

The differences $[g_i\{z_{i(j)}\} - \{\alpha_{ij} - \beta_{ij} z_{i(j)}\}]$ converge to zero as n tends to infinity. Consequently, the differences $(1/n)\{(\partial \ln L/\partial\mu_i)-(\partial \ln L^*/\partial\mu_i)\}$ and $(1/n)\{(\partial \ln L/\partial\delta_i)-(\partial \ln L^*/\partial\delta_i)\}$ (i = 1, 2) are equal to zero asymptotically. For a rigorous proof see, for example, Vaughan and Tiku (2000). Therefore, the MML estimators are asymptotically equivalent to the ML estimators; Bhattacharya 1985; Vaughan and Tiku 2000.

1) Since $\partial \ln L^* / \partial\mu_i$ (i = 1, 2) can be written as

$$\frac{\partial \ln L}{\partial \mu_i} \cong \frac{\partial \ln L^*}{\partial \mu_i} = \frac{(b+1)m_i}{\sigma_i^2}\{\hat{\mu}_i(\delta_i, \phi_i, \sigma_i) - \mu_i\} \qquad (i = 1, 2)$$

where $\hat{\mu}_i(\delta_i, \phi_i, \sigma) = \bar{v}_{i[.]} - \delta_i \bar{u}_{i[.]} - \frac{\Delta_i}{m_i}\sigma_i$. Therefore, the MML estimators $\hat{\mu}_i(\delta_i, \phi_i, \sigma_i)$ (i = 1, 2) are conditionally (i.e. for known δ_i , ϕ_i and σ_i) the MVB estimators of μ_i (i=1, 2) with variances $\sigma_i^2 / m_i(b + 1)$ (i = 1, 2).

2) Similarly, $\partial \ln L^* / \partial\delta_i$ (i = 1, 2) can be written as

$$\frac{\partial \ln L}{\partial \delta_i} \cong \frac{\partial \ln L^*}{\partial \delta_i} = \frac{(b+1)m_i}{k\sigma_i^2}\left[\frac{1}{m_i}\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2\right]\left(\hat{\delta}_i(\phi_i, \sigma_i) - \delta_i\right)$$

(i = 1, 2) where $\hat{\delta}_i(\phi_i, \sigma_i) = G_i - H_i\sigma_i$;

$$G_i = \frac{\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)v_{i[j]}}{\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} \quad \text{and} \quad H_i = \frac{\sum_{j=1}^{n}\Delta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)}{\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2} \quad (i = 1, 2).$$

Thus, the MLL estimators $\hat{\delta}_i(\phi_i, \sigma_i)$ (i = 1, 2) are conditionally (i.e. for known ϕ_i and σ) the MVB estimators with variances $\dfrac{\sigma_i^2}{(b+1)\sum_{j=1}^{n}\beta_{ij}\left(u_{i[j]} - \bar{u}_{i[.]}\right)^2}$ (i = 1, 2) .

5. Efficiencies of the Estimators

The LS estimators are widely used regardless of the underlying distribution which we will show results in enormous loss of efficiency. Since it is very difficult to work out the expected values and the variance-covariances of the LS estimators even asymptotically, they are simulated.

The LS estimators are computed exactly the same way as MML estimators since the LS estimators can be obtained from the modified likelihood equations simply by equating α_{ij} and β_{ij} to 0 and 1, respectively. However, the LS estimators $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ so obtained need to be corrected for bias; $(\psi(b) - \psi(1))\hat{\sigma}_i$ must be subtracted to obtain the bias-corrected LS estimator of μ_i while the estimator of σ_i has to be divided by $(\psi'(b) + \psi'(1))^{1/2}$.

The x-values are generated from a uniform distribution U(0, 1). The values of b considered are b = 0.5, 1.0, and 2.0. $\mu_1 = \mu_2 = 0$, $\delta_1 = \delta_2 = 1$, $\phi_1 = \phi_2 = 0.5$, $\sigma_1 = 1$ and $\sigma_2 = 1.5$ are chosen, without loss of generality. For n = 50, the simulated means and variances of LS and MML estimators of μ_i, δ_i, ϕ_i, and σ_i and the relative efficiencies of the LS estimators defined as E = 100(variance of the MML estimator/variance of the LS estimator) are given in Table 1 for the case (ii) $x_{i,t}$ change with $y_{i,t}$.

Table 1. Values of (1) Mean, (2) Variance, and Relative Efficiency E of the Least Squares Estimators for the Generalized Logistic Distribution.

b		δ_1	δ_2	ϕ_1	ϕ_2	σ_1	σ_2	μ_1	μ_2
0.5	LS	1.0012	1.0013	0.4563	0.4571	0.9659	1.4323	-0.1750	-0.2670
(1)	MML	1.0015	1.0010	0.4675	0.4689	0.9808	1.4547	-0.1231	-0.1834
	LS	0.0865	0.0872	0.0160	0.0158	0.0201	0.0468	0.2734	0.6188
(2)	MML	0.0662	0.0681	0.0131	0.0138	0.0156	0.0350	0.2375	0.5365
	E	**76.5**	**78.1**	**81.9**	**87.3**	**77.6**	**74.8**	**86.9**	**86.7**
1.0	LS	0.9946	0.9985	0.4547	0.4560	0.9573	1.4399	-0.0064	-0.0060
(1)	MML	0.9939	0.9999	0.4576	0.4602	0.9716	1.4586	-0.0055	-0.0069
	LS	0.0429	0.0397	0.0167	0.0167	0.0154	0.0337	0.0816	0.1926
(2)	MML	0.0403	0.0370	0.0162	0.0159	0.0138	0.0309	0.0755	0.1751
	E	**93.9**	**93.2**	**97.0**	**95.2**	**89.6**	**91.7**	**92.5**	**90.9**
2.0	LS	1.0016	1.0015	0.4534	0.4619	0.9563	1.4386	0.1345	0.1680
(1)	MML	1.0021	1.0020	0.4614	0.4656	0.9701	1.4564	0.1079	0.1409
	LS	0.0280	0.0294	0.0168	0.0162	0.0149	0.0363	0.1237	0.2667
(2)	MML	0.0248	0.0264	0.0150	0.0145	0.0129	0.0316	0.1117	0.2454
	E	**88.6**	**89.8**	**89.3**	**89.5**	**86.6**	**87.1**	**90.3**	**92.0**

From the table, it can be seen that the LS estimators are considerably less efficient than the MML estimators. Similar results were obtained for other ϕ_i , μ_i and n values, but are not reported for conciseness. In fact, the relative efficiencies of the LS estimators decrease as the sample size n increases.

Similar results were obtained for the first case where $x_{i,t}$ are design variables in Türker (2002) so we do not reproduce them here.

6. Hypothesis Testing

The main interest here is to test $H_0 : \delta_1 = \delta_2$. We consider two cases: 1) $\sigma_1 = \sigma_2$, and 2) $\sigma_1 \neq \sigma_2$.

Case 1) The test statistic

$$t = \frac{\hat{\delta}_1 - \hat{\delta}_2}{\hat{\sigma}\sqrt{\dfrac{1}{(b+1)\sum_{j=1}^{n}\beta_{1j}\left(\hat{u}_{1[j]} - \bar{\hat{u}}_{1[.]}\right)^2} + \dfrac{1}{(b+1)\sum_{j=1}^{n}\beta_{2j}\left(\hat{u}_{2[j]} - \bar{\hat{u}}_{2[.]}\right)^2}}} \tag{16}$$

is used to test the null hypothesis H_0: $\delta_1 = \delta_2$.

Here $\hat{\sigma}$ is the pooled estimator of σ which is given by $\hat{\sigma} = \sqrt{\left(\hat{\sigma}_1^2 + \hat{\sigma}_2^2\right)/2}$.

For given values of ϕ_i (< 1 in absolute value), the conditional distribution (i.e. σ known) of $\hat{\delta}_i$ is asymptotically normal as n tends to infinity. Since $\hat{\phi}_i$ and $\hat{\sigma}$ converge to ϕ_i and σ, respectively, as n becomes large, the asymptotic distribution of t is normal N(0, 1). The asymptotic power function is Prob$\{Z \geq z_\alpha - |\lambda|\}$ where Z is a standard normal variate, z_α is its $100(1 - \alpha)\%$ point and λ is the noncentrality parameter,

$$\lambda^2 = \frac{(\delta_1 - \delta_2)^2}{\sigma^2\left(\dfrac{1}{(b+1)\sum_{j=1}^{n}\beta_{1j}\left(u_{1[j]} - \bar{u}_{1[.]}\right)^2} + \dfrac{1}{(b+1)\sum_{j=1}^{n}\beta_{2j}\left(u_{2[j]} - \bar{u}_{2[.]}\right)^2}\right)}. \tag{17}$$

Case 2) The test statistic

$$t = \frac{\hat{\delta}_1 - \hat{\delta}_2}{\sqrt{\dfrac{\hat{\sigma}_1^2}{(b+1)\sum_{j=1}^{n}\beta_{1j}\left(\hat{u}_{1[j]} - \bar{\hat{u}}_{1[.]}\right)^2} + \dfrac{\hat{\sigma}_2^2}{(b+1)\sum_{j=1}^{n}\beta_{2j}\left(\hat{u}_{2[j]} - \bar{\hat{u}}_{2[.]}\right)^2}}} \tag{18}$$

is used to test the null hypothesis H_0: $\delta_1 = \delta_2$.

For given value of ϕ_i (< 1 in absolute value), the conditional distribution (i.e. σ_i known) of $\hat{\delta}_i$ is asymptotically normal as n tends to infinity. Since $\hat{\phi}_i$ and $\hat{\sigma}_i$ converge to ϕ_i and σ_i , respectively, as n becomes large, the asymptotic distribution of t is normal N(0, 1). The asymptotic power function is Prob$\{Z \geq z_\alpha - |\lambda|\}$ where Z is a standard normal variate, z_α is its 100(1 - α)% point and λ is the noncentrality parameter,

$$\lambda^2 = \frac{(\delta_1 - \delta_2)^2}{\left(\dfrac{\sigma_1^2}{(b+1)\sum_{j=1}^{n}\beta_{1j}\left(u_{1[j]} - \bar{u}_{1[.]}\right)^2} + \dfrac{\sigma_2^2}{(b+1)\sum_{j=1}^{n}\beta_{2j}\left(u_{2[j]} - \bar{u}_{2[.]}\right)^2}\right)}. \tag{19}$$

To obtain the test statistics based on the LS estimators, simply replace the MML estimators by LS estimators in (16) and (18). Also equate β_{ij} values to 1 and the multiplier (b + 1) to 1 in (16)-(19). For a comparison of the power of the tests based on the MML estimators and LS estimators, we have done simulations. Power of the test statistics based on the LS and the MML estimators are calculated by using both the simulated variances and the asymptotic variances of the estimator of (δ_1 - δ_2). Also the theoretical power values are calculated. We considered $\mu_1 = \mu_2 = 0$, $\phi_1 = \phi_2 = 0.5$ and $\delta_1 = 1$ without loss of generality. For n = 50 type I errors with the use of the simulated variances are given in Table 2; the graphs of the power functions are given in Figure 1.

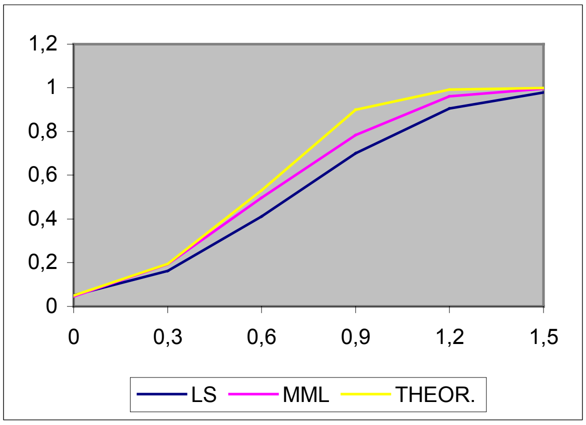
Table 2. Type I Error Values

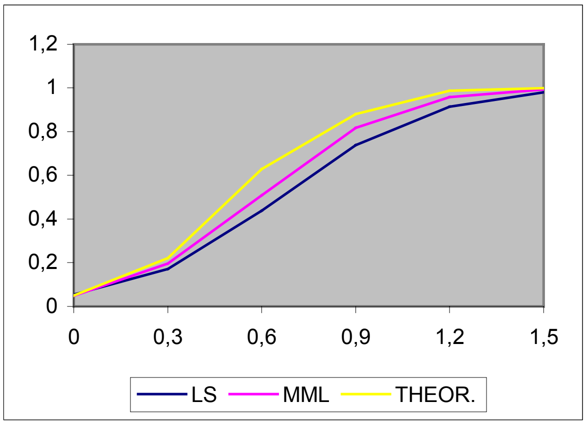
	$\sigma_1 = 1.0$ and $\sigma_2 = 1.5$		**$\sigma_1 = \sigma_2 = 1.0$**	
b	**LS**	**MML**	**LS**	**MML**
0.5	0.054	0.049	0.051	0.045
1.0	0.052	0.048	0.052	0.049
2.0	0.050	0.051	0.050	0.047

From the table, it can be seen that in both cases, the test based on the MML estimators has type I error less than the test based on the LS estimators. Also, in all situations, the tests based on the MML estimators are more powerful which can be seen from the figures.

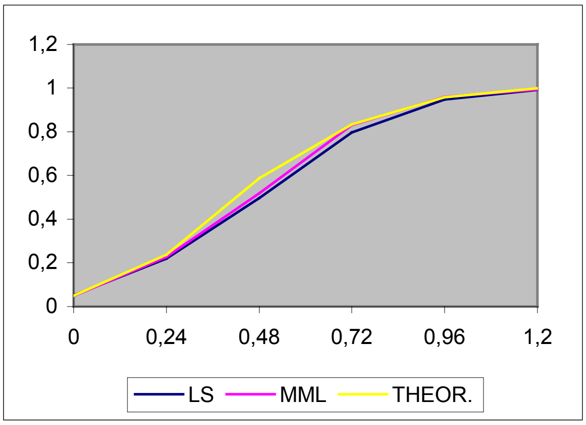
$\sigma_1 = \sigma_2 = 1.0$ $\qquad\qquad$ $\sigma_1 = 1.0$ and $\sigma_2 = 1.5$

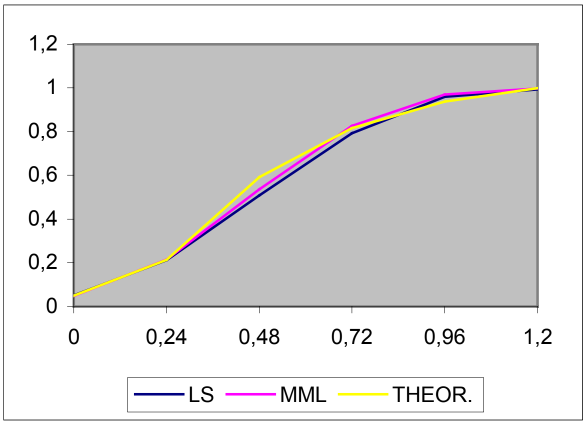
b = 0.5



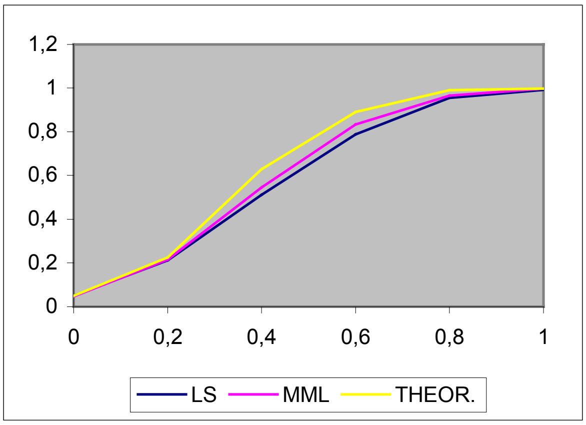


b = 1.0





b = 2.0



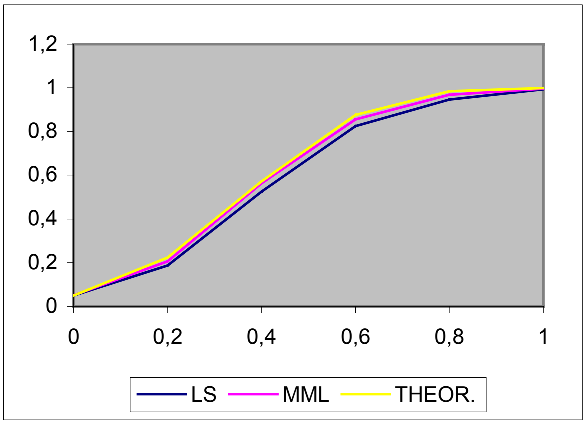


Fig. 1. Graph of Power Functions of Test Statistics Based on LS, MML Estimators and Theoretical Power Function for H_0: $\delta_1 = \delta_2$.

For large n ($\geq$ 100), asymptotic variances of $\hat{\delta}_i$ (i = 1, 2) can be used with similar results. Türker (2002) obtained similar results for the first case where $x_{i,t}$ are design variables so we do not reproduce them here.

7. Q-Q Plots

To locate a plausible distribution, we construct a Q-Q plot by plotting the order statistics $y_{(i)}$ ($1 \leq i \leq n$) of a random sample of size n (called sample quantiles) against the population quantiles $t_{(i)}$ defined as follows.

Let the underlying distribution be the location-scale type, i.e., $(1/\sigma)\, f((y-\mu)/\sigma)$. Writing $z = (y - \mu) / \sigma$, the distribution of z is f(z). The functional form f is not known. For an assumed f, we obtain $t_{(i)}$ from the equation

$$\int_{-\infty}^{t_{(i)}} f(z)dz = \frac{i}{n_i + 1}, \qquad (1 \leq i \leq n).$$

We plot $y_{(i)}$ against $t_{(i)}$ ($1 \leq i \leq n$). If we get a straight line (or closest to such), then f is a plausible distribution (model) for the data. In fact, we calibrate with several density functions and choose the one which gives us closest to a straight line.

To locate a plausible error distribution for our model

$$Y_t = \phi Y_{t-1} + \mu + \delta\,(X_t - \phi X_{t-1}) + \varepsilon_t\,,$$

we construct a Q-Q plot for the observations

$$w_t = y_t - \tilde{\phi}\, y_{t-1} - \tilde{\delta}\,(X_t - \tilde{\phi}\, X_{t-1})$$

where $\tilde{\delta}$ and $\tilde{\phi}$ are the LSE. We calibrate with a few conceivable distributions and choose the one which gives closest to a straight line pattern. We follow it up with a statistical analysis based on the MMLE.

8. Conclusion

The use of LS estimators in autoregressive models when the innovations have a non-normal distribution, which is more prevalent in practice, results in loss of efficiency as compared to ML estimators. However, the ML estimators have numerous computational difficulties. Therefore, it is recommended to use the MML estimators which have explicit solutions. In this study, the MML esitmators for two regressive models where the innovations come from generalized logistic family are derived and showed to be efficient. It is shown that LS estimators have less efficiencies compared to the MML estimators. Also, the test statistic for testing the equivalance of the parameters ($H_0 : \delta_1 = \delta_2$) are derived and its properties are studied and found that it has less type I error and more powerfull than the corresponding test statistic based on the LS estimators.

Acknowledgement

I am grateful to Prof. Dr. Moti Lal Tiku and Assoc. Prof. Dr. Ayşen Akkaya for their guidance and advice. It was a great honor for me to study especially with Prof. Dr. Moti Lal Tiku.

References

- Akkaya, A. and Tiku, M. L. (2001), "Estimating parameters in autoregressive models in non-normal situations: asymmetric innovations", Commun. Statist. Theory Meth., 30(3), 517-536.
- Anderson, R. L. (1949), "The problem of autocorrelation in regression analysis", J. Amer. Statist. Assoc., 44, 113-127.
- Barnett, V. D. (1966), "Evaluation of the maximum likelihood estimator when the likelihood equation has multiple roots", Biometrika, 53, 151-165.
- Beach, C. M., and Mackinnon, J. G. (1978), "A maximum likelihood procedure for regression with autocorrelated errors", Econometrics, 46, 51-58.
- Bhattacharya, G. K. (1985), "The asymptotics of maximum likelihood and related estimators based on Type II censored data", J. Amer. Statist. Assoc., 80, 398-404.
- Cochrane, D. and Orcutt, G. H. (1949), "Application of least squares regression to relationships containing autocorrelated error terms", J. Amer. Statist. Assoc., 44, 32-61.
- Cogger, K. O. (1990), "Robust time series analysis-an L_1 approach", Robust Regression, (K. D. Lawrence and J. L. Arthur, Eds.), New York: Marcel Dekker.
- Dielman, T. E., and Pfaffenberger, R. C. (1989), "Efficiency of ordinary least square for linear model with autocorrelation", J. Amer. Statist. Assoc., 84, 242-248.
- Durbin, J. (1960), "Estimation of parameters in time series regression model", JRSS B, 22, 139-153.
- Elveback, L. R., Guillier, C. L., Kerting, F. R. (1970), "Health, normality and the ghost of Gauss", J. American Medical Assoc., 211, 69-75.
- Gallant, A. R., and Goebel, J. J. (1967), "Nonlinear regression with autocorrelated error", J. Amer. Statist. Assoc., 71, 961-967.

- Kramer, W. (1980), “Finite sample efficiency of ordinary least square in the linear regression with autocorrelated errors”, J. Amer. Statist. Assoc., 75, 1005-1009.
- Lee, K. R., Kapadia, C. H., and Dwight, B. B. (1980), “On estimating the scale parameter of Rayleigh distribution from censored samples”, Statist. Hefte, 21, 14-20.
- Magee, L., Ullah, A. and Srivastava, V. K. (1987), “Efficiency of estimators in the regression model with first-order autoregressive errors”, Specification analysis in the linear model, Internat. Lib. Econom., London: Routledge and Kegan Paul, 81-98.
- Maller, R. A. (1989), “Regression with autoregressive errors-some asymptotic results”, Statistics, 20, 23-39.
- Nagaraj, N. K., and Fuller, W. A. (1992), “Least squares estimation of the linear model with autoregressive errors”, New Directions in time series analysis, Part I, IMA Vol. Math. Appl. 45, New York: Springer, 215-225.
- Pearson, E. S. (1932), “The analysis of variance in cases of nonnormal variation”, Biometrika, 23, 114-133.
- Pearson, E. S. and Hartley, H. O. (1972), Biometrika Tables for Statisticians, Vol. II, Cambridge: University Press.
- Puthenpura, S., and Sinha, N. K. (1986), “Modified maximum likelihood method for the robust estimation of system parameters from very noisy data”, Automatica, 22, 231-235.
- Schäffler, S. (1991), “Maximum likelihood estimation for linear regression model with autoregressive errors”, Statistics, 22, 191-198.
- Tiao, C. H., and Tan, W. Y. (1966), “Bayesian Analysis of random-effect models in the analysis of variance, II: effect of autocorrelated errors”, Biometrika, 53, 477-495.
- Tiku, M. L., Tan, W. Y., and Balakrishnan, N. (1986), Robust Inference, New York: Marcel Dekker.
- Tiku, M. L.,and Suresh, R. P.(1992),“A new method of estimation for location and scale parameters”, J. Stat. Plann. And Inf., 30, 281-292.
- Tiku, M. L.,Wong, W. K.,and Bian, G.(1999),“Estimating parameters in autoregressive models in non-normal situations; symmetric innovations”, Commun. Statist. Theory Meth., 28(2), 315-341.
- Tse, Y. K. (1991), “Price and volume in the Tokyo stock exchange: an explanatory study”, Japanese Financial Market Research (W. T. Ziemba, W. Bailey and Y. Hamano, Eds.), New York: Elsevier Science Publishers B. V., 91-119.

- Tukey, J. W. (1970), Explanatory Data Analysis, mimeographed preliminary edition, Princeton: University Press.
- Türker, Ö. (2002), Autoregressive Models: Statistical Inference and Applications, Ph. D. Thesis, Middle East Technical University, Ankara.
- Vaughan, D. C. (1992), “On the Tiku-Suresh method of estimation”, Commun. Statist. Theory Meth., 21, 451-469.
- Vaughan, D. C. (2002), “The generalized secant hyperbolic distribution and its properties”, Commun. Statist. Theory Meth., 31(2), 219-238.
- Vaughan, D. C., and Tiku, M. L. (2000), “Estimation and hypothesis testing for a non-normal bivariate distribution and applications”, J. Mathematical and Computer Modelling, 32, 53-67.
- Velu, R., and Gregory, C. (1987), “Reduced rank regression with autoregressive errors”, Econometrics, 35, 317-335.
- Weiss, G. (1990), “Least absolute error estimation in the presence of serial correlation”, J. of Econometrics, 44, 127-158.