2007

Estimating parameters in autoregressive models in non-normal situations: symmetric innovations

Moti L Tiku

Wing Keung Wong
Hong Kong Baptist University, awong@hkbu.edu.hk

Guorui Bian

Citation

Tiku, Moti L, Wing Keung Wong, and Guorui Bian. "Estimating parameters in autoregressive models in non-normal situations: symmetric innovations." *Communications in Statistics - Theory and Methods* 28.2 (2007): 315-341.

ESTIMATING PARAMETERS IN AUTOREGRESSIVE MODELS IN NON-NORMAL SITUATIONS: SYMMETRIC INNOVATIONS

Moti L Tiku
McMaster University and Wilfrid Laurier University
Canada

Wing-Keung Wong **Guorui Bian**
National University of Singapore
Singapore

Keywords. Autoregression; Nonnormality; Maximum likelihood; Modified maximum likelihood; Least Squares; Robustness.

Abstract. The estimation of coefficients in a simple regression model with autocorrelated errors is considered. The underlying distribution is assumed to be symmetric, one of Student's t family for illustration. Closed form estimators are obtained and shown to be remarkably efficient and robust. Skew distributions will be considered in a future paper.

1. INTRODUCTION

The estimation of coefficients in a simple regression model with autocorrlelated errors is an important problem and has received a great deal of attention in the literature. Most of the work reported is, however, based 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 Machinnon (1978), Kramer (1980), Magee et al (1987), Velu and Gregory (1987), Dielman and Pfuffenberger (1989), Maller (1989), Cogger (1990), Weiss (1990), Schäffler (1991), Nagaraja et al (1992), Tan and Lin (1993). The paper by Tan and Lin (1993) is of particular interest. They assumed normality but based their estimators on censored samples. They showed that the resulting estimators are robust to plausible deviations from normality. In recent years, however, it has been recognized that the underlying distribution is, in most situations, basically not normal; see, for example, Huber (1981) and Tiku et al (1986). The problem, therefore, is to develop efficient estimators of coefficients in autoregressive models when the underlying distribution is non-normal. Naturally, one would prefer closed form estimators which are fully efficient (or nearly so). Preferably, these estimators should also be robust to plausible deviations from an assumed model. That is exactly what has been achieved in this paper. The underlying distribution is assumed to be symmetric and of the type $(1/\sigma)f((y-\mu)/\sigma)$, one of Student's t family for illustration. The method of modified maximum likelihood estimation (Tiku 1967, 1968, 1980; Tiku and Suresh 1992) is invoked. The resulting estimators are explicit functions of sample observations and are asymptotically fully efficient. They are almost fully efficient for small sample sizes; fully efficient estimators do not exist for small sample sizes. The estimators are also shown to be remarkably robust. The relative efficiencies of the LS (least squares) estimators are investigated and shown to be generally quite low. A test for the regression coefficient $H_0 : \delta = 0$ is formulated. The methodology developed in this paper can (hopefully) be extended to skew distributions. That will be the subject matter of a future paper.

2. AUTOREGRESSIVE MODEL

Consider a simple autoregressive model

$$\begin{aligned} y_t &= \mu' + \delta x_t + e_t \\ e_t &= \phi\, e_{t-1} + a_t \qquad (t = 1, 2, 3, \cdots, n) \end{aligned} \tag{1}$$

where

$$y_t = \text{observed value of a random variable } y \text{ at time } t$$
$$x_t = \text{value of a nonstochastic design variable } x \text{ at time } t$$
$$\phi = \text{Autoregressive coefficient } (|\phi| \leq 1).$$

It is assumed that the innovations a_t are iid. The autoregressive model (1) has many applications. For example, in predicting future stock prices the effect of an intervention might persist for some time. Numerous other applications of the model above are in agricultural, biological and biomedical problems besides business and economics; see, for example, Anderson (1949), Durbin (1960), Tiao and Tan (1966), Beach and Machinnon (1978), Cogger (1990), Weiss (1990), Schäffler (1991).

Assume that the common distribution of a_t's is symmetric and is, for illustration, given by

$$f(a;p) \propto \frac{1}{\sigma}\left\{1 + \frac{a^2}{k\sigma^2}\right\}^{-p} \qquad (-\infty < a < \infty) \tag{2}$$

where $k = 2p - 1$ and $p \geq 2$. Note that $E(a) = 0$ and $V(a) = \sigma^2$. It may be noted that $t = \sqrt{(k/\nu)}\,(a/\sigma)$ has Student's t distribution with $\nu = 2p - 1$ df (degree of freedom). For $1 \leq p < 2$, k is equated to 1 in which case σ in (2) is simply a scale parameter. In the first place we assume that p is known.

3. MODIFIED LIKELIHOOD

An alternative form of the model (1) is

$$y_t - \phi\, y_{t-1} = \mu + \delta(x_t - \phi\, x_{t-1}) + a_t \qquad (1 \leq t \leq n)\,. \tag{3}$$

Conditional on y_0, the likelihood function is

$$L \propto \left(\frac{1}{\sigma}\right)^n \Pi_{i=1}^n \left[1 + \frac{z_i^2}{k}\right]^{-p} \tag{4}$$

where $z_t = (1/\sigma)\{(y_t - \phi y_{t-1}) - \mu - \delta(x_t - \phi x_{t-1})\}$ $(1 \leq t \leq n)$; see Hamilton (1994, p123) for numerous advantages of conditional likelihoods. Let

$$z_{(i)} = \frac{(y_{[i]} - \phi y_{[i-1]}) - \mu - \delta(x_{[i]} - \phi x_{[i-1]})}{\sigma} \tag{5}$$

denote the order statistics of z_i $(1 \leq i \leq n)$ arranged in ascending order of magnitude. Note that $y_{[i-1]} = y_{[i]-1}$, $x_{[i-1]} = x_{[i]-1}$ and $(y_{[i]}, x_{[i]})$ is the pair (y_i, x_i) that determines $z_{(i)}$; $(y_{[i]}, x_{[i]})$ may be called concomitants of $z_{(i)}$ $(1 \leq i \leq n)$. Since complete sums are invariant to ordering, the likelihood equations for estimating μ, δ, σ, ϕ are

$$\frac{\partial \ln L}{\partial \mu} = \frac{2p}{k\sigma} \sum_{i=1}^{n} g(z_{(i)}) = 0 \tag{6}$$

$$\frac{\partial \ln L}{\partial \delta} = \frac{2p}{k\sigma} \sum_{i=1}^{n} (x_{[i]} - \phi x_{[i]-1}) g(z_{(i)}) = 0 \tag{7}$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma} \sum_{i=1}^{n} z_{(i)} g(z_{(i)}) = 0 \tag{8}$$

$$\frac{\partial \ln L}{\partial \phi} = \frac{2p}{k\sigma} \sum_{i=1}^{n} (y_{[i]-1} - \delta x_{[i]-1}) g\{z_{(i)}\} = 0 \tag{9}$$

where $g(z) = z/\{1 + (1/k)z^2\}$.

Equations (6)–(9) have to be solved by iterative methods which is a formidable task. Moreover, Barnett (1966) and Lee et al (1980) point out some fundamental difficulties with iterative solutions, e.g., the iterations might converge to wrong values or not converge at all. See also Pearson and Hartley (1972, p89) who give examples where the iterations involved in determining ML (maximum likelihood) estimates do not converge rapidly enough. Besides, the solutions provided by different iterative methods are not neccessarily identical (Barnett 1966).

To obtain efficient closed form estimators, we invoke Tiku's method of modified likelihood estimation which is by now well established (Smith et al 1973, Lee et al 1980, Tan 1985, Tiku, et al 1986, Schneider 1986, Vaughan 1992a). Let $t_{(i)} = E\{z_{(i)}\}$ $(1 \leq i \leq n)$ be the expected values of the standardized order statistics. Since $g(z)$ is almost linear in a small interval $c \leq z \leq d$ (Tiku 1967, 1968; Tiku and Suresh 1992) and realizing that under some very general regularity conditions $z_{(i)}$ converges to $t_{(i)}$ as n becomes large, we use the linear approximations given by a Taylor series expansion:

$$\begin{aligned} g\{z_{(i)}\} &\simeq g\{t_{(i)}\} + (z_{(i)} - t_{(i)}) \left\{ \frac{\partial g(z)}{\partial z} \right\}_{z=t_{(i)}} \\ &= \alpha_i + \beta_i \, z_{(i)} \qquad (\, 1 \leq i \leq n \,) \,. \end{aligned} \tag{10}$$

This gives

$$\alpha_i = \frac{(2/k) t_{(i)}^3}{[1 + (1/k)\, t_{(i)}^2]^2} \qquad \text{and} \qquad \beta_i = \frac{1 - (1/k) t_{(i)}^2}{[1 + (1/k)\, t_{(i)}^2]^2} \qquad (i = 1, 2, \cdots, n) \,. \tag{11}$$

The values of $t_{(i)}$ are readily available for $n \leq 20$, in Tiku and Kumra (1981) for $p = 2(.5)10$ and in Vaughan (1992b, 1994) for $p = 1$ and 1.5. For $n > 20$, $t_{(i)}$ are obtained from the equation:

$$\int_{-\infty}^{t_{(i)}} f(z)dz = \frac{i}{n+1} \qquad (1 \leq i \leq n)\,. \tag{12}$$

In evaluating (12), it is helpful to remember that $\sqrt{(k/\nu)}z$ has Student's t distribution with $\nu = 2p - 1$ df. Note that

$$\alpha_i = -\alpha_{n-i+1} \quad , \qquad \beta_i = \beta_{n-i+1} \quad (1 \leq i \leq n) \qquad \text{and} \qquad \sum_{i=1}^{n} \alpha_i = 0\,;$$

this follows from the fact that for symmetric distributions $t_{(i)} = -t_{(n-i+1)}$. Modified likelihood equations are obtained by incorporating (10) in (6) and (9):

$$\frac{\partial\, ln\, L}{\partial\, \mu} \simeq \frac{\partial\, ln\, L^*}{\partial\, \mu} = \frac{2p}{k\sigma}\sum_{i=1}^{n}\{\alpha_i + \beta_i\, z_{(i)}\} = 0 \tag{13}$$

$$\frac{\partial\, ln\, L}{\partial\, \delta} \simeq \frac{\partial\, ln\, L^*}{\partial\, \delta} = \frac{2p}{k\sigma}\sum_{i=1}^{n}(x_{[i]} - \phi x_{[i]-1})\{\alpha_i + \beta_i\, z_{(i)}\} = 0 \tag{14}$$

$$\frac{\partial\, ln\, L}{\partial\, \sigma} \simeq \frac{\partial\, ln\, L^*}{\partial\, \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma}\sum_{i=1}^{n} z_{(i)}\{\alpha_i + \beta_i\, z_{(i)}\} = 0 \tag{15}$$

$$\frac{\partial\, ln\, L}{\partial\, \phi} \simeq \frac{\partial\, ln\, L^*}{\partial\, \phi} = \frac{2p}{k\sigma}\sum_{i=1}^{n}(y_{[i]-1} - \delta x_{[i]-1})\{\alpha_i + \beta_i\, z_{(i)}\} = 0\,. \tag{16}$$

Note that the differences $\{g(z_{(i)}) - (\alpha_i + \beta_i\, z_{(i)})\}$ converge to zero as n becomes large. Consequently, $(1/n)\{(\partial\, ln\, L/\partial\, \mu) - (\partial\, ln\, L^*/\partial\, \mu)\}$, $(1/n)\{(\partial\, ln\, L/\partial\, \delta) - (\partial\, ln\, L^*/\partial\, \delta)\}$, etc., converge to zero. This is primarily the reason that the modified likelihood equations are asymptotically equivalent to the corresponding likelihood equations (Tiku 1970, Bhattacbaryya 1985, Tiku et al 1986, Tiku and Suresh 1992).

4. THE MML ESTIMATORS

The MML estimators of μ, δ, and σ (for a given ϕ) are the solutions of (13) — (16):

$$\hat{\mu} = \overline{w}_{[.]} - \hat{\delta}\, \overline{u}_{[.]} \tag{17}$$

$$\hat{\delta} = G + H\hat{\sigma} \tag{18}$$

$$\hat{\sigma} = \frac{B + \sqrt{B^2 + 4n\, C}}{2n} \tag{19}$$

where

$$\begin{aligned}
w_{[i]} &= y_{[i]} - \phi y_{[i]-1} \qquad u_{[i]} = x_{[i]} - \phi x_{[i]-1} \qquad (1 \le i \le n) \\
\overline{w}_{[.]} &= \frac{1}{m}\sum_{i=1}^{n} \beta_i \, w_{[i]} \qquad \overline{u}_{[.]} = \frac{1}{m}\sum_{i=1}^{n} \beta_i \, u_{[i]} \\
G &= \frac{\sum_{i=1}^{n} \beta_i \, w_{[i]} u_{[i]} - m\overline{w}_{[.]}\overline{u}_{[.]}}{\sum_{i=1}^{n} \beta_i u_{[i]}^2 - m\overline{u}_{[.]}^2} \\
H &= \frac{\sum_{i=1}^{n} \alpha_i \, u_{[i]}}{\sum_{i=1}^{n} \beta_i u_{[i]}^2 - m\overline{u}_{[.]}^2} \\
B &= \frac{2p}{k}\sum_{i=1}^{n} \alpha_i \, \{w_{[i]} - Gu_{[i]}\} \\
C &= \frac{2p}{k}\sum_{i=1}^{n} \beta_i \, \{(w_{[i]} - \overline{w}_{[.]}) - G(u_{[i]} - \overline{u}_{[.]})\}^2 \\
&= \frac{2p}{k}\{\sum_{i=1}^{n} \beta_i \, (w_{[i]} - \overline{w}_{[.]})^2 - G\sum_{i=1}^{n} \beta_i \, (w_{[i]} - \overline{w}_{[.]})(u_{[i]} - \overline{u}_{[.]})\} \, .
\end{aligned}$$

Note that $\sum_{i=1}^{n} \beta_i u_{[i]}^2 - m\overline{u}_{[.]}^2 = \sum_{i=1}^{n} \beta_i \, (u_{[i]} - \overline{u}_{[.]})^2$ and $m = \sum_{i=1}^{n} \beta_i$.
Incorporating (17) – (19) in (16), the MML estimator of ϕ is obtained:

$$\hat{\phi} = K + L\hat{\sigma} \tag{20}$$

where

$$\begin{aligned}
K &= \frac{\sum_{i=1}^{n} \beta_i \, (y_{[i]} - \hat{\delta} x_{[i]-1}) \, (y_{[i]-1} - \hat{\delta} x_{[i]-1}) - \hat{\mu} \sum_{i=1}^{n} \beta_i \, (y_{[i]-1} - \hat{\delta} x_{[i]-1})}{\sum_{i=1}^{n} \beta_i (y_{[i]-1} - \hat{\delta} x_{[i]-1})^2} \\
L &= \frac{\sum_{i=1}^{n} \alpha_i \, (y_{[i]-1} - \hat{\delta} x_{[i]-1})}{\sum_{i=1}^{n} \beta_i \, (y_{[i]-1} - \hat{\delta} x_{[i]-1})^2} \, .
\end{aligned}$$

It is clear that the MML estimators above have all closed form algebraic expressions and are, therefore, easy to compute. Moreover, they are asymptotically equivalent to the maximum likelihood estimators.

<u>Computations:</u> Since $a_i = \sigma z_i$ and σ is a positive constant, the order statistics $z_{(i)}$ are determined by the order statistics $a_{(i)}$ $(1 \le i \le n)$. To initialize ordering of

$$a_i = y_i - \mu - \delta x_i - \phi y_{i-1} - \gamma x_{i-1} \qquad (\, \gamma = -\delta\phi \,) \, . \tag{21}$$

We ignore the constraint $\gamma = -\delta\phi$ (Durbin 1960, Tan and Lin 1993) and calculate the least square estimators μ^*, δ^*, ϕ^* and γ^*:

$$\begin{pmatrix} \mu^* \\ \delta^* \\ \phi^* \\ \gamma^* \end{pmatrix} = \begin{pmatrix} n & \sum x_i & \sum y_{i-1} & \sum x_{i-1} \\ \sum x_i & \sum {x_i}^2 & \sum y_{i-1}x_i & \sum x_i x_{i-1} \\ \sum y_{i-1} & \sum y_{i-1}x_i & \sum {y_{i-1}}^2 & \sum y_{i-1}x_{i-1} \\ \sum x_{i-1} & \sum x_i x_{i-1} & \sum y_{i-1}x_{i-1} & \sum {x_{i-1}}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i y_{i-1} \\ \sum y_i x_{i-1} \end{pmatrix}$$

each sum is carried over $i = 1, 2, \cdots, n$. Initially, therefore,

$$a_{(i)} = y_{[i]} - \mu^* - \delta^* x_{[i]} - \phi^* y_{[i]-1} - \gamma^* x_{[i]-1} \qquad (1 \le i \le n)\,. \tag{22}$$

Using the initial concomitants $(y_{[i]}\,,\, x_{[i]})$ $(1 \le i \le n)$ determined by (22), the MML estimators $\hat{\mu}$, $\hat{\delta}$ and $\hat{\sigma}$ are calculated from equations (17) – (19) with $\phi = \phi^*$. The MML estimator $\hat{\phi}$ is then calculated from (20). In the second iteration, the order statistics $z_{(i)}$ are determined by ordering

$$a_i = (y_i - \hat{\phi} y_{i-1}) - \hat{\mu} - \hat{\delta}(x_i - \hat{\phi} x_{i-1}) \qquad (1 \le i \le n)$$

and revised concomitant pairs $(y_{[i]}\,,\, x_{[i]})$ obtained. The estimators $\hat{\mu}$, $\hat{\delta}$ and $\hat{\sigma}$ are calculated from (17) – (19) and $\hat{\phi}$ is then calculated from (20). For more iterations are carried out till the estimates stablize. In all our computations partly presented in this paper, no more than three iterations were needed for the estimates to stablize. The reason for that is that the MML estimators only depend on $(y_{[i]}\,,\, x_{[i]})$ and the concomitant indices $[i]$ are determined by the relative magnitudes, not neccessarily the true values, of z_i $(1 \le i \le n)$. In the rest of the paper we assume that $|\phi| < 1$. The case $|\phi| = 1$ will be studied in another paper.

5. ASYMPTOTIC RESULTS

Since $\partial\, ln\, L^*/\partial\, \mu$, $\partial\, ln\, L^*/\partial\, \delta$ and $\partial\, ln\, L^*/\partial\, \phi$ are, as said earlier, asymptotically equivalent to $\partial\, ln\, L/\partial\, \mu$, $\partial\, ln\, L/\partial\, \delta$ and $\partial\, ln\, L/\partial\, \phi$, respectively, we have the following asymptotic results.

LEMMA 1: The MML estimator, $\hat{\mu}(\delta, \phi)$ is conditionally (for known δ and ϕ) the MVB (minimum variance bound) estimator with variance $(p+1)(p-\frac{3}{2})\sigma^2/np(p-\frac{1}{2})$.

Proof: This follows from the fact that $\sum_{i=1}^{n} \alpha_i = 0$ and

$$\frac{\partial\, ln\, L^*}{\partial\, \mu} = \frac{2mp}{k\sigma^2}\{\hat{\mu}(\delta, \phi) - \mu\}$$

where

$$\hat{\mu}(\delta,\phi) = \frac{1}{m}\sum_{i=1}^{n}\beta_i\{(y_{[i]} - \phi\, y_{[i]-1}) - \delta(x_{[i]} - \phi\, x_{[i]-1})\}$$
$$= \overline{w}_{[.]}(\phi) - \delta\overline{u}_{[.]}(\phi)$$

and (see the Appendix)

$$lim_{n\to\infty}\frac{m}{n} = \frac{p-\frac{1}{2}}{p+1}\,.$$

LEMMA 2: The MML estimator $\delta(\phi,\sigma)$ is conditionally (for known ϕ and σ) the MVB estimator with variance

$$\frac{(p-\frac{3}{2})\frac{\sigma^2}{np}}{\frac{1}{n}\sum_{i=1}^{n}\beta_i\{u_{[i]}(\phi) - \overline{u}_{[.]}(\phi)\}^2}\,.$$

Proof: This follows from the fact that

$$\frac{\partial\, ln\, L^*}{\partial\,\delta} = \frac{2np}{k\sigma^2}\left(\frac{1}{n}\sum_{i=1}^{n}\beta_i\{u_{[i]}(\phi) - \overline{u}_{[.]}(\phi)\}^2\right)\{\hat{\delta}(\phi,\sigma) - \delta\}$$

where $\hat{\delta}(\phi,\sigma) = G_0 + \sigma H_0$ and

$$G_0 = \frac{\sum_{i=1}^{n}\beta_i u_{[i]}(\phi)\{w_{[i]}(\phi) - \overline{w}_{[.]}(\phi)\}}{\sum_{i=1}^{n}\beta_i\{u_{[i]}(\phi) - \overline{u}_{[.]}(\phi)\}^2}$$
$$H_0 = \frac{\sum_{i=1}^{n}\alpha_i u_{[i]}(\phi)}{\sum_{i=1}^{n}\beta_i\{u_{[i]}(\phi) - \overline{u}_{[.]}(\phi)\}^2}\,.$$

Since $\frac{1}{n}\sum_{i=1}^{n}\beta_i(y_{[i]-1} - \delta x_{[i]-1})^2$ converges to its expected value as n becomes large (see the Appendix), we have the following result.

LEMMA 3: The MML estimator, $\hat{\phi}(\mu,\delta,\sigma)$ is conditionally (for known μ, δ and σ) the MVB estimator with variance

$$\frac{(p-\frac{3}{2})\frac{\sigma^2}{np}}{\frac{1}{n}\sum_{i=1}^{n}\beta_i(y_{[i]-1} - \delta x_{[i]-1})^2}\,.$$

The result follows from the fact that

$$\frac{\partial\, ln\, L^*}{\partial\,\phi} = \frac{2np}{k\sigma^2}\left[\frac{1}{n}\sum_{i=1}^{n}\beta_i(y_{[i]-1} - \delta x_{[i]-1})^2\right]\{\hat{\phi}(\mu,\delta,\sigma) - \phi\}$$

where $\hat{\phi}(\mu, \delta, \sigma) = K_0 + \sigma L_0$ and

$$K_0 = \frac{\sum_{i=1}^n \beta_i (y_{[i]} - \delta x_{[i]})(y_{[i]-1} - \delta x_{[i]-1}) - \mu \sum_{i=1}^n \beta_i (y_{[i]-1} - \delta x_{[i]-1})}{\sum_{i=1}^n \beta_i (y_{[i]-1} - \delta x_{[i]-1})^2}$$

$$L_0 = \frac{\sum_{i=1}^n \alpha_i (y_{[i]-1} - \delta x_{[i]-1})}{\sum_{i=1}^n \beta_i (y_{[i]-1} - \delta x_{[i]-1})^2}.$$

COMMENT: In certain practical situations, one is primarily interested in statistical inferences about μ, δ and σ. In that regard, we have the following results for $\hat{\mu}$, $\hat{\delta}$ and $\hat{\sigma}$ (given by (17)–(19)).

<u>THEOREM 1:</u> For a given ϕ, the MML estimators, $\hat{\mu}$, $\hat{\delta}$ and $\hat{\sigma}$ are asymptotically unbiased with covariance matrix

$$COV_\phi(\hat{\mu}, \hat{\delta}, \hat{\sigma}) \simeq \frac{(p+1)\sigma^2}{n(p-\frac{1}{2})} \begin{bmatrix} \frac{(p-3/2)}{p}\left(1 + \frac{\overline{u}^2}{s_{uu}}\right) & -\frac{(p-3/2)}{p}\frac{\overline{u}}{s_{uu}} & 0 \\ -\frac{(p-3/2)}{p}\frac{\overline{u}}{s_{uu}} & \frac{(p-3/2)}{p}\frac{1}{s_{uu}} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \quad (23)$$

where $u_i = x_i - \phi x_{i-1}$ $(1 \le i \le n)$, $n\overline{u} = \sum_{i=1}^n u_i$ and $n\, s_{uu} = \sum_{i=1}^n (u_i - \overline{u})^2$.

Proof: The asymptotic unbiasedness follows from Taylor series expansions (Kendall and Stuart 1979, p52) of $\frac{\partial\, ln\, L^*}{\partial\, \mu}$, $\frac{\partial\, ln\, L^*}{\partial\, \delta}$ and $\frac{\partial\, ln\, L^*}{\partial\, \sigma}$. For example, a Taylor expansion yields the following result (for large n):

$$E(\hat{\mu}) \simeq \mu - E\left(\frac{\frac{\partial\, ln\, L^*}{\partial\, \mu}}{\frac{\partial^2\, ln\, L^*}{\partial\, \mu^2}}\right). \quad (24)$$

Since $E(\partial\, ln\, L^* / \partial\, \mu) = 0$ (see the Appendix), $E(\hat{\mu}) \simeq \mu$. Similarly, $E(\hat{\delta}) \simeq \delta$ and $E(\hat{\sigma}) \simeq \sigma$ for large n.

The information matrix $I_\phi(\hat{\mu}, \hat{\delta}, \hat{\sigma})$ is given by the expected values of the second derivatives $-E(\partial^2\, ln\, L^* / \partial\, \mu^2)$, $-E(\partial^2\, ln\, L^* / \partial\, \mu \partial \delta), \cdots$ Their asymptotic values can readily be evaluated; see the Appendix. The inverse I^{-1} then gives the asymptotic covariance matrix (23).

<u>Remark:</u> In practice, the unknown parameter ϕ is replaced by $\hat{\phi}$. Thus u_i above are replaced by $\hat{u}_i = x_i - \hat{\phi} x_{i-1}$ $(1 \le i \le n)$.

<u>THEOREM 2:</u> For a given ϕ, the asymptotic distribution of $\sqrt{n}(\hat{\mu} - \mu, \hat{\delta} - \delta)$ is bivariate normal with mean vector $(0, 0)$ and covariance matrix

$$COV_\phi(\hat{\mu}, \hat{\delta}) = \frac{(p+1)(p-\frac{3}{2})\sigma^2}{n(p-\frac{1}{2})} \begin{bmatrix} 1+\frac{\overline{u}^2}{s_{uu}} & -\frac{\overline{u}}{s_{uu}} \\ -\frac{\overline{u}}{s_{uu}} & \frac{1}{s_{uu}} \end{bmatrix} \qquad (p \geq 2)\,. \tag{25}$$

Proof: Since the mixed cumulants of $\frac{\partial\, ln\, L^*}{\partial\, \mu}$, $\frac{\partial\, ln\, L^*}{\partial\, \delta}$ are completely determined by (Bartlett 1953)

$$\kappa_{ij} = -E\left(\frac{\partial^{i+j}\; ln\; L^*}{\partial\, \mu^i\, \partial\, \delta^j}\right)$$

and $\kappa_{ij} = 0$ for all $i + j \geq 3$, the joint distribution of $\frac{\partial\, ln\, L^*}{\partial\, \mu}$ and $\frac{\partial\, ln\, L^*}{\partial\, \delta}$ is bivariate normal (asymptotically). The result then follows from the fact that $\hat{\mu}$ and $\hat{\delta}$ are linear functions of $\frac{\partial\, ln\, L^*}{\partial\, \mu}$ and $\frac{\partial\, ln\, L^*}{\partial\, \delta}$.

COROLLARY: The marginal distribution (asymptotic) of $\hat{\delta}$ is normal with mean δ and variance $(p+1)(p-\frac{3}{2})\sigma^2/p(p-\frac{1}{2})S_{uu}$, $S_{uu} = ns_{uu}$.

Relative Efficiency: The asymptotic covariance matrix of the normal-theory (i.e., $p = \infty$ in (4)) estimators of μ and δ, for a given ϕ, is

$$\frac{\sigma^2}{n} \begin{bmatrix} 1+\frac{\overline{u}^2}{s_{uu}} & -\frac{\overline{u}}{s_{uu}} \\ -\frac{\overline{u}}{s_{uu}} & \frac{1}{s_{uu}} \end{bmatrix} .$$

For a given ϕ, the asymptotic relative efficiency of the normal-theory (Gaussian) estimators of μ and δ to the MML estimators is, therefore,

$$RE = \frac{(p+1)(p-\frac{3}{2})}{p(p-\frac{1}{2})} \qquad (p \geq 2)$$

which is always less than 1 unless $p = \infty$ in which case the MML estimators reduce to the Gaussian estimators and $RE = 1$.

LEMMA 4: The MML estimators, $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$ are asymptotically unbiased and their asymptotic covariance matrix is I^{-1}, where $I(\hat{\mu}, \hat{\delta}, \hat{\sigma}, \hat{\phi})$ is the matrix:

$$\frac{np(p-\frac{1}{2})}{(p+1)(p-\frac{3}{2})\sigma^2} \begin{bmatrix} 1 & \overline{u} & 0 & \frac{\mu}{1-\phi} \\ \overline{u} & \frac{1}{n}\sum_{i=1}^n u_i^2 & 0 & \frac{\mu\overline{u}}{1-\phi} + \frac{\delta\sum_{i=1}^n d_i u_i}{n(1-\phi)} \\ 0 & 0 & \frac{2(p-\frac{3}{2})}{p} & 0 \\ \frac{\mu}{1-\phi} & \frac{\mu\overline{u}}{1-\phi} + \frac{\delta\sum_{i=1}^n d_i u_i}{n(1-\phi)} & 0 & \frac{\sigma^2}{1-\phi} + \frac{\mu^2}{(1-\phi)^2} + \frac{\delta^2\sum_{i=1}^n d_i^2}{n(1-\phi)^2} \end{bmatrix} \tag{26}$$

where $d_i = x_i - x_{i-1}$ and $u_i = x_i - \phi x_{i-1}$ $(1 \le i \le n)$.

Proof: The unbiasedness follows from Taylor series expansions, as in (24). The elements of the information matrix (26) are the values of $-E(\partial^2 \ln L^*/\partial \mu^2)$, $-E(\partial^2 \ln L^*/\partial \mu \partial \delta), \cdots$, $-E(\partial^2 \ln L^*/\partial \phi^2)$ which can readily by evaluated from results given in the Apendix.

COMMENT: The estimators $\hat{\mu}$, $\hat{\delta}$ and $\hat{\phi}$ are asymptotically uncorrelated with $\hat{\sigma}$. This is indeed an interesting result.

REMARK: The information matrix (26) is exactly the same as that of the ML (maximum likelihood) estimators given by $-E(\partial^2 \ln L/\partial \mu^2)$, $-E(\partial^2 \ln L/\partial \mu \partial \delta), \cdots, -E(\partial^2 \ln L/\partial \phi^2)$. This was to be expected since the MML estimators are asymptotically equivalent to the ML estimators.

6. SIMULATIONS

For $p \ge 2$, the MML estimators $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$ are asymptotically fully efficient and are, therefore, more efficient than the Gaussian estimators unless $p = \infty$ (normal innovations) in which case the MML estimators are identical to the Gaussian estimators. To investigate their efficiencies for small sample sizes, we considered n = 20, 30, 60 and 100. The following x-values (common to all y-samples) were considered which represent a very wide range of symmetric designs:

(a). the values x_i $(i = 0, 1, \cdots, n)$ generated from a uniform distribution $U(-1, 1)$,

(b). the values x_i generated from a normal distribution $N(0, 1)$ (see also Tan and Lin 1993),

(c). the values x_i generated from a Cauchy distribution.

The design points are listed in Tables 1, 2 and 3. respectively, for $n = 30$. The values of p considered were p = 1, 1.5, 2.5, 3.5, 4.5, 6 and 10. It may be noted, however, that for $p \le 3$ a few extreme β_i coefficients in (11) are negative as a consequence of which $\hat{\sigma}$ may cease to be real for some samples; see also Tiku and Suresh (1992) and Vaughan (1992a). To remedy this situation, as suggested by Tiku and Suresh (1992), whenever, $\beta_i < 0$ it is equated to zero and so is the corresponding α_i. For $2 \le p \le 3$, however, very few β_i coefficients are negative (and are small in magnitude) and equating them (and the corresponding α_i) to zero has no defrimental effect on their biases and efficiencies; see, for example, Vaughan (1992a). For $1 \leq p < 2$, the extreme α_i and β_i are very small since they are of the order $1/t_{(i)}$ and $1/t_{(i)}^2$, respectively, and $t_{(i)}$ are large; see Vaughan (1992b, 1994).

For the simulations we chose, without any loss of generality, $\mu = 0$, $\delta = 1$ and $\sigma = 1$ in the model (1). We simulated from 10,000 Monte Carlo runs the means and the variances and the covariances of the estimators. The values of the mean, (bias)^2 and MSE (mean square error) are reported in Tables 1–3 from $n = 30$ and $p = 1.5$, 2.5, 3.5, 4.5 and 6; covariances are not reported since their values turned out to be very small due to the fact that $\mu = 0$ and $\sum_{i=1}^{n} x_i \simeq 0$ and $\sum_{i=1}^{n} x_{i-1} \simeq 0$ (symmetry of the designs (a)–(c)). Only the values for $n = 30$ and $p = 1.5$, 2.5, 3.5 and 6 are given. The standard errors in these values are well within ± 0.003. It is clear that the MML estimators are considerably more efficient than the Gaussian estimators. In fact, the relative efficiency of the Gaussian estimators as determined by the ratio of the sum of their mean square errors to that of the MML estimators is quite low and decreases as $|\phi|$ increases or n increases. We have not reported the results for $n = 20$, 60 and 100 for conciseness. For $p = 1$ Gaussian estimators have zero efficiency, and for $p = 10$ they are only slightly less efficient than the MML estimators. That is the reason for not reporting the values for $p = 1$ and 10. It may be noted that biases in both the Gaussian and the MML estimators for μ, δ and σ are generally very low. As an estimator of ϕ, however, the Gaussian estimator has a larger downward bias than the MML estimator. For $n \geq 50$, the simulated variances of $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$ were found to be only slightly larger than those given by the asymptotic covariance matrix (26). Realizing that the MVB (minimum variance bound) estimators do not exist for small n, it follows that the MML estimators are highly efficient. As expected, the Gaussian and MML estimators of δ are most affected by the type of design used. Other than $\hat{\delta}$, the MML estimators $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$ are amazingly stable since their MSE hardly change from design to design (Table 1–3).

7. ROBUSTNESS

Due to practical necessities, the issue of robustness is very important and has received a great deal of attention in recent years; see, for example, Huber (1981) and Tiku et al (1986). A robust estimator is fully efficient (or nearly so) for an assumed model but maintains high efficiency for plausible alternatives to that model (estimation robustness). To investigate the robustness of the estimators $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$, we assumed the underlying distribution to be $f(a; p)$ with $p = 3.5$ (essentially Student's t distributions with 6 df). Of course, any other finite value of p can be chosen with similar results. The value $p = 3.5$ is however, of particular interest since it lies between the two extremes, $p = 1$ (Cauchy) and $p = \infty$ (normal). The first four moments of $f(a; 3.5)$ are all finite but all its even moments of order greater than four do not exist. We considered a large number of alternatives considered by Huber (1981), Tiku et al (1986), Tan and Lin (1993), and Bian and Tiku (1995). We found the MML estimators remarkably robust. For illustration, we report our simulated values in Table 4 for the following alternatives to $f(a; 3.5)$, for design (b); $\mu = 0$, $\delta = 1$ and $\sigma = 1$:

The family $f(a;p)$ with (1) $p = 2$, (2) $p = 2.5$, (3) $p = 4.5$, (4) $p = 6$;

(5) Dixon's scale outlier model with $(n - r)$ observations coming from $N(0, 1)$ and r observations (we do not know which) coming from $N(0, 4^2)$; $r = [\frac{1}{2} + 0.1n]$.

(6) Mixture model with 90% observations coming from $N(0, 1)$ and 10% observations (chosen at random) coming from $f(a; 2.5)$.

The random numbers generated from (5) were divided by $\sqrt{2.5}$ so that the population variance is 1. We report the results only for $n = 30$ for conciseness. The values for $n = 20$, 60 and 100 are not reported for conciseness. In fact, the relative efficiencies of the Gaussian estimators as compared to the MML estimators are lower for $n = 60$ and 100 than for $n = 30$ (Table 4). It is clear that the MML estimators are remarkably robust.

COMMENT: The robustness of the MML estimators was to be expected for the following reasons:

For large n, the quantities H and L in (18) and (20) and $B/\sqrt{nC}$ in (19) are very small so that $\hat{\sigma} \simeq \sqrt{C/n}$. Ignoring H,L and $B/\sqrt{nC}$ (equivalently, ignoring the α-sums in (13)–(16)), the estimators $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$ are the solutions of the equations

$$\begin{aligned} \frac{\partial \ln L^*}{\partial \mu} &\simeq \frac{2p}{k\sigma}\sum_{i=1}^{n}\beta_i\, z_{(i)} = 0\,, \\ \frac{\partial \ln L^*}{\partial \delta} &\simeq \frac{2p}{k\sigma}\sum_{i=1}^{n}(x_{[i]} - \phi x_{[i]-1})\beta_i\, z_{(i)} = 0\,, \\ \frac{\partial \ln L^*}{\partial \sigma} &\simeq -\frac{n}{\sigma} + \frac{2p}{k\sigma}\sum_{i=1}^{n}\beta_i\, z_{(i)}^2 = 0\,, \\ \frac{\partial \ln L^*}{\partial \phi} &\simeq \frac{2p}{k\sigma}\sum_{i=1}^{n}(y_{[i]-1} - \delta x_{[i]-1})\beta_i\, z_{(i)} = 0\,. \end{aligned} \tag{27}$$

As is clear from (11), β_i is of order and $1/t_{(i)}^2$ $(1 \le i \le n)$ and $t_{(i)}^2$ is a decreasing (till the middle value) and then an increasing sequence of positive numbers. Therefore, β_i follow umbrella ordering as a consequence of which $e_{(i)} = \sigma z_{(i)}$ at both ends receive small weights and those in the middle receive large weights. Thus, the influence of extreme residuals is automatically depleted. This phenomenon leads to the robustness of the MML estimators; see also Wong et al (1996). The Gaussian estimators are given by (27) with $\beta_i = 1$ $(1 \leq i \leq n)$. Consequently, all the residuals $e_{(i)}$ $(1 \leq i \leq n)$ receive the same weight and there is no mechanism in them to deplete the influence of extreme residuals in situations where the underlying distribution has long tails.

REMARK: It can be argued that the value of p is not known in practice. A plausible value of p can be obtained from a Q-Q plot of the order statistics of the estimated residuals

$$\hat{e}_i = (y_i - \hat{\phi}_0 y_{i-1}) - \hat{\mu}_0 - \hat{\delta}_0(x_i - \hat{\phi}_0 x_{i-1}) \qquad (1 \leq i \leq n) .$$

$\hat{\mu}_0$, $\hat{\delta}_0$ and $\hat{\phi}_0$ being some convenient initial estimates (Gaussian, for example); see the special Appendix C in Tiku et al (1986). Since the MML estimators are robust, a value of p so determined serves well. Alternatively, p may be obtained from the following equation (with μ, δ, σ and ϕ replaced by the MML estimators $\hat{\mu}$, $\hat{\delta}$, $\hat{\sigma}$ and $\hat{\phi}$, respectively)

$$\frac{\partial \, ln \, L}{\partial \, p} = 0 \tag{28}$$

which can be solved by graphical interpolation as in, for example, Tiku (1968a p137). This is particularly useful if n is large since $\hat{p}$ converqes to p.

FUTURE WORK : Testing, the null hypotheses (i) $\delta = 0$ (ii) $\phi = 0$ and (iii) $\phi = 1$, are of great intest. See also Wong et al (1996) and Tiku and Wong (1997). We are in the process of developing appropriate procedures. We are also in the process of extending the techniques above to skew distributions.

APPENDIX

To find the asymptotic value of m/n, we have $(p \geq 2)$

$$\begin{aligned} lim_{n\to\infty}\frac{m}{n} &= lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{1-(1/k)t_{(i)}^2}{[1+(1/k)\,t_{(i)}^2]^2} = \int_{-\infty}^{\infty}\frac{1-(1/k)z^2}{[1+(1/k)\,z^2]^2}f(z)dz \\ &= \int_{-\infty}^{\infty}\left\{2\left(1+\frac{z^2}{k}\right)^{-2} - \left(1+\frac{z^2}{k}\right)^{-1}\right\}f(z)dz \\ &= \frac{p-\frac{1}{2}}{p+1} \end{aligned} \tag{29}$$

since $\int_{-\infty}^{\infty}\{1+(1/k)z^2\}^{-j}f(z)dz = \sqrt{k}\Gamma(\frac{1}{2})\Gamma(j-\frac{1}{2})/\Gamma(j)$;

$$f(z) = \frac{1}{\sqrt{k}\beta(\frac{1}{2}, p-\frac{1}{2})}\left(1+\frac{z^2}{k}\right)^{-p} \qquad -\infty < z < \infty\,. \tag{30}$$

To find the expected values of the frist and second derivatives of $ln\,L^*$, We consider in particular $E(\partial\, ln\, L^*/\partial\,\phi)$ and $-E(\partial^2\, ln\, L^*/\partial\,\phi^2)$. The expected values of all other derivatives can be obtained in a similar fashion. Now, $t_{(i)} = E\{z_{(i)}\}$ and z_i and y_{i-1} $(1 \le i \le n)$ are independent of each other and complete sums are invariant to ordering. Asmptotically, therefore $(p > 2)$

$$\begin{aligned} E\left(\frac{\partial\, ln\, L^*}{\partial\,\phi}\right) &\simeq \frac{2p}{k\sigma}E\sum_{i=1}^{n}(y_{[i]-1}-\delta x_{[i]-1})\left\{\frac{(2/k)z_{(i)}^3}{[1+(1/k)\,z_{(i)}^2]^2}+\frac{z_{(i)}-(1/k)\,z_{(i)}^3}{[1+(1/k)\,z_{(i)}^2]^2}\right\} \\ &= \frac{2p}{k\sigma}\sum_{i=1}^{n}E(y_{i-1}-\delta x_{i-1})E\left\{\frac{z_i+(1/k)\,z_i^3}{[1+(1/k)\,z_i^2]^2}\right\} = 0 \end{aligned} \tag{31}$$

since the expected value of any odd function $\eta(z)$ is zero, i.e. $\int_{-\infty}^{\infty}\eta(z)\,f(z)\,dz = 0$,

$$\begin{aligned} -E\left(\frac{\partial^2\, ln\, L^*}{\partial\,\phi^2}\right) &= \frac{2p}{k\sigma^2}E\sum_{i=1}^{n}\beta_i(y_{[i]-1}-\delta x_{[i]-1})^2 \\ &\simeq \frac{2p}{k\sigma^2}\sum_{i=1}^{n}E(y_{i-1}-\delta x_{i-1})^2E\left\{\frac{1-(1/k)z_i^2}{[1+(1/k)\,z_i^2]^2}\right\} \\ &= \frac{2p}{k\sigma^2}\sum_{i=1}^{n}E(y_{i-1}-\delta x_{i-1})^2\,\frac{p-\frac{1}{2}}{p+1} \\ &= \frac{p(p-\frac{1}{2})}{(p+1)(p-\frac{3}{2})}\sum_{i=1}^{n}\{E(y_{i-1}^2)-2\delta x_{i-1}E(y_{i-1})+\delta^2x_{i-1}^2\} \\ &= \frac{p(p-\frac{1}{2})}{(p+1)(p-\frac{3}{2})}\sum_{i=1}^{n}\left\{\frac{\sigma^2}{1-\phi^2}+\left[\frac{\mu+\delta(x_i-\phi x_{i-1})}{1-\phi}\right]^2\right. \\ &\qquad \left. -2\delta x_{i-1}\left[\frac{\mu+\delta(x_i-\phi x_{i-1})}{1-\phi}\right]\delta^2x_{i-1}^2\right\} \\ &= \frac{np(p-\frac{1}{2})}{(p+1)(p-\frac{3}{2})}\left\{\frac{\sigma^2}{1-\phi^2}+\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\mu+\delta(x_i-\phi x_{i-1})}{1-\phi}-\delta x_{i-1}\right]^2\right\} \\ &= \frac{np(p-\frac{1}{2})}{(p+1)(p-\frac{3}{2})}\left\{\frac{\sigma^2}{1-\phi^2}+\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\mu+\delta(x_i-x_{i-1})}{1-\phi}\right]^2\right\} \end{aligned} \tag{32}$$

which reduces to the last element in the matrix (25) since for large n

$$\frac{1}{n}\sum_{i=1}^{n} x_i \simeq \frac{1}{n}\sum_{i=1}^{n} x_{i-1} \, .$$

ACKNOWLEDGEMENT

The senior author would like to thank the NSERC for a research grant.

References

Anderson, R.L. (1949), "The problem of autocorrelation in regression analysis," *J Amer Statist Assoc,* 44, 113-127.

Barnett, V.D. (1966a), "Evaluation of the maximum likelihood estimator when the likelihood equation has multiple roots," *Biometrika,* 53, 151-165.

Barnett, V.D. (1966b), "Order statistics estimators of the location of the Cauchy distribution," *J Amer Statist Assoc,* 61, 1205-1218.

Bartlett, M.S. (1953), "Approximate confidence intervals," *Biometrika,* 40, 12-19.

Bhattacharyya, G.K. (1985), "The asymptotics of maximum likelihood and related estimators based on Type II censored data," *J Amer Statist Assoc,* 80, 398-404.

Beach, C.M., and Mackinnon, J.G. (1978), "a maximum likelihood procedure for regression with autocorrelated errors," *Econometrika,* 46, 51-58.

Bian, G., and Tiku, M.L. (1995), "Bayesian inference based on robust priors and MML estimators, Part I, symmetric location–scale distributions," *Statistics,* (to appear).

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." In *Robust Regression,* (Eds., K.D. Lawrence and J.L. Arthur): Marcel Dekker, New York.

Dielman, T.E., and Pfaffenberger, R.C. (1989), "Efficiency of ordinary least square for linear model with autocorrelation," *J Amer Statist Assoc,* 84, 248.

Durbin, J. (1960), "Estimation of parameters in time-series regression model," *J Roy Stat Soc B,* 22, 139-153.

Gallant, A.R., and Goebel, J.J. (1976), "Nonlinear regression with autocorrelated error," *J Amer Statist Assoc,* 71, 961-967.

Hamilton, J.D. (1994), *Time series analysis,* Princeton University Press, New Jersey.

Huber, P.J. (1981), *Robust Statistics,* John Wiley, New York.

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

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,* 81-98. Internat. Lib. Econom: Routledge and Kegan Paul, London.

Maller, R.A. (1989), "Regression with autoregressive errors—some asymptotic results," *Statistics,* 20, 23-39.

Nagaraj, N.K., and Fuller, W.A. (1992), "Least sqaures estimation of the linear model with autoregressive errors," *New Directions in time series analysis, Part I,* 215-225, IMA Vol Math Appl, 45: Springer, New York.

Pearson, E.S., and Hartley, H.O. (1972), *Biometrika Tables for Statisticians,* Vol II: University Press, Cambridge.

Schäffler, S. (1991), "Maximum likelihood estimation for linear regression model with autoregressive errors," *Statistics,* 22, 191-198.

Schneider, H. (1986), *Truncated and censored samples from normal populations,* New York Marcel Dekker.

Smith, W.B., Zeis, C.D., and Syler, G.W. (19739), "Three parameter lognormal estimation from censored data," *J Indian Statistical Association,* 11, 15-31.

Tan, W.Y. (1985), "On Tiku's robust procedure – a Bayesian insight," *J Statist Plann and Inf,* 11, 329-340.

Tan, W.Y., and Lin, V. (1993), "Some robust procedures for estimating parameters in an autoregressive model," *Sankhya B,* 55, 415-435.

Tiao, C.H., and Tan, W.Y. (1966), "Bayesian analysis of random-effect models in the analysis of variance, II: effect of autocorrelated error," *Biometrika,* 53, 477-495.

Tiku, M.L. (1967), "Estimating the mean and standard deviation from censored normal samples," *Biometrika,* 54, 155-165.

Tiku, M.L. (1968a), "Estimating the parameters of log-normal distribution from censored samples," *J Amer Stat Assoc,* 63, 134-140.

Tiku, M.L. (1968b), "Estimating the parameters of normal and logistic distributions from censored samples," *Aust J Statist,* 10, 64-74.

Tiku, M.L. (1970), "Monte Carlo study of some simple estimators in censored normal samples," *Biometrika,* 57, 207-211.

Tiku, M.L. (1980), "Robustness of MML estimators based on censored samples and robust test statistics," *J Stat Plann and Inf,* 4, 123-43.

Tiku, M.L. and Kumar, S. (1981), "Expected values and variances and covariances of order statistics for a family of symmetrical distributions (Student's t)," *Selected Tables in Mathematical Statistics 8, American mathematical Society, Providence,* RI: 141-270.

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., and Wong, W.K. (1997), "Testing for a unit root in an AR(1) model using three and four moment approximation: symmetric distributions," *Commun Stat Simul,* (to appear).

Vaughan, D.C. (1992a), "On the Tiku-Suresh method of estimation," *Commun Stat Theory Meth,* 21, 451-469.

Vaughan, D.C. (1992b), "Expected values, variances and covariances of order statistics for Student's t distribution with two degrees of freedom," *Commun Stat Simul,* 21, 391-404.

Vaughan, D.C. (1994), "The exact values of the expected values, variances and covariances of the order statistics from the Cauchy distribution," *J Stat Comput Simul,* 49, 21-32.

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.

Wong, W.K., Tiku, M.L., and Bian, G. (1996), "Time series models with non-normal innovations, symmetric location-scale distributions," Technical Report.