Open Journal of Statistics, 2014, 4, 620-629
Published Online September 2014 in SciRes. http://www.scirp.org/journal/ojs
http://dx.doi.org/10.4236/ojs.2014.48058

Modified Maximum Likelihood Estimation in Autoregressive Processes with Generalized Exponential Innovations

Bernardo Lagos-Álvarez[1], Guillermo Ferreira[1], Emilio Porcu[2]

[1]Department of Statistics, Universidad de Concepción, Concepción, Chile
[2]Department of Mathematics, University Federico Santa María, Valparaíso, Chile
Email: bla@udec.cl, gferreir@udec.cl, emilio.porcu@usm.cl

Received 19 June 2014; revised 8 July 2014; accepted 26 July 2014

Abstract

We consider a time series following a simple linear regression with first-order autoregressive errors belonging to the class of heavy-tailed distributions. The proposed model provides a useful generalization of the symmetrical linear regression models with independent error, since the error distribution covers both correlated innovations following a Generalized Exponential distribution. Furthermore, we derive the modified maximum likelihood (MML) estimators as an efficient alternative for estimating model parameters. Finally, we investigate the asymptotic properties of the proposed estimators. Our findings are also illustrated through a simulation study.

Keywords

Autoregressive Time Series Model, Maximum Likelihood, Modified Maximum Likelihood, Least Squares, Generalized Exponential

1. Introduction

The common model for a stationary time series is the stationary and invertible autoregressive model of order $p\left(AR\left(p\right)\right)$ where the usual assumption is that the innovations $\{\epsilon_t\}$ are identically and independently distributed (IID) according to a Gaussian distribution with zero mean and variance $\sigma^2 > 0$.

Recent and past literatures agree in that the assumption of Gaussianity is a way too restrictive in order to deal with applications (see [1] and [2] with the references therein). On the other hand, [3] assumed $\{\epsilon_t\}$ has a Laplace distribution and computes the maximum likelihood (ML) estimators by using iterative methods. [2] have used the modified likelihood function proposed by [4] which is based on censored normal samples [5] and have

How to cite this paper: Lagos-Álvarez, B., Ferreira, G. and Porcu, E. (2014) Modified Maximum Likelihood Estimation in Autoregressive Processes with Generalized Exponential Innovations. *Open Journal of Statistics*, **4**, 620-629.
http://dx.doi.org/10.4236/ojs.2014.48058

studied the robustness properties of the resulting estimators. In this context, [6] generated non-Gaussian distributions through transformations of a Gaussian variate.

[7] considered the Huber *M*-estimation, which is valid under heavy-tailed symmetric distributions, and uses different forms of contaminated Gaussian to compute the influence functionals (IF) of parameter estimates and gross-error sensitivity for the IF. In this context, [8] and [9] have studied the rate of convergence of the least squares (LS) estimators. It may be noted that *M*-estimation is not valid for skewed distributions, and has the problem of inefficient estimates for short-tailed symmetric distributions; this has been widely shown by [1] in the classical framework of IID observations.

[10] obtained approximations to some likelihood functions in the context of state space models as considered by [11]. Besides, [12] considered an asymmetric Laplace distribution for the innovations of an autoregressive and moving average model and of a generalized autoregressive conditional heteroscedastic model.

The main proposal of our paper is based on the use of modified likelihood as introduced by [13] [14] and [15] under the framework of IID observations, in order to estimate the parameters in the context of simple linear regression with stationary and invertible autoregressive errors of order one with innovations represented by Generalized Exponential distribution; for more details on these distributions the reader refers to [16]. This method is notorious for giving asymptotically fully efficient estimators (for example, see [17]-[20]).

The outline of the paper is as follows. In Section 2 we define the regression linear model with autoregressive errors, where the underlying distribution of the innovations is a Generalized Exponential distribution. In Section 3 we propose the MML estimators as a powerful methodology to deal with ML estimators which are intractable in the case of a Generalized Exponential distribution. In Section 4 we study the asymptotic properties of the proposed estimators. The main advantages of the proposed estimators are discussed via simulation studies in Section 5. Finally discussions and observations appear in Section 6 of the proposed model and the specific numerical results, attaching an Appendix which displays the details of asymptotic results.

2. The Model

We denote $\{Y_t, t = 0, \pm 1, \cdots\}$ a time series and the following model

$$\begin{aligned} Y_t &= \mu^* + \delta X_t + \eta_t, \\ \eta_t &= \phi \eta_{t-1} + \epsilon_t. \end{aligned} \tag{1}$$

where X_t is the value of a fixed design variable X at time t, η_t is the error, assumed to be modeled through a non-Gaussian stationary autoregressive model, μ^* is a constant, ϕ is the autoregressive coefficient, with $|\phi| < 1$, and ϵ_t is the innovation, distributed according to a Generalized Exponential distribution (GEd), given by

$$f(\epsilon_t; \lambda, \alpha) = \alpha \lambda \left(1 - \mathrm{e}^{-\lambda \epsilon_t}\right)^{\alpha - 1} \mathrm{e}^{-\lambda \epsilon_t}, \ \epsilon_t \geq 0, \ \alpha > 0, \ \lambda > 0. \tag{2}$$

The corresponding cumulative distribution function is given by

$$F(\epsilon_t; \lambda, \alpha) = \left(1 - \mathrm{e}^{-\lambda \epsilon_t}\right)^{\alpha}, \ \epsilon_t \geq 0, \ \alpha > 0. \tag{3}$$

Notably, λ and α play, respectively, the role of scale and shape parameters. The $\mathrm{GEd}(\lambda, \alpha)$ has a similar form to the Gamma and Weibull distributions. See the survey in [21] for some recent developments on GEd, distributions.

3. Modified Maximum Likelihood Estimators

The model in Equation (1) can be written as

$$Y_t - \phi Y_{t-1} = \mu + \delta\left(X_t - \phi X_{t-1}\right) + \epsilon_t, \ \{\epsilon_t\} \sim \mathrm{IIDGEd}(\lambda, \alpha), \tag{4}$$

or

$$\phi(B) Y_t = \mu + \delta \phi(B) X_t + \epsilon_t, \tag{5}$$

whit $\mu := \mu^*(1 - \phi)$, $\phi(B) = 1 - \phi B$ is the autoregressive polynomial, and B is the backward shift operator.

Conditional on $Y_0 = y_0$, the likelihood function for the parameter vector $\boldsymbol{\beta}^{\mathrm{T}} = (\mu, \delta, \phi, \alpha, \lambda)$ in model (4) is given by

$$L(\boldsymbol{\beta}) = \alpha^n \lambda^n \prod_{t=1}^{n} \left(1 - \mathrm{e}^{-z_t}\right)^{\alpha - 1} \mathrm{e}^{-z_t}, \tag{6}$$

where $z_t = \lambda \epsilon_t$, with $\epsilon_t = \phi(B) y_t - \mu - \delta\phi(B) x_t$ or $z_t = \lambda(w_t - \mu)$, w_t given by

$$w_t = \phi(B) y_t - \delta\phi(B) x_t. \tag{7}$$

The log-likelihood is given by

$$l(\boldsymbol{\beta}) = n\left[\ln(\alpha) + \ln(\lambda)\right] - \sum_{t=1}^{n} z_t + (\alpha - 1)\sum_{t=1}^{n} \ln\left(1 - \mathrm{e}^{-z_t}\right). \tag{8}$$

For convenience we introduce at this point the following reparameterization: $\lambda := 1/\sigma$ and $\alpha = -b$. Then the density function of ϵ_t is given by

$$f(\epsilon_t; b, \sigma) = -\frac{b\mathrm{e}^{-\epsilon_t/\sigma}}{\sigma\left(1 - \mathrm{e}^{-\epsilon_t/\sigma}\right)^{b+1}}, \quad \epsilon_t > 0, \tag{9}$$

where $\sigma > 0$ and $b < 0$. Its cumulative distribution function is

$$F(\epsilon_t; b, \sigma) = \left(1 - \mathrm{e}^{-\epsilon_t/\sigma}\right)^{-b}, \quad \epsilon_t > 0. \tag{10}$$

Now $z_t = \epsilon_t/\sigma$, and note that $\{Z_t\} \sim \text{IID GEd}(\sigma = 1, b)$ is the standardized member of the GE family. The log-likelihood for the parameter vector $\boldsymbol{\beta}^{\mathrm{T}} = (\mu, \delta, \phi, \sigma, b)$ then becomes

$$l(\boldsymbol{\beta}) = n\left[\ln(-b) + \ln(\sigma)\right] - \sum_{t=1}^{n} z_t + (b + 1)\sum_{t=1}^{n} \ln\left(1 - \mathrm{e}^{-z_t}\right). \tag{11}$$

Also note that if we consider the parameter b as fixed, then the log-likelihood for the reduced parameter vector, $\boldsymbol{\beta}_1^{\mathrm{T}} = (\mu, \delta, \phi, \sigma)$, is proportional to

$$l(\boldsymbol{\beta}_1) \propto -n\ln(\sigma) - \sum_{t=1}^{n} z_t - (b + 1)\sum_{t=1}^{n} \ln\left(1 - \mathrm{e}^{-z_t}\right). \tag{12}$$

For notational simplicity, let us write $h(z) := \left(\mathrm{e}^{z} - 1\right)^{-1}$. Then, direct inspection shows that first derivatives of the log-likelihood function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_1$ can be written as:

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\beta})}{\partial \mu} &= \frac{\partial l(\boldsymbol{\beta}_1)}{\partial \mu} = \frac{1}{\sigma}\left[n + (b+1)\sum_{t=1}^{n} h(z_t)\right], \\
\frac{\partial l(\boldsymbol{\beta})}{\partial \delta} &= \frac{\partial l(\boldsymbol{\beta}_1)}{\partial \delta} = \frac{1}{\sigma}\left[\sum_{t=1}^{n}(x_t - \phi x_{t-1}) + (b+1)\sum_{t=1}^{n}(x_t - \phi x_{t-1}) h(z_t)\right], \\
\frac{\partial l(\boldsymbol{\beta})}{\partial \phi} &= \frac{\partial l(\boldsymbol{\beta}_1)}{\partial \phi} = \frac{1}{\sigma}\left[\sum_{t=1}^{n}(y_{t-1} - \delta x_{t-1}) + (b+1)\sum_{t=1}^{n}(y_{t-1} - \delta x_{t-1}) h(z_t)\right], \\
\frac{\partial l(\boldsymbol{\beta})}{\partial \sigma} &= \frac{\partial l(\boldsymbol{\beta}_1)}{\partial \sigma} = \frac{1}{\sigma}\left[-n + \sum_{t=1}^{n} z_t + (b+1)\sum_{t=1}^{n} z_t h(z_t)\right], \\
\frac{\partial l(\boldsymbol{\beta})}{\partial b} &= \frac{n}{b} + \sum_{t=1}^{n} z_t + \sum_{t=1}^{n} \ln\left[h(z_t)\right].
\end{aligned} \tag{13}$$

The likelihood equations are expressions in terms of intractable functions $\ln\left(\mathrm{e}^{z} - 1\right)$, which lead no explicit solutions, using as alternative numeric iterative methods for get the solutions.

In order to obtain efficient closed form estimators, we consider Tiku's method of modified likelihood estimation, which is by now well established, see [22] (Chapter 6). For given values of μ, δ, b, and ϕ, let $Z_{1:n} \le Z_{2:n} \le \cdots \le Z_{n:n}$ be the order statistics of $Z_1, \cdots, Z_n$. Let $\mu_{1:n} \le \mu_{2:n} \le \cdots \le \mu_{n:n}$, with $\mu_{t:n} = \mathbb{E}\{Z_{t:n}\}$, be the expected values of the standardized order statistics.

A standard Taylor expansion of $h(z)$ around $z=\mu_{t:n}$ up to first order allows us to obtain

$$h(z_{t:n}) \approx h(\mu_{t:n}) + [z_{t:n}-\mu_{t:n}] \left.\frac{\partial h(z)}{\partial z}\right|_{z=\mu_{t:n}} = a_t - b_t z_{t:n}, \quad t=1,\cdots,n, \tag{14}$$

where $a_t = \left(e^{\mu_{t:n}}-1\right)^{-1} + b_t\mu_{t:n}$ and $b_t = e^{\mu_{t:n}}\left(e^{\mu_{t:n}}-1\right)^{-2}$. A closed form expression for $\mu_{t:n}$ has been calculated by [23], namely

$$\mu_{t:n} = \frac{\Gamma(n+1)}{\Gamma(t)\Gamma(n-t+1)} \sum_{i=0}^{n-t} \frac{(-1)^i}{t+i} \binom{n-t}{i} \left[\Psi\left(1-b(t+i)\right) - \Psi(1)\right], \quad t=1,\cdots,n, \tag{15}$$

where $\psi(x) := \mathrm{d}\ln\Gamma(x)/\mathrm{d}x$ is the Digamma function. However, for large *n*, and using the Delta Method, we have the well-known approximation for $\mu_{t:n}$ for sufficiently large *n*, $\mu_{t:n} \approx \lambda_{t:n}$, with

$$\lambda_{t:n} = F^{-1}\left(\mathbb{E}\left\{F\left(Z_{t:n}\right)\right\}\right) = -\ln\left[1-\left(\frac{t}{n+1}\right)^{-1/b}\right], \tag{16}$$

where F^{-1} is the inverse of the cumulative distribution function of ε_t, see for instance [22]. Since $h(z)$ is locally linear ([14] [15]), under some very general regularity conditions, $z_{t:n}$ converges to $\mu_{t:n}$ as the sample size becomes large, in a small interval not containing the zero value.

Plugging (14) into (13), we obtain the approximated derivative of the log-likelihood function for $\boldsymbol{\beta}_1$, which can be written as

$$\frac{\partial l(\boldsymbol{\beta}_1)}{\partial \mu} \simeq \frac{\partial l^*(\boldsymbol{\beta}_1)}{\partial \mu} = \frac{1}{\sigma}\left[n + (b+1)\sum_{t=1}^{n}\left(a_t - b_t z_{t:n}\right)\right], \tag{17}$$

$$\frac{\partial l(\boldsymbol{\beta}_1)}{\partial \delta} \simeq \frac{\partial l^*(\boldsymbol{\beta}_1)}{\partial \delta} = \frac{1}{\sigma}\left[\sum_{t=1}^{n}\left(x_{[t]} - \phi x_{[t]-1}\right) + (b+1)\sum_{t=1}^{n}\left(x_{[t]} - \phi x_{[t]-1}\right)\left(a_t - b_t z_{t:n}\right)\right], \tag{18}$$

$$\frac{\partial l(\boldsymbol{\beta}_1)}{\partial \phi} \simeq \frac{\partial l^*(\boldsymbol{\beta}_1)}{\partial \phi} = \frac{1}{\sigma}\left[\sum_{t=1}^{n}\left(y_{[t]-1} - \delta x_{[t]-1}\right) + (b+1)\sum_{t=1}^{n}\left(y_{[t]-1} - \delta x_{[t]-1}\right)\left(a_t - b_t z_{t:n}\right)\right], \tag{19}$$

$$\frac{\partial l(\boldsymbol{\beta}_1)}{\partial \sigma} \simeq \frac{\partial l^*(\boldsymbol{\beta}_1)}{\partial \sigma} = \frac{1}{\sigma}\left[-n + \sum_{t=1}^{n} z_{t:n} + (b+1)\sum_{t=1}^{n} z_{t:n}\left(a_t - b_t z_{t:n}\right)\right]. \tag{20}$$

The zeros of the above system of equations are the MML estimators of $\boldsymbol{\beta}_1$. For the sake of clarity, let $\Delta_{[t]} = y_{[t]} - \delta x_{[t]}$, where $y_{[t]}$ and $x_{[t]}$ are the concomitants, the associate values of y and x for $z_{t:n}$, of $z_{t:n} = \left[\phi(B)y_{[t]} - \delta\phi(B)x_{[t]} - \mu\right]/\sigma$. Then, from Equations (17) and (18) we get

$$\begin{pmatrix} \sum_{t=1}^{n} b_t & \sum_{t=1}^{n} b_t\phi(B)x_{[t]} \\ \sum_{t=1}^{n} b_t\phi(B)x_{[t]} & \sum_{t=1}^{n} b_t\left(\phi(B)x_{[t]}\right)^2 \end{pmatrix} \begin{pmatrix} \mu \\ \delta \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{n} b_t\phi(B)y_{[t]} - \sigma\sum_{t=1}^{n}\left(\frac{1}{b+1}+a_t\right) \\ \sum_{t=1}^{n} b_t\phi(B)x_{[t]}\phi(B)y_{[t]} - \sigma\sum_{t=1}^{n}\left(\frac{1}{b+1}+a_t\right)\phi(B)x_{[t]} \end{pmatrix}.$$

Then, from the Equations (17) and (19) we get

$$\begin{pmatrix} \sum_{t=1}^{n} b_t & \sum_{t=1}^{n} b_t\Delta_{[t-1]} \\ \sum_{t=1}^{n} b_t\Delta_{[t-1]} & \sum_{t=1}^{n} b_t\Delta_{[t-1]}^2 \end{pmatrix} \begin{pmatrix} \mu \\ \phi \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{n} b_t\Delta_{[t]} - \sigma\sum_{t=1}^{n}\left(\frac{1}{b+1}+a_t\right) \\ \sum_{t=1}^{n} b_t\Delta_{[t]}\Delta_{[t-1]} - \sigma\sum_{t=1}^{n}\left(\frac{1}{b+1}+a_t\right)\Delta_{[t-1]} \end{pmatrix}.$$

Defining the *n*-dimensional vectors, $1=(1,\cdots,1)^{\mathrm{T}}$, $\boldsymbol{a}=(b+1)^{-1}1+(a_1,\cdots,a_n)^{\mathrm{T}}$, $\boldsymbol{y}=\left(y_{[1]},\cdots,y_{[n]}\right)^{\mathrm{T}}$, and $\boldsymbol{x}=\left(x_{[1]},\cdots,x_{[n]}\right)^{\mathrm{T}}$, the identities above lead to the following expressions for the MML estimators of $\boldsymbol{\beta}_1$:

$$\begin{pmatrix} \hat{\mu} \\ \hat{\delta} \end{pmatrix} = \left(U^{\mathrm{T}}DU\right)^{-1}\left[U^{\mathrm{T}}D\phi(B)\boldsymbol{y} - \hat{\sigma}U^{\mathrm{T}}\boldsymbol{a}\right], \tag{21}$$

$$\begin{pmatrix}\hat{\mu}\\ \hat{\phi}\end{pmatrix} = \left(V^{\mathrm{T}}DV\right)^{-1}\left[V^{\mathrm{T}}D\left(\boldsymbol{y}-\hat{\delta}\boldsymbol{x}\right)-\hat{\sigma}V^{\mathrm{T}}\boldsymbol{a}\right], \tag{22}$$

$$\hat{\sigma} = \frac{c+\sqrt{c^2-4nd}}{2n}, \text{ or } \hat{\sigma} = \frac{c-\sqrt{c^2-4nd}}{2n}, \tag{23}$$

where

$$\underset{n\times n}{\underbrace{D}} = \text{Diagonal}\left(b_1,\cdots,b_n\right), \ \underset{n\times 2}{\underbrace{U}} = \left(1,\hat{\phi}(B)\boldsymbol{x}\right), \ \underset{n\times 2}{\underbrace{V}} = \left(1,B\left(\boldsymbol{y}-\hat{\delta}\boldsymbol{x}\right)\right),$$

and

$$c = \sum_{t=1}^{n}\left[1+(b+1)a_t\right]\left(\hat{\phi}(B)y_{[t]}-\hat{\mu}-\hat{\delta}\hat{\phi}(B)x_{[t]}\right),$$

$$d = (b+1)\sum_{t=1}^{n}\left(\hat{\phi}(B)y_{[t]}-\hat{\mu}-\hat{\delta}\hat{\phi}(B)x_{[t]}\right)^2 b_t.$$

Furthermore, note that setting the expression for $\partial l(\boldsymbol{\beta})/\partial b$ in (13) to zero and solving for b, while substituting $z_{t:n} = \left[\hat{\phi}(B)y_{[t]}-\hat{\delta}\hat{\phi}(B)x_{[t]}-\hat{\mu}\right]/\hat{\sigma}$ gives

$$\hat{b}^{-1} = \frac{1}{n}\sum_{t=1}^{n}\ln\left(1-\exp\left\{\left[\hat{\mu}-\hat{\phi}(B)\left(y_{[t]}-\hat{\delta}x_{[t]}\right)\right]/\hat{\sigma}\right\}\right). \tag{24}$$

We note that the coefficients b_t's are positive. It is expected if the $e_{t:n}$'s have positive values, then $\ln\left(1-\exp\left\{e_{t:n}\right\}\right)$'s are all negatives, so $\hat{b}$ is negative. And if d is negative, no complex roots occur for $\hat{\sigma}$. Moreover $c^2-4nd > c^2$, and $\sqrt{c^2-4nd} > c$, resulting as an estimator for σ, $\hat{\sigma} = \left(c+\sqrt{c^2-4nd}\right)/2n$. We observe that these estimates involve the μ parameter.

These facts suggest that it is possible to obtain MML estimators of $\boldsymbol{\beta}$ by using the following iterative procedure. As a starting point, consider the LS estimator for (μ,δ,ϕ,γ), with $\gamma = -\delta\phi$, which is given by

$$\left(\tilde{\mu},\tilde{\delta},\tilde{\phi},\tilde{\gamma}\right)^{\mathrm{T}} = S^{-1}L, \tag{25}$$

where

$$L = \left(1^{\mathrm{T}}Y \ \ Y^{\mathrm{T}}X \ \ Y^{\mathrm{T}}Y_1 \ \ Y^{\mathrm{T}}X_1\right)^{\mathrm{T}}, \tag{26}$$

$$Y = \left(y_1,\cdots,y_n\right)^{\mathrm{T}}, \ X = \left(x_1,\cdots,x_n\right)^{\mathrm{T}}, \ X_1 = BX = \left(x_0,x_1,\cdots,x_{n-1}\right)^{\mathrm{T}}, \tag{27}$$

$$Y_1 = \begin{pmatrix} y_0\\ y_1\\ \vdots\\ y_{n-1}\end{pmatrix} = BY, \text{ and } S = \begin{pmatrix} 1^{\mathrm{T}}1 & 1^{\mathrm{T}}X & 1^{\mathrm{T}}Y_1 & 1^{\mathrm{T}}X_1\\ 1^{\mathrm{T}}X & X^{\mathrm{T}}X & Y_1^{\mathrm{T}}X & X^{\mathrm{T}}X_1\\ 1^{\mathrm{T}}Y_1 & Y_1^{\mathrm{T}}X & Y_1^{\mathrm{T}}Y_1 & Y_1^{\mathrm{T}}X_1\\ 1^{\mathrm{T}}X & X^{\mathrm{T}}X_1 & Y_1^{\mathrm{T}}X_1 & X^{\mathrm{T}}X\end{pmatrix}. \tag{28}$$

We suggest the following routine for the numerical computation of the MML estimator. Initialize with $y_0 = 0 = x_0$ and $b = -1$ (an exponential distribution).

Step 0. Set $\left(\hat{\mu}^{(0)},\hat{\delta}^{(0)},\hat{\phi}^{(0)},\hat{\gamma}^{(0)}\right) = \left(\tilde{\mu}^{(0)},\tilde{\delta}^{(0)},\tilde{\phi}^{(0)},\tilde{\gamma}^{(0)}\right)$ the LSE from (25).

Step 1. Get $\left(\hat{\mu}^{(1)},\hat{\delta}^{(1)}\right)$ from (21) using $\left(\hat{\phi}^{(0)},\hat{\sigma}^{(0)},b^{(0)}\right)$, and $\hat{\phi}^{(1)}$ from (22) using $\left(\hat{\mu}^{(1)},\hat{\delta}^{(1)},\hat{\sigma}^{(0)},b^{(0)}\right)$. Update $\hat{\sigma}^{(1)}$ with (23); $\hat{b}^{(1)}$ with (24).

Step 2. Evaluate the expressions (15)-(17) in (μ,δ,ϕ,γ) with the initial estimated values $\left(\hat{\mu}^{(0)},\hat{\delta}^{(0)},\hat{\phi}^{(0)},\hat{\gamma}^{(0)}\right)$.

Step 3. Get from (7) the initial estimates for the $w_n = \left\{w_t, t = 1,\cdots,n\right\}$. Sort the set w_n saving the corresponding concomitants values of $y_t, y_{t-1}, x_t, x_{t-1}$, say $y_{[t]}, y_{[t]-1}, x_{[t]}, x_{[t]-1}$.

Step 4. With the values of Step 3, get $\hat{\sigma}^{(0)}$ from (23) and get $\hat{b}^{(0)}$ from (24), to obtain the complete initial values vector $\boldsymbol{\beta}^{(0)}$.

Step *i*. Get $\left(\hat{\mu}^{(i)},\hat{\delta}^{(i)}\right)$ from (21) using $\left(\hat{\phi}^{(i-1)},\hat{\sigma}^{(i-1)},b^{(i-1)}\right)$, and $\hat{\phi}^{(i)}$ from (22) using $\left(\hat{\mu}^{(i)},\hat{\delta}^{(i)},\hat{\sigma}^{(i-1)},b^{(i-1)}\right)$. Update $\hat{\sigma}^{(i)}$ with (23), $b^{(i)}$ with (24).

The steps are repeated until convergence is achieved.

Remark. *The stopping criteria is given by*

$$C\left(\hat{\beta}_s\right)=\left|\hat{\beta}_s-\hat{\beta}_{s-1}\right|<10^{-3}.$$

4. Asymptotic Equivalence and Efficiency

The asymptotic equivalence of MML and ML estimators is based on the fact that $h\left(z_{t:n}\right)-a_t-b_t z_{t:n}$ converges to zero as *n* tends to infinity. Thus, following [17] we have that the differences, $\left(1/n\right)\left|\partial l\left(\boldsymbol{\beta}\right)/\partial\mu-\partial l^*\left(\boldsymbol{\beta}_1\right)/\partial\mu\right|$ and $\left(1/n\right)\left|\partial l\left(\boldsymbol{\beta}\right)/\partial\delta-\partial l^*\left(\boldsymbol{\beta}_1\right)/\partial\delta\right|$ tend to zero asymptotically. Therefore, the MML and ML estimators are asymptotically equivalent.

On the other hand, if we know the values of ϕ and σ, asymptotically, the MML estimators $\hat{\mu}\left(\phi,\sigma\right)$ and $\hat{\delta}\left(\phi,\sigma\right)$ are unbiased for μ and δ. Namely, let $\boldsymbol{\beta}_2=\left(\mu,\delta\right)^{\mathrm{T}}$ be the parameter vector and by applying the standard Taylor expansion in a neighborhood of $\boldsymbol{\beta}_2^*=\left(\mu^*,\delta^*\right)^{\mathrm{T}}$ we have (see [24])

$$\hat{\boldsymbol{\beta}}_2=\boldsymbol{\beta}_2-\left[\frac{\partial^2 l^*\left(\boldsymbol{\beta}_2\right)}{\partial\boldsymbol{\beta}_2^2}\right]_{\boldsymbol{\beta}_2^*}^{-1}\frac{\partial l^*\left(\boldsymbol{\beta}_2\right)}{\partial\boldsymbol{\beta}_2}.$$

Using the results (5.7.5), p. 115 of [25] and Lemma 1 in the Appendix, we show that $\mathbb{E}\left\{\hat{\boldsymbol{\beta}}_2-\boldsymbol{\beta}_2\right\}=0$ for large *n*. The unbiasedness property of $\left(\phi,\sigma\right)$ is analogous to the previous case.

Furthermore, if we know the values of ϕ and σ, the MML estimators $\hat{\mu}\left(\phi,\sigma\right)$ and $\hat{\delta}\left(\phi,\sigma\right)$ are unbiased and normally distributed with variance-covariance matrix

$$\Sigma\left(\phi,\sigma\right)=\frac{\sigma^2}{n}\frac{1}{b\left(b+1\right)}\frac{1}{v_{22}-v_{11}^2}\begin{pmatrix}v_{22} & -v_{12}\\ -v_{12} & v_{11}\end{pmatrix}. \tag{29}$$

knowledge of the values of ϕ and σ. Observe that

$$\lim_{n\to\infty}\frac{1}{n}U^{\mathrm{T}}DU=-b\begin{pmatrix}v_{11} & v_{12}\\ v_{12} & v_{22}\end{pmatrix},$$

with $v_{11}=1$, $v_{12}=\sum_{t=1}^{n}\phi\left(B\right)x_{[t]}$ and $v_{22}=\sum_{t=1}^{n}\left(\phi\left(B\right)x_{[t]}\right)^2$. Thus, we have (29) for the asymptotic variance-covariance matrix of $\hat{\mu}\left(\phi,\sigma\right)$.

The asymptotic behavior of variance for $\hat{\sigma}$ (say $\operatorname{Var}\left\{\hat{\sigma}\right\}$) and $\hat{\phi}$ (say $\operatorname{Var}\left\{\hat{\phi}\right\}$) can be deduced from the arguments in [26]. We can thus show that $\operatorname{Var}\left\{\hat{\phi}\right\}\simeq 1/v_{33}$, where

$$v_{33}=\frac{n}{\sigma^2}\left[1-\mathcal{M}^{(1)}\left(0;b\right)-2b\mathcal{M}^{(1)}\left(-1;b+1\right)+\frac{b\left(b+1\right)}{b+2}\mathcal{M}^{(2)}\left(-1;b+2\right)\right].$$

Analogously, we have $\operatorname{Var}\left\{\hat{\sigma}\right\}\simeq 1/v_{44}$, where

$$v_{44}=\frac{n}{\sigma}\left[\mu_\eta^{(1)}+\left(b+1\right)\left\{\mu_\eta^{(1)}\left(\frac{b}{b+1}\mathcal{M}\left(-1;b+1\right)+\frac{b}{b+2}\mathcal{M}^{(1)}\left(-1;b+2\right)\right)\right.\right.$$
$$\left.\left.+\frac{1}{\sigma}\left(\mu^{*2}+2\mu_\eta^{(1)}+\mu_\eta^{(2)}\right)\frac{b}{b+2}\mathcal{M}\left(-1;b+2\right)\right\}\right].$$

where $\mathcal{M}^{(j)}\left(t;b+k\right)$ is *j*th derivative of the moment generating function of $GEd\left(1,b+k\right)$, $\mu_\eta^{(r)}$ is the *r*th non-central moment of $\left\{\eta_t\right\}$ given by the Lemma 2 of the Appendix for $r=1,2$.

Table 1. The simulated values of mean, bias square and mean square error of the LS estimators $(\tilde{\mu},\tilde{\delta},\tilde{\phi},\tilde{\sigma})$ and the MML estimators $(\bar{\mu},\bar{\delta},\bar{\phi},\bar{\sigma})$ and $n=100$.

		$b=-1$				$b=-2$			
		Mean	Var	Bias	MSE	Mean	Var	Bias	MSE
$\mu=0.0$	$\tilde{\mu}$	−0.000886	0.033584	−0.000886	0.033582	0.986002	0.024213	0.986002	0.996410
	$\bar{\mu}$	0.001300	0.029693	0.000130	0.023692	0.435218	0.018573	0.435218	0.216985
$\delta=1.0$	$\tilde{\delta}$	1.000554	0.034930	0.000554	0.034926	1.002698	0.024228	0.002698	0.024233
	$\bar{\delta}$	1.000043	0.020870	0.000043	0.020868	1.001481	0.015923	0.001481	0.015923
$\phi=-0.8$	$\tilde{\phi}$	−0.787714	0.004023	0.012286	0.004173	−0.786830	0.004093	0.013171	0.004566
	$\bar{\phi}$	−0.790068	0.003618	0.009932	0.003516	−0.786149	0.003588	0.013051	0.004480
$\sigma=1.0$	$\tilde{\sigma}$	3.211265	0.352510	2.211265	5.242168	2.239461	0.176988	1.239461	1.713233
	$\bar{\sigma}$	1.032998	0.008841	0.032998	0.009929	1.055952	0.009012	0.055952	0.012142
$\mu=0.0$	$\tilde{\mu}$	0.001376	0.035637	0.001376	0.035635	1.015149	0.036122	1.015149	1.066645
	$\bar{\mu}$	0.001175	0.025692	0.001175	0.029690	0.462208	0.032993	0.462208	0.252625
$\delta=1.0$	$\tilde{\delta}$	0.999292	0.034422	−0.000708	0.034420	0.999049	0.024585	−0.001951	0.024583
	$\bar{\delta}$	0.998947	0.032174	−0.001053	0.032172	0.998650	0.021924	−0.001350	0.021923
$\phi=0.1$	$\tilde{\phi}$	0.087611	0.012722	−0.012389	0.012375	0.085178	0.029938	−0.014822	0.013157
	$\bar{\phi}$	0.095462	0.010529	−0.004538	0.010548	0.054797	0.011023	−0.005203	0.011049
$\sigma=1.0$	$\tilde{\sigma}$	3.158387	0.327490	2.158387	4.986093	2.192925	0.165809	1.192925	1.588861
	$\bar{\sigma}$	1.025339	0.009202	0.025339	0.009844	1.046575	0.008996	0.046575	0.011165
$\mu=0.0$	$\tilde{\mu}$	−0.001417	0.055431	−0.001417	0.055427	1.403107	0.219596	1.403107	2.188284
	$\bar{\mu}$	0.000813	0.044283	0.000813	0.044279	0.779076	0.127095	0.779076	0.834031
$\delta=1.0$	$\tilde{\delta}$	0.999761	0.035422	−0.002239	0.035419	0.998107	0.025840	−0.001893	0.025841
	$\bar{\delta}$	0.998032	0.019583	−0.001968	0.019585	0.978544	0.016259	−0.001456	0.016260
$\phi=0.8$	$\tilde{\phi}$	0.762102	0.004769	−0.037898	0.006204	0.778123	0.006862	−0.071877	0.011227
	$\bar{\phi}$	0.777242	0.003312	−0.022758	0.005830	0.741970	0.006476	−0.058030	0.009843
$\phi=1.0$	$\tilde{\sigma}$	3.205789	0.353528	2.205789	5.218997	2.362176	0.226820	1.362176	2.082321
	$\bar{\sigma}$	1.032757	0.009014	0.032757	0.010086	1.076499	0.010523	0.076499	0.016374
$\mu=0.0$	$\tilde{\mu}$	0.012309	0.298050	0.018309	0.298172	2.962612	0.421637	2.962612	8.998682
	$\bar{\mu}$	0.015270	0.249466	0.015270	0.249674	1.252334	0.319509	1.252334	1.887818
$\delta=1.0$	$\tilde{\delta}$	1.005434	0.309207	0.005434	0.309206	0.999057	0.218581	−0.000943	0.218560
	$\bar{\delta}$	0.997079	0.172751	−0.002921	0.172742	1.001011	0.137776	0.001011	0.137763
$\phi=-0.8$	$\tilde{\phi}$	−0.786810	0.004018	0.013190	0.004782	−0.785997	0.004676	0.014353	0.004787
	$\bar{\phi}$	−0.788509	0.004374	0.011491	0.004505	−0.785681	0.004337	0.014219	0.004541
$\sigma=3.0$	$\tilde{\sigma}$	29.042450	29.572310	26.042450	707.778770	20.152190	15.007650	17.152190	309.203850
	$\bar{\sigma}$	3.089355	0.071391	0.089355	0.079368	3.137597	0.079450	0.137597	0.098375
$\mu=0.0$	$\tilde{\mu}$	−0.009324	0.328285	−0.009324	0.328339	3.038606	0.426984	3.038606	9.550080
	$\bar{\mu}$	−0.006528	0.271379	−0.006528	0.271395	1.362092	0.374527	1.362092	2.229783
$\delta=1.0$	$\tilde{\delta}$	0.991023	0.307443	−0.008977	0.307493	0.997002	0.289937	−0.002998	0.285924
	$\bar{\delta}$	0.996085	0.284231	−0.003915	0.284218	0.893295	0.233959	−0.006705	0.233980
$\phi=0.1$	$\tilde{\phi}$	0.085039	0.019795	−0.014961	0.014518	0.099978	0.012785	−0.011022	0.015646
	$\bar{\phi}$	0.094117	0.011222	−0.005883	0.011256	0.088109	0.000979	−0.000891	0.010979
$\sigma=3.0$	$\tilde{\sigma}$	28.484590	26.924310	25.484590	676.385740	19.810370	13.435800	16.810370	296.022870
	$\bar{\delta}$	3.074408	0.072055	0.074408	0.077584	3.126289	0.074052	0.126289	0.089994

5. Simulation Study

In order to have some indications of the robustness aspects of the MML estimates of μ, δ, ϕ and σ against LSE estimates, we performed a small numerical study similar to the one presented by [26] for the generalized logistic model. We consider the following AR(1) Generalized Exponential model:

$$Y_t - \phi Y_{t-1} = \mu + \delta\left(X_t - \phi X_{t-1}\right) + \epsilon_t,\ \{\epsilon_t\} \sim \text{IID GEd}\left(\lambda, \alpha\right) \quad (30)$$

where $X_t \sim N\left(0,1\right)$. Additionally, our simulation study considers different scenarios, sketched as follows:

1) $\mu = 0$, $\delta = 1$, $\phi = -0.8$ and $\sigma = 1$,
2) $\mu = 0$, $\delta = 1$, $\phi = 0.1$ and $\sigma = 1$,
3) $\mu = 0$, $\delta = 1$, $\phi = 0.8$ and $\sigma = 1$,
4) $\mu = 0$, $\delta = 1$, $\phi = -0.8$ and $\sigma = 3$,
5) $\mu = 0$, $\delta = 1$, $\phi = -0.1$ and $\sigma = 3$.

Without loss of generality, we have considered the parameter b as a constant value given by $b = -1$ and -2. The summaries of Monte Carlo study for μ, δ, ϕ and σ, come from the four measures, the mean, 100 × $(\text{Bias})^2$, variance and mean squared error (MSE) for both the LS and the MML estimators. Finally, we use sample size $n = 100$ and 10,000 replications. **Table 1** displays the results from the simulations with the biases, variance and MSE of the parameters estimates. The results suggest that the MML estimators are considerably more efficient than the LS estimators for all parameters.

6. Conclusion

In this paper, we have studied a regression linear model with first-order autoregressive errors belonging to a class of asymmetric distributions; more specifically the underlying distribution for the innovations is a Generalized Exponential distribution. We have developed a complete asymptotic theory for the MML estimators in these models. In addition, we have shown that the MML estimators are robust and efficient, as depicted by the numerical study presented in Section 5 for the AR(1) GE model. We thus claim that the MML estimator is a very good alternative to estimate autoregressive models with asymmetric innovations (see [26] and [27], among others as example). The R codes may be obtained from the authors upon request in order to analyze such models.

Acknowledgements

The first author would like to thank for the support from DIUC 213.014.022-1.0, established by the Universidad de Concepción. The second author gratefully acknowledges the financial support from ECOS-CONICYT C10E03, established by the Chilean Government and DIUC 213.014.021-1.0 from the Universidad de Concepción and the third author was supported by Fodecyt grant 1130647.

References

[1] Tiku, M.L., Tan, W.Y. and Balakrishnan, N. (1986) Robust Inference. Marcel Dekker, Inc., New York.

[2] Tan, W.Y. and Lin, V. (1993) Some Robust Procedures for Estimating Parameters in an Autoregressive Model. *Sankhya B*, **55**, 415-435.

[3] Damsleth, E. and El-Shaarawi, A.H. (1989) ARMA Models with Double Exponentially Distributed Noise. *Journal of the Royal Statistical Society B*, **51**, 61-69.

[4] Tiku, M.L. (1980) Robustness of MML Estimators Based on Censored Samples and Robust Test Statistics. *Journal of Statistical Planning and Inference*, **4**, 123-143. http://dx.doi.org/10.1016/0378-3758(80)90002-6

[5] Tan, W.Y. (1985) On Tiku's Robust Procedure—A Bayesian Insight. *Journal of Statistical Planning and Inference*, **11**, 329-340. http://dx.doi.org/10.1016/0378-3758(85)90038-2

[6] Swift, A.L. (1995) Modelling and Forecasting Time Series with a General Non-Normal Distribution. *Journal of Forecasting*, **14**, 45-66. http://dx.doi.org/10.1002/for.3980140105

[7] Martin, R.D. and Yohai, V.J. (1986) Influence Functionals for Time Series. *Annals of Statistics*, **14**, 781-818. http://dx.doi.org/10.1214/aos/1176350027

[8] Bhansali, R.J. (1997) Robustness of the Autoregressive Spectral Estimate for Linear Process with Infinite Variance. *Journal Time Series Analysis*, **18**, 213-229. http://dx.doi.org/10.1111/1467-9892.00047

[9] Davis, R.A. and Resnick, S. (1986) Limit Theory for the Sample Covariance and Correlation Functions of Moving Averages. *Annals of the Institute of Statistics*, **14**, 533-558. http://dx.doi.org/10.1214/aos/1176349937

[10] Durbin, J. and Koopman, S.J. (1997) Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models. *Biometrika*, **84**, 669-684. http://dx.doi.org/10.1093/biomet/84.3.669

[11] Kitagawa, G. (1987) Non-Gaussian State-Space Modelling of Nonstationary Time Series (with Discussion). *Journal of the American Statistical Association*, **82**, 1032-1063.

[12] Trindade, A.A. and Zhu, Y. (2010) Time Series Models with Asymmetric Laplace Innovations. *Journal of Statistical Computation and Simulation*, **80**, 1317-1333.

[13] Tiku, M.L. (1967) Estimating the Mean and Standard Deviation from Censored Normal Samples. *Biometrika*, **54**, 155-165. http://dx.doi.org/10.2307/2283834

[14] Tiku, M.L. (1968) Estimating the Parameters of Log-Normal Distribution from Censored Samples. *Journal of the American Statistical Association*, **63**, 134-140. http://dx.doi.org/10.2307/2283834

[15] Tiku, M.L. and Suresh, R.P. (1992) A New Method of Estimation for Location and Scale Parameters. *Journal of Statistical Planning and Inference*, **30**, 281-292. http://dx.doi.org/10.1111/1467-842X.00072

[16] Gupta, R.D. and Kundu, D. (1999) Generalized Exponential Distributions. *Australian and New Zealand Journal of Statistics*, **41**, 173-188. http://dx.doi.org/10.1111/1467-842X.00072

[17] Vaughan, D.C. and Tiku, M.L. (2000) Estimation and Hypothesis Testing for a Non-Normal Bivariate Distribution with Applications. *Journal of Mathematical and Computer Modelling*, **32**, 53-67. http://dx.doi.org/10.1016/S0895-7177(00)00119-9

[18] Bhattacharyya, G.K. (1985) The Asymptotics of Maximum Likelihood and Related Estimators Based on Type II Censored Data. *Journal of the American Statistical Association*, **80**, 398-404. http://dx.doi.org/10.1080/01621459.1985.10478130

[19] Tiku, M.L. (1970) Some Notes on the Relationship between the Distribution of Central and Non-Central *F*. *Biometrika*, **57**, 175-179. http://dx.doi.org/10.1093/biomet/57.1.175

[20] Tiku, M.L. (1970) Monte Carlo Study of Some Simple Estimators in Censored Normal Samples. *Biometrika*, **57**, 207-210.

[21] Gupta, R.D. and Kundu, D. (2007) Generalized Exponential Distribution: Existing Methods and Some Recent Developments. *Journal of Statistical Planning and Inference*, **137**, 3537-3547. http://dx.doi.org/10.1016/j.jspi.2007.03.030

[22] Balakrishnan, N. and Cohen, A.C. (1991) Order Statistics and Inference. Academic Press, Waltham.

[23] Raqab, M.Z. and Ahsanullah, M. (2001) Estimation of Location and Scale Parameters of Generalized Exponential Distribution Based on Order Statistics. *Journal of Statistical Computation and Simulation*, **69**, 109-124.

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

[25] Tiku, M.L. and Akkaya, A.D. (2004) Robust Estimation and Hypothesis Testing. New Age International (P) Publishers, New Delhi, 337.

[26] Wong, W.K. and Bian, G. (2005) Estimating Parameters in Autoregressive Models with Asymmetric Innovations. *Statistics and Probability Letters*, **71**, 61-70. http://dx.doi.org/10.1016/j.spl.2004.10.022

[27] Tiku, M.L., Wong, W.K. and Bian, G. (1999) Time Series Models with Asymmetric Innovations. *Communications in Statistics—Theory and Methods*, **28**, 1331-1360. http://dx.doi.org/10.1080/03610929908832360

Appendix

Lemma 1. Let $Z \sim GE(1,b)$, $U = \exp(-Z)$, $q \in \mathbb{N} \cup \{0\}$ and p, $s \in \mathbb{R}$ such that $p > -b$ and $-1 < s < b + p$, then

$$E\left\{Z^q U^s / (1-U)^p\right\} = \frac{b}{b+p} \mathcal{M}^{(q)}(-s; b+p)$$

where $\mathcal{M}^{(j)}(t; b+k)$ is jth derivative of the moment generating function of $\mathrm{GEd}(1, b+k)$.

Proof

$$E_Z\left\{Z^q U^s / (1-U)^p\right\} = \int_0^\infty \frac{z^q \mathrm{e}^{-sz}}{\left(1-\mathrm{e}^{-z}\right)^p} \frac{b\mathrm{e}^{-z}}{\left(1-\mathrm{e}^{-z}\right)^{b+1}} \mathrm{d}z = \int_0^\infty z^q \mathrm{e}^{-sz} \frac{b\mathrm{e}^{-z}}{\left(1-\mathrm{e}^{-z}\right)^{b+p+1}} \mathrm{d}z = \frac{b}{b+p} \mathcal{M}^{(q)}(-s; b+p).$$

Lemma 2. For the process $\{\eta_t\}$ defined as a stationary autoregressive model, $\eta_t = \phi\eta_t + \varepsilon_t$, ϕ is the autoregressive coefficient, with $|\phi| < 1$, and ϵ_t is distributed according to a GEd. The first and second moment are given by

$$\mu_\eta^{(1)} = \frac{\phi}{1-\phi^2} \mu_\varepsilon \text{ and } \mu_\eta^{(2)} = \frac{\phi^2}{1-\phi^2} \mu_\varepsilon^{(2)}.$$

Proof is deduced by using the moment generating function of $\varepsilon \sim \mathrm{GEd}(\sigma, b)$

$$\mathcal{M}(t; \sigma, b) = \frac{\Gamma(1+b)\Gamma(1-\sigma t)}{\Gamma(b-\sigma t+1)}, \quad -b < \sigma t < 1, \tag{31}$$

(see [21]). Moreover, for the $\mu_\eta^{(1)}$, we used

$$\lim_{k\to\infty} E\{\eta_t\} = \lim_{k\to\infty} \phi^k E\{\eta_{t-k}\} + \sum_{j=0}^{k} \phi^k \mu_\varepsilon = \mu_\eta^{(1)}$$

and for $\mu_\eta^{(2)}$, we used

$$\lim_{k\to\infty} E\left\{\mu_t^2\right\} = \lim_{k\to\infty} \sum_{j=0}^{k} \phi^{2k} \mu_\varepsilon^{(2)} = \mu_\eta^{(2)}.$$