ADAPTIVE INFERENCE OF AUTOREGRESSIVE MODELS UNDER NONNORMALITY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BEGÜM YENTÜR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
STATISTICS

SEPTEMBER 2019

Approval of the thesis:

ADAPTIVE INFERENCE OF AUTOREGRESSIVE MODELS UNDER NONNORMALITY

submitted by **BEGÜM YENTÜR** in partial fulfillment of the requirements for the degree of **Master of Science in Statistics Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Ayşen Dener Akkaya
Head of Department, **Statistics** ________________

Prof. Dr. Ayşen Dener Akkaya
Supervisor, **Statistics, METU** ________________

Assoc. Prof. Dr. Özlem Türker Bayrak
Co-Supervisor, **Inter-Curricular Courses, Çankaya Uni.** ________________

Examining Committee Members:

Assoc. Prof. Dr. Ebru Yüksel Haliloğlu
Management, TOBB Uni. of Economics and Technology ________________

Prof. Dr. Ayşen Dener Akkaya
Statistics, METU ________________

Prof. Dr. M. Qamarul Islam
Statistics, METU ________________

Date: 10.09.2019

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Begüm Yentür

Signature:

ABSTRACT

ADAPTIVE INFERENCE OF AUTOREGRESSIVE MODELS UNDER NONNORMALITY

Yentür, Begüm
Master of Science, Statistics
Supervisor: Prof. Dr. Ayşen Dener Akkaya
Co-Supervisor: Assoc. Prof. Dr. Özlem Türker Bayrak

September 2019, 80 pages

Recently, non-normal innovations in autoregressive models have become prevalent in many applications. In this case, it is known that the least squares (LS) estimators are neither efficient nor robust. Also, obtaining maximum likelihood (ML) estimators requires numerical solution which is a formidable task. To overcome these difficulties modified maximum likelihood (MML) estimation technique is used to obtain the estimators of the model parameters. In this method, although explicit solution can be found, the necessity of knowing the shape parameter becomes a drawback especially in machine data processing. That is why, in this thesis adaptive modified maximum likelihood (AMML) methodology which combines MML estimation technique with Huber's M-estimation procedure is used so that the shape parameter can also be estimated. Expectation Maximization (EM) algorithm is also used in order to obtain Maximum Likelihood Estimators (MLEs) numerically. Then, through a simulation study, efficiency and robustness properties of the estimators are discussed and compared with each other. Finally, test statistics are proposed for the crucial parameters of the model. The power comparisons of the test statistics under each estimation technique are presented.

Keywords: Autocorrelation, Modified Maximum Likelihood, Robustness, Regression.

ÖZ

NORMAL OLMAYAN DURUMDA OTOREGRESİF MODELLERİN UYARLAMALI ÇIKARSAMASI

Yentür, Begüm
Yüksek Lisans, İstatistik
Tez Danışmanı: Prof. Dr. Ayşen Dener Akkaya
Ortak Tez Danışmanı: Doç. Dr. Özlem Türker Bayrak

Eylül 2019, 80 sayfa

Son zamanlarda otoregresif modellerde normal olmayan hatalara birçok uygulamada rastlanmaya başlanmıştır. Bu durumda en küçük kareler tahmin edicilerinin etkin ve sağlam olmadıkları bilinmektedir. Ayrıca, en çok olabilirlik tahmin edicilerini elde etmek sayısal çözüm teknikleriyle mümkündür ancak bu da zordur. Bu zorlukların üstesinden gelmek için uyarlanmış en çok olabilirlik tahmin yöntemi geliştirilmiştir. Bu yöntemde, açık bir çözüm bulunabilse de şekil parametresinin biliniyor olma gerekliliği makine veri işleme sürecinde dezavantaj oluşturmaktadır. Bu nedenden dolayı, bu tezde şekil parametresinin de tahmin edilebilmesi için Huber'ın M-tahmin yöntemi ile uyarlanmış en çok olabilirlik (EÇO) tahmin tekniğini birleştiren adaptif uyarlanmış en çok olabilirlik metodolojisi kullanılmıştır. Ayrıca, en çok olabilirlik tahmin edicilerini numerik olarak elde etmek için koşullu beklenti maksimizasyonu algoritması kullanılmıştır. Daha sonra simülasyon çalışmasıyla tüm tahmin edicilerin etkinlik ve sağlamlık özellikleri tartışılmış ve tahmin ediciler birbirleriyle karşılaştırılmıştır. Son olarak modelin kritik parametreleri için test istatistikleri önerilmiştir. Her bir tahmin yöntemi altında test istatistiklerinin güç karşılaştırmaları sunulmuştur.

Anahtar Kelimeler: Otokorelasyon, Uyarlanmış En Çok Olabilirlik, Sağlamlık, Regresyon.

To my beloved family…

ACKNOWLEDGEMENTS

To begin with, I want to express my very profound gratitude to my supervisor Prof. Dr. Ayşen Dener Akkaya for her guidance, patience, kindness and critics throughout the study. I could not be able to finish this thesis without her support and calming attitude.

I am also really thankful to Assoc. Prof. Dr. Özlem Türker Bayrak for her helpful suggestions, valuable feedback, contributions and insightful comments during my thesis.

My completion of this work could not have been accomplished without the support of my special friends Gamze, Ezgi, Görkem, Göksu, Remin and Doğa. They listened to me whenever I am in need. Also, special thanks to my valuable colleagues for their support. I am lucky to have all of you in my life.

Above all, words are inadequate to express my gratitude to my family for showing confidence and patience not just during this process but throughout everything I do. They teach me not to give up. I am always grateful to them for being there for me. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS	
AMML	Adaptive Modified Maximum Likelihood
BAN	Best Asymptotically Normal
ECM	Expectation Conditional Maximization
EM	Expectation Maximization
LS	Least Squares
LSE	Least Squares Estimator
LTS	Long Tailed Symmetric
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator
MML	Modified Maximum Likelihood
MMLE	Modified Maximum Likelihood Estimator
MSE	Mean Squared Error

CHAPTER 1

INTRODUCTION

Statistical methods are commonly used in many fields to analyse the data. One of these methods is regression. It is very prevalent to observe autocorrelation in regression analysis. This problem gives rise to autoregressive models popular in different areas like engineering, finance, economics etc. Most of these models are based on the normality assumption (Durbin, 1960; Kramer, 1980; Schaffler, 1991; Tan and Lin, 1993). But later on, it is seen that the assumption of normality is violated in many applications (Huber, 1981 and Tiku et al., 1986). In this study, we consider the first-order autoregressive model (AR(1)) with a single explanatory variable where the error term is having a distribution from Long Tailed Symmetric (LTS) family. It is known that under non-normality, least squares estimators (LSEs) are neither efficient nor robust (Islam and Tiku, 2004; Bayrak and Akkaya, 2010) and maximum likelihood estimators (MLEs) cannot be obtained analytically since the likelihood functions involve non-linear terms. Therefore, they can only be obtained via numerical solutions by the use of algorithms like Expectation Maximization (EM), Expectation Conditional Maximization (ECM), Newton-Raphson etc. However, due to the convergence problems, using iterative methods is troublous and especially for small samples they induce bias (Barnett, 1966; Puthenpura and Sinha, 1986; Vaughan, 1992). Besides, their properties are not known theoretically so that inference about them becomes doubtful. To avoid these difficulties, modified maximum likelihood (MML) methodology was introduced by Tiku (1967) and developed by Tiku and Suresh (1992). This method linearizes the intractable non-linear terms in the likelihood equations so that the analytical solutions can be obtained. MML estimators have three very desirable properties: (i) explicit functions of sample observations, (ii) almost fully efficient for small sample sizes, (iii) best asymptotically normal (BAN).

Moreover, they are known to be robust. However, MML method assumes known shape parameter. This assumption is unrealistic when machine learning methods are considered. In these cases, it is assumed that the distribution belongs to a member of a broad class of distributions (Hampel et al., 1986) and the shape parameter has to be estimated. This problem is solved by integrating the logic of MML with M-estimators called as Adaptive Modified Maximum Likelihood (AMML) estimation procedure (Tiku and Sürücü, 2009; Dönmez, 2010; Akkaya and Bayrak, 2018).

The aim and the contribution of the thesis is to obtain the AMML estimators and the test statistics for AR(1) model with a single explanatory variable where the error term is having a distribution from LTS family. We compare the estimators' efficiency and robustness properties with four different methods via simulation study. These methods are Least Squares, Maximum Likelihood (ML) with ECM algorithm, Modified Maximum Likelihood, and Adaptive Modified Maximum Likelihood.

This thesis consists of four chapters. An introduction can be found in Chapter 1. In Chapter 2, estimation procedures are explained step by step for each technique and AMML estimators are derived. In Chapter 3, properties of the estimators are discussed via simulation. In Chapter 4, test statistics for the significance of the model parameters are derived. Also, test statistics' power properties are discussed through a simulation study. Finally, in the last chapter the findings are summarized and the work that has been done is concluded.

CHAPTER 2

AUTOREGRESSIVE MODEL AND ESTIMATION OF THE PARAMETERS

In this chapter, estimation methods, namely the least squares, maximum likelihood, modified maximum likelihood and adaptive modified maximum likelihood are mentioned. The EM and ECM algorithms which are numerical estimation methods to be used for obtaining maximum likelihood estimators are also reviewed.

Regression analysis is a widely used statistical method used for modelling and investigating the relationship between variables. More specifically, it aims to describe the relationship between two or more variables of interest.

In regression models, autocorrelation often arise. Correlated errors give rise to autoregressive models. So, autoregressive regression models started to become popular.

For our case, consider a simple autoregressive model

$$y_t = \mu' + \gamma x_t + a_t \tag{2.1}$$

$$a_t = \phi a_{t-1} + \varepsilon_t$$

where

y_t = observed value of a random variable y at time t

x_t = pre-determined value of a non-stochastic design variable at time t

ϕ = autoregressive coefficient ($-1<\phi<1$);

ε_t is assumed to be independent and identically distributed (iid) and called as error term.

Model (2.1) is equivalent to

$$y_t - \phi y_{t-1} = \mu + \gamma(x_t - \phi x_{t-1}) + \varepsilon_t \quad (1 \leq t \leq n). \tag{2.2}$$

This model has numerous applications in different fields. For $\gamma = 0$, equation (2.2) reduces to time series AR(1) model. Because of the parameter $\gamma\phi$, this model is nonlinear. Two models are existent for the initial value, y_0 (Vinod and Shenton, 1996);

Model A: y_0 is constant, $y_0 = 0$ in particular,

Model B: y_0 is random.

Since Model B is flexible, we preferred to work with this. Also, the likelihood functions (conditional to y_0) for Model A and Model B is exactly the same.

Traditionally, errors are assumed to be normal but in many real life problems it is seen that non-normal distributed errors are more common. In this study, it is assumed that the error term ε_t in the model (2.2) have one of the distribution in LTS family of distributions

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{k}} \frac{1}{B\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left(1 + \frac{\varepsilon^2}{k\sigma^2}\right)^{-p}, \qquad -\infty < \varepsilon < \infty \tag{2.3}$$

where p is the shape and σ is the scale parameter. B(a,b)= $\Gamma(a)\Gamma(b)/\Gamma(a+b)$ where $\Gamma(.)$ is the Gamma Function. Note that k=2p-3 and p≥2. Also, E(ε)=0, V(ε)=σ^2 and $\sqrt{\upsilon/k}\,(\varepsilon/\sigma)$ is distributed Student's t with $\upsilon = 2p - 1$ degrees of freedom. The kurtosis value of the distribution is $3(p - 3/2)/(p - 5/2)$. Since LTS is a symmetric distribution, only kurtosis values are shown in Table 2.1 for some values of p.

Table 2.1 Kurtosis Values of LTS Distribution for Different Shape Parameter Values

p	2.5	3.5	5	10	∞
Kurtosis	∞	9	4.2	3.4	3

As clearly seen from the Table 2.1, the kurtosis values are always greater than 3. When p approaches to ∞, the distribution in equation (2.3) reduces to N(0,1). For $1 \leq p < 2$, Var(ε) does not exist.

Different methods can be used to estimate the model parameters.

2.1 Least Squares Estimation

Least squares method is commonly used for parameter estimation in statistics. This method is based on the idea of minimizing the sum of squared errors in order to find the optimal parameter values.

The classical assumptions are:

1. Linearity (Model should be linear in parameters),
2. No endogeneity,
3. Normality and homoscedasticity,
4. No autocorrelation,
5. No multicollinearity.

If these assumptions are violated, robust methods give more reliable estimates.

This method both have advantages and disadvantages. To mention about advantages, no distribution assumption is needed in order to find the estimators. If the errors are iid with E(e)=0 and V(e)=σ^2 then the LSEs are fully efficient as MLEs but the LSEs may lose efficiency if the distribution is not normal. One of the disadvantage is that LSEs are sensitive to outliers. In other words, it is not robust to existence of data anomalies. Another disadvantage is when the errors have a non-zero mean, or the variance is a function of σ^2, the LSEs becomes biased and requires adjustment.

In our case, the least squares estimators of the model (2.2) under LTS can be written as

$$\tilde{\mu} = \frac{\sum_{i=1}^{n} y_i}{n} - \tilde{\phi}\frac{\sum_{i=1}^{n} y_{i-1}}{n} - \tilde{\gamma}\frac{\sum_{i=1}^{n}(x_i - \tilde{\phi}x_{i-1})}{n},$$

$$\tilde{\gamma} = \frac{\sum_{i=1}^{n} u_i v_i - n\,\bar{u}\bar{v}}{\sum_{i=1}^{n} u_i^2 - n\bar{u}^2},$$

$$\tilde{\phi} = \frac{\sum_{i=1}^{n}(y_i - \tilde{\gamma}x_i)(y_{i-1} - \tilde{\gamma}x_{i-1}) - \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{\gamma}x_i)\sum_{i=1}^{n}(y_{i-1} - \tilde{\gamma}x_{i-1})}{\sum_{i=1}^{n}(y_{i-1} - \tilde{\gamma}x_{i-1})^2 - \frac{1}{n}\left[\sum_{i=1}^{n}(y_{i-1} - \tilde{\gamma}x_{i-1})\right]^2},$$

$$\tilde{\sigma} = \frac{\sqrt{\sum_{i=1}^{n}\left[y_i - \tilde{\phi}y_{i-1} - \tilde{\mu} - \tilde{\gamma}\left(x_i - \tilde{\phi}x_{i-1}\right)\right]^2}}{n-3}. \tag{2.4}$$

where $v_i = y_i - \tilde{\phi}y_{i-1},\ u_i = x_i - \tilde{\phi}x_{i-1}$

2.2 Maximum Likelihood Estimation

It assumes known functional form of the distribution. The likelihood function conditional on the initial value, $y_0 = \varepsilon_0/\sqrt{1-\phi^2}$, where ε_0 is independent value and has the same distibution as that of ε_i ($1 \le i \le n$) is

$$L \propto \sigma^{-n}\prod_{i=1}^{n}\left(1 + \frac{\varepsilon_i^2}{k\sigma^2}\right)^{-p}. \tag{2.5}$$

For known p, the likelihood equations are obtained in terms of

$$z_i = \{(y_i - \phi y_{i-1}) - \mu - \gamma(x_i - \phi x_{i-1})\}/\sigma = \varepsilon_i/\sigma. \tag{2.6}$$

After taking logarithm of the likelihood function, partial derivatives of the ln(L) function are calculated with respect to the model parameters and equating them equal to zero give the following:

$$\frac{\partial \ln L}{\partial \mu} = \frac{2p}{k\sigma}\sum_{i=1}^{n} g(z_i) = 0 \tag{2.7}$$

$$\frac{\partial \ln L}{\partial \gamma} = \frac{2p}{k\sigma}\sum_{i=1}^{n} g(z_i)(x_i - \phi x_{i-1}) = 0 \tag{2.8}$$

$$\frac{\partial \ln L}{\partial \phi} = \frac{2p}{k\sigma}\sum_{i=1}^{n} g(z_i)\,[y_{i-1} - \gamma(x_{i-1})] = 0 \tag{2.9}$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma}\sum_{i=1}^{n} g(z_i) z_i = 0 \tag{2.10}$$

where $g(z_i) = \frac{z_i}{1+\left(\frac{z_i^2}{k}\right)}$.

Realizing that equations (2.7)-(2.10) include non-linear function $g(z_i)$, the MLEs cannot be obtained explicitly. However, the estimators can be obtained by using numerical methods.

Asymptotic Properties of MLE

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with a parameter θ, and $\hat{\theta}_{ML}$ denote the MLE of it. Then, under some regularity conditions,

1. $\hat{\theta}_{ML}$ is asymptotically consistent:
$$\lim_{n\to\infty} P(\left|\hat{\theta}_{ML} - \theta\right| > \epsilon) = 0,$$
2. $\hat{\theta}_{ML}$ is asymptotically unbiased:
$$\lim_{n\to\infty} E[\hat{\theta}_{ML}] = \theta,$$
3. As n becomes large, $\hat{\theta}_{ML}$ is approximately normally distributed. More precisely, the random variable
$$\frac{\hat{\theta}_{ML} - \theta}{\sqrt{Var(\hat{\theta}_{ML})}}$$
converges in distribution to N(0,1).

2.3 Expectation Maximization (EM) Algorithm

One of the most popular and efficient algorithms in the literature that can be used to obtain a numerical solution for the nonlinear ML equations is the EM algorithm. It is an iterative method that is first introduced and given its name by Dempster, Laird, Rubin (1977). It depends on missing or latent (unobserved) variables and performs through Expectation step (E-step) and Maximization step (M-step).

Algorithm:

1. Starts with the initial guess of the parameters.
2. E-Step: Use the observed data in order to estimate the value of the missing data.
3. M-Step: Update the parameters by using the complete data generated after the expectation step.
4. Repeat the steps 2 through 4 until convergence.

Advantages of EM

- In each iteration the likelihood function increases for sure.
- Both E-step and M-step are easy to implement in many problems since they do not require heavy preparatory analytical work. It is a derivative free algorithm.
- It can also be used to obtain estimates in the missing data cases.

Disadvantages of EM

- Convergence speed slow down especially when approaching to local optima.
- Both forward and backward probabilities are required (In numerical optimization only forward probabilities are required).
- It is very sensitive to the initial values.

2.3.1 ML Estimation Using Expectation Conditional Maximization (ECM)

EM is impractical if it is difficult to maximize the complete-data log-likelihood function. To overcome this deficiency, the ECM algorithm is introduced by Meng and Rubin (1993). It is the extension of EM algorithm which consist of the EM and a sequence of conditional maximization (CM) steps instead of M-step.

Haghbin and Nematollahi (2013) considered hierarchical model for Student's t distribution which helps providing simple analytic expressions for the ML functions. The hierarchical model is given in the proposition.

Proposition: Let $\tau_1, \dots, \tau_n$ be iid from $\Gamma\left(\frac{\upsilon}{2}, \frac{\upsilon}{2}\right)$ and $\varepsilon_t|\tau_t \sim N(\mu, \frac{(\upsilon-2)\sigma^2}{\upsilon\tau_t})$ for t=1,2,…,n, where j≠t, τ_t are independent of ε_j and $\varepsilon_t|\tau_t$ are independent of $\varepsilon_j|\tau_{j.}$ Then (i) $\varepsilon_t \sim \left(\upsilon, \mu, \sqrt{\frac{\upsilon-2}{\upsilon}}\sigma\right)$; and (ii) $\varepsilon_t|\tau_t \sim \Gamma\left(\frac{\upsilon+1}{2}, \frac{\upsilon}{2} + \frac{\upsilon(\varepsilon_t-\mu)^2}{2(\upsilon-2)\sigma^2}\right)$.

In the proposition, the latent variable τ_t refers to the missing data, υ is the degrees of freedom, Γ is the Gamma Function and $\varepsilon_t = \left(y_t - \mu - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right)$.

Nduka et al. (2018) adopted this procedure to the multiple autoregression model with Student's t distribution. In this study, we use the procedure of Nduka et al. (2018) and apply it to simple autoregression model given in equation (2.2). The methodology is given as follows:

Let Y denotes the observed data, Z = $\{\tau_t,\ t = 1, \dots, n\}$ denote the latent variables and $\theta = (\mu, \gamma, \phi, \sigma, \upsilon)$ be collectively the vector of unknown parameters so that the complete data becomes {Z,Y} with joint densities.

$$f(\varepsilon_1, \dots, \varepsilon_n, \tau_1, \dots \tau_n|\varepsilon_0, \theta) = \prod_{t=1}^{n} f_N\ (\varepsilon_t|\tau_t, \varepsilon_{t-1}; \theta)\ f_G\ (\tau_t;\upsilon). \tag{2.11}$$

As mentioned in Nduka et al. (2018), equation (2.11) can be called as a state space model which consists Normal as the structural distribution, τ_t as the missing parameters and Gamma as the mixing distribution. Due to this mixture structure, equation (2.11) can be split into two different functions as

$$l_c\,(\theta|\varepsilon,\tau) = l_N\,(\mu,\gamma,\phi,\sigma|y,\tau) + l_G\,(\upsilon|\tau) \tag{2.12}$$

where first part of the equation (2.12) refers to the log-likelihood of the conditional Normal distribution as

$$l_N\,(\theta|y,\tau) \propto -\frac{n}{2}\log(\sigma^2) - \frac{\upsilon}{2(\upsilon-2)\sigma^2} - \sum_{t=1}^{n}\left(y_t - \mu - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right)^2 \tau_t \tag{2.13}$$

and the second part refers to the log-likelihood of the conditional Gamma distribution

$$l_G = \left(\frac{n}{2}\right) v\log\left(\frac{\upsilon}{2}\right) - n\log\Gamma\left(\frac{\upsilon}{2}\right) + \left(\frac{\upsilon}{2}\right)\sum_{t=1}^{n}\log(\tau_t) - \left(\frac{\upsilon}{2}\right)\sum_{t=1}^{n}\tau_t - \sum_{t=1}^{n}\log(\tau_t). \tag{2.14}$$

The complete-data log-likelihood, l_c for the hierarchical representation can be obtained by combining equations (2.13) and (2.14) as follows:

$$\begin{aligned} l_c\left(\theta; Y, Z^{(1)}\right) = {} & \sum_{t=1}^{n}\left\{\frac{1}{2}\left[\upsilon\log\left(\frac{\upsilon}{2}\right) + \log(\sigma^2)\right] - \log\Gamma\left(\frac{\upsilon}{2}\right)\right\} + \frac{\upsilon}{2}\sum_{t=1}^{n}\{\log(\tau_t) - \tau_t\} \\ & - \sum_{t=1}^{n}\log(\tau_t) - \frac{\upsilon}{2\sigma^2(\upsilon-2)}\left[\sum_{t=1}^{n}\{y_t - \mu - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\}^2\tau_t\right]. \end{aligned} \tag{2.15}$$

Therefore, the estimation of the parameters can be split into two parts as the CM of $(\mu,\gamma,\phi,\sigma,\upsilon)$ from the conditional normal log-likelihood function and the CM of υ from the Gamma log-likelihood function where both parts are conditioned on the missing variable, τ. Thus, to obtain the estimates, the missing parameter τ must be estimated first. The estimation procedures are given below:

E-step: We need conditional posterior distribution of τ_t, given in equation (2.16), in order to derive the latent variable τ_t and the parameters $(\mu,\gamma,\phi,\sigma,\upsilon)$.

$$f(\tau_t|\varepsilon_t) = \frac{\tau_t^{\frac{\upsilon+1}{2}-1}\left(\frac{\upsilon(\upsilon-2)\sigma^2+\upsilon\varepsilon_t^2}{2(\upsilon-2)\sigma^2}\right)}{\Gamma\left(\frac{\upsilon+1}{2}\right)} \exp\left\{-\left(\frac{\upsilon(\upsilon-2)\sigma^2+\upsilon\varepsilon_t^2}{2(\upsilon-2)\sigma^2}\right)\tau_t\right\} \quad (2.16)$$

From the posterior distribution above, the conditional expectations at kth iteration are obtained:

$$w_t = E\,(\tau_t|Y,\theta) = \frac{\left(\upsilon^{(k)}+1\right)\left(\upsilon^{(k)}-2\right)}{\upsilon^{(k)}}\;\frac{\sigma^{(k)^2}}{(\upsilon^{(k)}-2)+\sigma^{(k)^2}+\varepsilon_t^{(k)^2}} \quad (2.17)$$

$$q_t = E\,[\log(\tau_t)|Y,\theta] = \psi\left(\frac{\left(\upsilon^{(k)}+1\right)}{2}\right) - \log\left(\frac{\upsilon^{(k)}}{2}\right) - \log\left(\frac{\left(\upsilon^{(k)}-2\right)\sigma^{(k)^2}+\varepsilon_t^{(k)^2}}{(\upsilon^{(k)}-2)\sigma^{(k)^2}}\right) \quad (2.18)$$

where $\psi(y) = d\log\Gamma(y)/dy$ indicates the digamma function and $\varepsilon_t = \left(y_t - \mu - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right)$.

CM-step: With the mixing parameter τ obtained in E-Step and υ=$\upsilon^{(k)}$, maximizing $l_N\,(\mu,\gamma,\phi,\sigma,\upsilon|y,\tau)$ in equation (2.15) with respect to $(\mu,\gamma,\phi,\sigma,\upsilon)$ yields the likelihood functions:

$$\frac{\partial l_N\,(\theta|y,\tau)}{\partial\mu} = \frac{\upsilon^{(k)}}{(\upsilon^{(k)}-2)\sigma^2}\sum_{t=2}^{n} w_t\left(y_t - \beta_0 - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right) = 0 \quad (2.19)$$

$$\frac{\partial l_N\,(\theta|y,\tau)}{\partial\gamma} = \frac{\upsilon^{(k)}}{(\upsilon^{(k)}-2)\sigma^2}\sum_{t=2}^{n} w_t(x_t - \phi x_{t-1})\left(y_t - \beta_0 - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right) = 0$$

(2.20)

$$\frac{\partial l_N\,(\theta|y,\tau)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{\upsilon^{(k)}}{2\sigma^4(\upsilon^{(k)}-2)}\sum_{t=2}^{n} w_t\left(y_t - \beta_0 - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right)^2 = 0$$

(2.21)

$$\frac{\partial l_N\,(\theta|y,\tau)}{\partial\phi} = \frac{\upsilon^{(k)}}{(\upsilon^{(k)}-2)\sigma^2}\sum_{t=2}^{n}(y_{t-1} - \gamma x_{t-1})\left(y_t - \beta_0 - \phi y_{t-1} - \gamma(x_t - \phi x_{t-1})\right) = 0.$$

(2.22)

These equations give close form solution represented as follows:

$$\beta^{(k+1)} = M^{-1}M_0 \tag{2.23}$$

where

$$\beta^{(k+1)} = \binom{\mu}{\gamma}$$

$$M = \begin{pmatrix} \sum w_t & \sum w_t\left(x_t - \phi^{(k)}x_{t-1}\right) \\ \sum w_t\left(x_t - \phi^{(k)}x_{t-1}\right) & \sum w_t\left(x_t - \phi^{(k)}x_{t-1}\right)^2 \end{pmatrix}$$

and

$$M_0 = \begin{pmatrix} \sum w_t\left(y_t - \phi^{(k)}y_{t-1}\right) \\ \sum w_t\left(x_t - \phi^{(k)}x_{t-1}\right)\left(y_t - \phi^{(k)}y_{t-1}\right) \end{pmatrix} \tag{2.24}$$

$$\phi^{(k+1)} = \frac{\sum_{t=2}^{n} w_t\left(y_{t-1} - \gamma^{(k)}x_{t-1}\right)\left(y_t - \mu^{(k)} - \gamma^{(k)}x_t\right)}{\sum_{t=1}^{n} w_t\left(y_{t-1} - \gamma^{(k)}x_{t-1}\right)^2}$$

$$\sigma^{(k+1)} = \left(\frac{\upsilon^{(k)}}{n\left(\upsilon^{(k)} - 2\right)}\sum_{t=1}^{n} w_t \varepsilon_t^{(k+1)^2}\right)^{\frac{1}{2}}. \tag{2.25}$$

Actually, Nduka et al. (2018) considered the unknown shape parameter case and added a second CM-step to the algorithm for this purpose as follows:

CM-Step 2: Given the mixing parameter τ with $\mu = \mu^{(k+1)}$, $\gamma = \gamma^{(k+1)}$, $\phi = \phi^{(k+1)}$ and $\sigma = \sigma^{(k+1)}$ fixed in each condition, υ can be estimated by maximizing $l_G\,(\upsilon|\tau)$ in equation (2.14) with respect to υ numerically. For this purpose, an algorithm like Newton-Raphson can be used. The derivatives can be shown as

$$\frac{\partial l_G}{\partial \upsilon} = \frac{n}{2} + \frac{n}{2}\log\left(\frac{\upsilon}{2}\right) - \frac{n}{2}\psi\left(\frac{\upsilon}{2}\right) + \frac{1}{2}\sum_{t=1}^{n}\left(q_t - w_t\right) \tag{2.26}$$

$$\frac{\partial^2 l_G}{\partial \upsilon^2} = \frac{n}{2\upsilon} - \frac{n}{4}\psi'\left(\frac{\upsilon}{2}\right). \tag{2.27}$$

Alternatively, υ can be estimated by finding the zero roots of

$$1 + \log\left(\frac{\upsilon}{2}\right) - \psi\left(\frac{\upsilon}{2}\right) + \frac{1}{n}\sum_{t=1}^{n}(q_t - w_t). \tag{2.28}$$

However, in their simulation study, they have seen that for some samples the degrees of freedom (df) parameter could not be estimated efficiently and are very biased. As mentioned in Nduka et al. (2018) since better estimates of the parameters can be obtained with any value of df, it is better to use ECM procedure with known df. Therefore, it is preferred to use ECM with known df in this thesis.

2.4 Modified Maximum Likelihood Estimation

Solving the equations (2.7)-(2.10) by iterative or numerical methods is known to have some drawbacks. Such as;

i. non-convergence,
ii. wrong-convergence (e.g., they correspond to local rather than global maximum of the likelihood function),
iii. multiple roots problems.

To avoid these difficulties, the modified maximum likelihood estimation technique is introduced by M. L. Tiku (1967) and developed by Tiku and Suresh (1992).

The steps of MML method proceeds as follows:

1. Express the likelihood equations (2.7)-(2.10) in terms of order statistics $z_{(i)} = \{(y_{[i]} - \phi y_{[i]-1}) - \mu - \gamma(x_{[i]} - \phi x_{[i]-1})\}/\sigma$ arranged in ascending order of magnitude. Note that, $(y_{[i]}, y_{[i]-1}; x_{[i]}, x_{[i]-1})$ are the concomitants of $z_{(i)}$ $(1 \leq i \leq n)$,

i.e., the pair $(y_{[j]}, y_{[j]-1})$ (j=[i]) associated with the i^{th} ordered value $z_{(i)}$ so that order of the original data is not lost. The numerical values of the equations do not change since the complete sums do not affect from the ordering.

2. Linearize the function $g(z_i)$ by using the first two terms of a Taylor series expansion around $t_{(i)} = E(z_{(i)})$.

$$g(z_{(i)}) \cong g(t_{(i)}) + (z_{(i)} - t_{(i)}) \left\{\frac{\partial g(z)}{\partial z}\right\}_{z=t(i)}$$

$$\cong \alpha_i + \beta_i z_{(i)} \quad (1 \le i \le n). \tag{2.29}$$

Remark: Under some very general regularity conditions; $z_{(i)}$ converges to $t_{(i)}$ as n becomes large (Tiku and Akkaya, 2004).

From the equation (2.29), the resulting α_i and β_i are obtained as

$$\alpha_i = (2/k)t_{(i)}^3/\{1 + (1/k)\, t_{(i)}^2\}^2, \quad \beta_i = \{1 - (1/k)\, t_{(i)}^2\}/\{1 + (1/k)\, t_{(i)}^2\}^2 . \tag{2.30}$$

Approximate values of $t_{(i)}$ are calculated by using the following equation:

$$\frac{\Gamma(p)}{\sqrt{k}\Gamma(1/2)\Gamma(p-1/2)} \int_{-\infty}^{t_{(i)}} \left(1 + \frac{z^2}{k}\right)^{-p} dz = \frac{i}{n+1} (1 \le i \le n). \tag{2.31}$$

In evaluating (2.31), consider $\sqrt{(k/\upsilon)z}$ is student's t distributed with υ =2p-1. Note that

$$t_{(i)} = t_{(n-i+1)},\ \alpha_i = -\alpha_{n-i+1},\ \beta_i = -\beta_{n-i+1}\ (1 \le i \le n) \text{ and } \textstyle\sum_{i=1}^{n} \alpha_i = 0$$

which is valid for symmetric distributions.

3. Finally, modified maximum likelihood equations can be obtained by incorporating $g(z_i)$ in (2.7)-(2.10) to give the following.

$$\frac{\partial \ln L}{\partial \mu} \cong \frac{\partial \ln L^*}{\partial \mu} = \frac{2p}{k\sigma} \sum_{i=1}^{n} (\alpha_i + \beta_i z_{(i)}) = 0 \tag{2.32}$$

$$\frac{\partial \ln L}{\partial \gamma} \cong \frac{\partial \ln L^*}{\partial \gamma} = \frac{2p}{k\sigma}\sum_{i=1}^{n}\left(\alpha_i + \beta_i z_{(i)}\right)(x_i - \phi x_{i-1}) = 0 \tag{2.33}$$

$$\frac{\partial \ln L}{\partial \phi} \cong \frac{\partial \ln L^*}{\partial \phi} = \frac{2p}{k\sigma}\sum_{i=1}^{n}\left(\alpha_i + \beta_i z_{(i)}\right)(y_{i-1} - \gamma x_{i-1}) = 0 \tag{2.34}$$

$$\frac{\partial \ln L}{\partial \sigma} \cong \frac{\partial \ln L^*}{\partial \sigma} - \frac{n}{\sigma} + \frac{2p}{k\sigma}\sum_{i=1}^{n}\left(\alpha_i + \beta_i z_{(i)}\right)z_i = 0. \tag{2.35}$$

Equations (2.32)-(2.35) are solved to obtain the following MML estimators:

$$\hat{\mu} = \bar{v}_{[.]} - \hat{\gamma}\,\bar{u}_{[.]}, \qquad \hat{\gamma} = G + H\hat{\sigma}, \qquad \hat{\phi} = K + D\hat{\sigma},$$

$$\hat{\sigma} = \left\{-B + \sqrt{B^2 + 4nC}\right\}\Big/\left[2\sqrt{n(n-3)}\right], \tag{2.36}$$

where

$v_{[i]} = y_{[i]} - \hat{\phi} y_{[i]-1}$, $\qquad \bar{v}_{[.]} = \frac{1}{m}\sum_{i=1}^{n}\beta_i v_{[i]}$,

$u_{[i]} = x_{[i]} - \hat{\phi} x_{[i]-1}$, $\qquad \bar{u}_{[.]} = \frac{1}{m}\sum_{i=1}^{n}\beta_i u_{[i]}$, $\qquad m = \sum_{i=1}^{n}\beta_i$,

$G = \{\sum_{i=1}^{n}\beta_i v_{[i]}u_{[i]} - m\bar{v}_{[.]}\bar{u}_{[.]}\}/\{\sum_{i=1}^{n}\beta_i u_{[i]}^2 - m\bar{u}_{[.]}^2\}$,

$H = \sum_{i=1}^{n}\alpha_i u_{[i]}/\{\sum_{i=1}^{n}\beta_i u_{[i]}^2 - m\bar{u}_{[.]}^2\}$,

$$K = \frac{\sum_{i=1}^{n}\beta_i(y_{[i]} - \hat{\gamma}x_{[i]})(y_{[i]-1} - \hat{\gamma}x_{[i]-1}) - \hat{\mu}\sum_{i=1}^{n}\beta_i(y_{[i]-1} - \hat{\gamma}x_{[i]-1})}{\sum_{i=1}^{n}\beta_i(y_{[i]-1} - \hat{\gamma}x_{[i]-1})^2},$$

$D = \sum_{i=1}^{n}\alpha_i\,(y_{[i]-1} - \hat{\gamma}x_{[i]-1})/\sum_{i=1}^{n}\beta_i(y_{[i]-1} - \hat{\gamma}x_{[i]-1})^2$,

$B = \frac{2p}{k}\sum_{i=1}^{n}\alpha_i\left(v_{[i]} - Gu_{[i]}\right)$,

$C = \frac{2p}{k}\sum_{i=1}^{n}\beta_i\left\{\left(v_{[i]} - \bar{v}_{[.]}\right) - G(u_{[i]} - \bar{u}_{[.]})\right\}^2$.

Note that

$\sum_{i=1}^{n}\beta_i u_{[i]}^2 - m\bar{u}_{[.]}^2 = \sum_{i=1}^{n}\beta_i\,(u_{[i]} - \bar{u}_{[.]})^2$ and from symmetry $\sum_{i=1}^{n}\alpha_i = 0$.

Remark: LS estimators can also be obtained from MML estimators by taking $\alpha_i = 0$ and $\beta_i = 1$.

Comment: The coefficients β_i ($1 \leq i \leq n$) have umbrella ordering, i.e., they have an increasing sequence until the middle value and then decrease in a symmetric fashion. Thus, if $\beta_1 \geq 0$ then all the β_i coefficients are positive. Also, $\hat{\sigma}$ is real and positive provided $\beta_i \geq 0$ for all i=1,2,…,n and highly efficient for all n. For large n and small p, β_1 (and possibly few other coefficients) can be negative therefore, $\hat{\sigma}$ cannot be real. In that case, we recast the linear approximation (2.29) by replacing α_i by α_i^* and β_i and β_i^* as follows:

$$\alpha_i^* = (1/k)t_{(i)}^3/\{1 + (1/k)\, t_{(i)}^2\}^2 \text{ and } \beta_i^* = 1/\{1 + (1/k)\, t_{(i)}^2\}^2 \qquad (2.37)$$

Since $\{z_{(i)} - t_{(i)}\} \cong 0$, this correction does not change the asymptotic properties of MMLEs. Consequently, we can write $g(z_{(i)}) \cong \alpha_i^* + \beta_i^* z_{(i)}$.

Small weights are assigned to extreme observations which depletes their effect. In long tailed symmetric distributions, this is useful to achieve robustness towards outliers and other data anomalies (Tiku and Akkaya, 2004).

Properties of MML

1. MML estimators are asymptotically equivalent to the maximum likelihood estimators (Vaughan and Tiku, 2000; Bhattacharyya, 1985).
2. The estimates are almost fully efficient even for small sample sizes.
3. Estimates are explicit functions of sample observations and no need for numerical computation.
4. The estimates are unbiased or have little bias.
5. This method assigns small weights to extreme values so estimators become robust.

2.5 Adaptive Modified Maximum Likelihood Estimation

MML estimation technique assumes a specific distribution. However, in some situations like in machine data processing, the underlying distribution may not be identified. Rather, assumed that it belongs to a member of a broad class of distributions (Hampel et al., 1986). Huber (1964) and his collaborators developed M-estimators with the assumption that the underlying distribution is LTS. Tiku and Sürücü (2009) suggested a modification to MMLEs for the location and scale parameters of the LTS and denoted the resulting estimators as MML30, whereas Dönmez (2010) named them as Revised Modified Maximum Likelihood estimators. Dönmez (2010) applied the methodology to regression and experimental design to one-way and two-way classification models under both LTS and generalized logistic distributions. Following, Akkaya and Bayrak (2018) called them Adaptive Modified Maximum Likelihood (AMML) estimators in memory of Moti Lal Tiku who was the pioneer of this idea. Also, they derived AMML estimators of AR(1) model and suggested a hypothesis testing procedure for the autoregressive parameter. In this study, the AMML estimation procedure is extended to autoregression models under LTS distribution.

2.5.1 M-Estimators

Let $X_1,\ X_2,\ X_3, \ldots, X_n$ be a random sample from a LTS distribution of type $(1/\sigma)\, f((x-\mu)/\sigma)$. Huber (1964) showed that the location parameter μ can be estimated as follows:

i) First express the log likelihood function as

$$\ln L = -n \ln \sigma \sum_{i=1}^{n} \ln f(z_i)\,, \quad z_i = (x_i - \mu/\sigma). \tag{2.38}$$

ii) If the functional form of f is known, the MLE can be obtained by taking the partial derivative with respect to parameter μ as

$$\frac{\partial \ln L}{\partial \mu} = -\frac{1}{\sigma}\sum_{i=1}^{n} \psi(z_i) = 0, \qquad \psi(z) = -f'(z)/f(z). \tag{2.39}$$

iii) Equation (2.39) can be redefined as $\sum_{i=1}^{n} w_i(x_i - \mu) = 0$ by writing $w_i = w_i(z) = \psi(z_i)/z_i$. The resulting estimator of μ is obtained as

$$\mu = \sum_{i=1}^{n} w_i\, x_i \Big/ \sum_{i=1}^{n} w_i. \tag{2.40}$$

Given σ and $\psi(z)$,equation (2.40) can be solved by iteration but in practice neither σ nor $\psi(z)$ are known. Huber (1964) proposed a function $\psi(z)$ as

$$\psi(z) = \begin{cases} z, & \text{if } |z| \le c \\ c\,\mathrm{sgn}(z), & \text{if } |z| > c \end{cases} \tag{2.41}$$

This corresponds to a normal distribution in the middle and double exponential in the tails. Huber argued that the popular choices of c are 1.345, 1.5 and 2 which correspond to 10%, 5% and 2.5% censoring, respectively on both sides of a normal distribution N(0,1).

Huber (1964, 1977), Gross (1976, 1977) and many others used $\tilde{\sigma}_0$=mad=median$|x_i - \mathrm{median}(x_i)|$ as an estimate for the unknown scale parameter σ. Later on, Huber (1981) and Birch and Myers (1982) proposed to change the value of this estimate to $\tilde{\sigma}_0$=mad/0.6745 and make it an unbiased estimator of σ based on the idea that E(mad)=0.6745 for the standart normal distribution.
It can be noted that, in case of N(0,1), mad is asymptotically unbiased estimate of σ. In an extensive numerical study (Gross, 1976), sixty five $\psi(z)$ functions were examined. Three of them found particularly useful. These are descending functions which means they decrease as |z| increases. Most popular ones are:

1. The wave function (Andrews, et al., 1972, Andrews, 1974)

$$\psi(z) = \begin{cases} \sin(z) & \text{if } |z| \le \pi \\ 0 & \text{if } |z| > \pi. \end{cases} \tag{2.42}$$

2. The bisquare function (Beaton and Tukey, 1974)

$$\varepsilon(z) = \begin{cases} x\,(1 - z^2)^2 & \text{if } |z| \le 1 \\ 0 & \text{if } |z| > 1. \end{cases} \tag{2.43}$$

3. The piecewise linear function (Hampel, 1974)

$$\varepsilon(z) = \begin{cases} |z| & \text{if } 0 \le |z| < a \\ a & \text{if } a \le |z| < b \\ \frac{c-|z|}{c-b} & \text{if } b \le |z| < c \\ 0 & \text{if } \le z \end{cases} \tag{2.44}$$

Different values of a, b and c yields different estimators.

Gross (1976) examined 25 representative estimators (out of 65) of μ and σ and remarked the three of them since they have similar efficiency and robustness properties. These are wave estimators W24, the bisquare estimators BS82 and the Hampel estimators H22. In this study, we only give the expressions for the estimator W24 as follows since the values of the M-estimators BS82 and H22 turn out to be similar to W24 (Dunnet, 1982, Tiku et al., 1986).

Taking $T_0 = \text{median}\,\{x_i\}$, $S_0 = \text{median}\{|x_i - T_0|\}$ (i=1,2,…,n) and

$z_i = (x_i - T_0)/hS_o$ ($|z_i| < \pi$) we have

$$\hat{\mu}_w = T_o + (hS_o)\tan^{-1}\left[\frac{\sum_{i=1}^{n}\sin(z_i)}{\sum_{i=1}^{n}\cos(z_i)}\right] \text{ and}$$

$$\hat{\sigma}_w = (hS_o)\left[n\frac{\sum_{i=1}^{n}\sin^2(z_i)}{(\sum_{i=1}^{n}\cos(z_i))^2}\right]^{1/2}, \tag{2.45}$$

where h=2.4

Remark: After various studies on defining coefficient h, Gross (1976) finally recommended to use h=2.4 depending on the simulation results. For this value of the coefficient, the estimators possess high efficiency and robustness.

Gross (1976) explicitly defines BS82 and H22.

To conclude,

- These three estimators, W24, BS82 and H22, censor observations. But for every sample, the number of censored observations can be different.
- M-estimators of μ are unbiased and remarkably efficient only for LTS distributions. For skew and short tailed distributions this method is not ideal but still not bad at all (Balci and Akkaya, 2015).
- M-estimators may underestimate σ, even asymptotically.

2.5.2 Influence Function and Breakdown Point

The influence function and the break down point are the most popular measures used for investigating the robustness of an estimator. They are almost equivalent to each other. The notion of breakdown point is presented by Hampel (1968, 1971). It is the fraction of the unusual data that the estimator can tolerate without being affected. For example, the median can tolerate up to 50% outliers whereas for the mean it is 0%. The influence function introduced by Hampel (1974) represents the effect of a single outlier. M-estimators are robust due to their bounded influence functions.

2.5.3 Adaptive Modified Maximum Likelihood Estimators

As distinct from MML method, $t_{(i)}$, α_i and β_i are estimated from the sample in AMML technique since the shape parameter p is not known in real life. Tiku and Sürücü (2009) integrated the advantages of Huber's and MML estimation methods to obtain robust estimators which is later on named as AMML estimation procedure by Akkaya and Bayrak (2018), as mentioned before. The steps of AMML procedure can be shown as follows:

1. First, revise the coefficients α_i and β_i.

$$\hat{\alpha}_i = (2/k)\hat{t}_i/\{1 + (1/k)\hat{t}_i^2\}^2, \quad \hat{\beta}_i = 1/\{1 + (1/k)\hat{t}_i^2\}^2. \tag{2.46}$$

We make this revision since MML estimators do not have bounded influence functions.

2. Define a new parameter; $\delta = -\phi\gamma$, and write the model (2.2) as a multiple regression model:

$$y_t = \mu + \phi y_{t-1} + \gamma x_t + \delta x_{t-1} + \varepsilon_t \ (1 \le t \le n) \ \ (-1<\phi<1). \qquad (2.47)$$

3. Take all the slope parameters equal as in Dönmez (2010), say θ, by assuming the effect of the all parameters are equal initially and express the model (2.47) as

$$y_t = \mu + \theta s_t + \varepsilon_t, \qquad (2.48)$$

where

$s_t = x_t - x_{t-1} + y_{t-1} \quad (1\le t\le n)$

4. An initial estimate of t_i values is obtained as

$$\tilde{t}_i = (y_i - T_0 - T_1 s_i)/s_0 \quad (1\le i\le n) \qquad (2.49)$$

where

$T_1 = \text{med}\{r_l\}$; $r_l = (y_{l+1} - y_l)/(s_{l+1} - s_l) \ \ (1 \le l \le n-1)$, $T_0 = \text{med}\{y_i - T_1 s_i\}$ $(1\le i \le n)$, and $S_0 = 1.483\,\text{med}\{|y_i - T_1 s_i - T_0|\}, (1 \le i \le n)$ are initial estimators of θ, μ, and σ, respectively.

5. In this method we do not need to use concomitants, since the complete sums do not affect from the ordering, so we use $\tilde{t}_i$ rather than $\tilde{t}_{(i)}$.

6. Incorporate $\tilde{t}_i$ in equation (2.46), then obtain the initial estimates of α_i and β_i as $\tilde{\alpha}_i$ and $\tilde{\beta}_i$, respectively.

7. Finally, incorporate $\tilde{\alpha}_i$ and $\tilde{\beta}_i$ values in equations (2.32)-(2.35) as in the MML method. AMML estimators are the solutions of these equations.

Alternatively, we propose another method for the initial values of the parameters other than the one given in steps (3) and (4), above. In this method T_0 and S_0 are used as the initial estimators of μ and σ, respectively and given as

$T_0 = \text{med}\{y_i - T_1 y_{i-1} - T_2 x_i - T_3 x_{i-1}\}$ and

$$S_0 = 1.483\text{med}\{y_i - T_1 y_{i-1} - T_2 x_i - T_3 x_{i-1} - T_0\}, \ (1 \le i \le n) \qquad (2.50)$$

where, $T_1 = \text{med}\{r_{1l}\}$, $r_{1l} = \frac{y_{l+1} - y_l}{y_l - y_{l-1}}$;

$T_2 = \text{med}\{r_{2l}\}$, $r_{2l} = \frac{y_{l+1} - y_l}{x_{l+1} - x_l}$;

$T_3 = \text{med}\{r_{3l}\}$, $r_{3l} = \frac{y_{l+1} - y_l}{x_l - x_{l-1}}$ $(1 \le l \le n-1)$.

The t_i values in equations (2.46) can be estimated by

$$\hat{t}_i = \frac{y_t - T_0 - T_1 y_{t-1} - T_2 x_t - T_3 x_{t-1}}{S_0}. \qquad (2.51)$$

Other steps (5-7) are exactly the same as given above.

2.5.3.1 Choice of k

Realize that $\hat{\alpha}_i$ and $\hat{\beta}_i$ depend on the value k (=2p-3). As indicated by Tiku and Sürücü (2009), if large k value is used, $\hat{\beta}_i$ $(1 \le i \le n)$ reduce to 1 which means estimators loose its efficiency and robustness. On the contrary, if small k value is used, estimators lose their efficiency drastically. That’s why choosing k=30 (p=16.5) is appropriate.

CHAPTER 3

SIMULATION STUDY

A good estimator should possess three desirable properties. These are unbiasedness, efficiency and robustness to the data anomalies. If the estimators have these properties, more accurate results can be obtained from the model under consideration. In this chapter, a comprehensive simulation study is conducted to compare the performances of the estimators and the results are illustrated.

3.1 Comparison Criterion

To compare the efficiency and robustness properties of the parameters and obtaining an insightful conclusion, comparison should be based on a reliable criterion. In order to explore the performances of the methods, we used the mean squared error (MSE) as a criterion. It is the average squared difference between the predicted values and the actual observations. Thus, MSE value of an estimator, $\hat{\theta}$ is:

$$\mathrm{MSE} = \mathrm{E}\left[\left(\hat{\theta} - \theta\right)^2\right]. \tag{3.1}$$

Equivalently,

$$\mathrm{MSE} = \mathrm{Var}\left(\hat{\theta}\right) + \mathrm{Bias}(\theta)^2. \tag{3.2}$$

Note that, a smaller MSE represents a better performance due to small bias and variability.

3.2 Convergence Criterion

In numerical algorithms, process starts with an initial guess and continues with the updated parameter values until convergence is achieved. For this reason, a convergence criterion is needed. To define, it is the difference between iterative values. When the difference becomes smaller than ε (usually taken as 10^{-5}), iteration process ends. In other words, it continues until there is no significant change between the estimated parameter values for consecutive iterations.

$$|\hat{\theta}^{k+1} - \hat{\theta}^{k}| < \varepsilon \tag{3.3}$$

where $\hat{\theta}$ indicates the estimated parameter value and k is the iteration number.

In this study, for ECM algorithm the same criterion is used.

3.3 Statistical Properties of the Estimators

The robustness issue is important since no one can be sure about the exact distribution. An estimator is called fully efficient if it is unbiased and has minimum variance. Also, an estimator is called robust if it is fully efficient (or nearly so) for an assumed distribution and maintains high efficiency for plausible alternatives. The efficiency and robustness properties of the estimators are evaluated under the assumption that error term comes from a LTS distribution with p=16.5 The simulation programs are written in R language. One of the codes is given for Model 1 in Appendix as an example. All the simulations are based on [100000/n] (integer value) Monte-Carlo runs. We consider the sample sizes n=30, 50 and 100. Also, μ, γ, ϕ, σ are taken to be 0, 1, 0.5 and 1, respectively without loss of generality.

We have generated random samples from the following twelve representative models.

(1) Normal with mean 0 and variance σ^2,

and the LTS family with

(2) p = 5.0;

(3) p = 3.5;

(4) p = 2.5;

(5) p = 2.0;

The outlier models where $(n-r)X_i$ come from $N(0,\sigma^2)$ and r (we do not know which) come from

(6) $N(0,4\sigma^2)$;

(7) $N(0,16\sigma^2)$; $r=[0.5+0.1n]$ (integer value)

and the mixture models

(8) $0.90N(0,\sigma^2)+0.10N(0,4\sigma^2)$;

(9) $0.90N(0,\sigma^2)+0.10N(0,16\sigma^2)$;

As extreme alternative sample models

(10) Student's t distribution with 2 degrees of freedom;

(11) Cauchy distribution and

(12) Slash (Normal/Uniform) distribution

are taken.

Note that although Models (1)-(9) have finite mean and variance, Model (10) has finite mean but variance does not exist. Also, both mean and variance of Model (11) and Model (12) do not exist.

According to the simulations, it can be seen from Table 3.1 that in Model (1), all the results are close to each other since it is a perfect case due to the proximity to normality. As mentioned in the earlier sections, LTS distribution approaches to N(0,1) when p goes to infinity. Besides, the kurtosis is infinite when p=2.5 and p=2 yielding the simulated variances of σ to be different as compared to p=5. As the distribution deviates from normality, σ is underestimated. This is more apparent in Model (5).

In Model (6)-(9), in order to make their variances equal to σ^2, the innovations are scale corrected. The results of the simulations are given in the Table 3.2 and Table 3.3. The interpretations are similar to misspecification ones shown in Table 3.1. However, estimated values of σ deviate from the assumed value (in a downward direction) in Model (7) and (9).

Usually, ECM and MML give similar results since MML is the approximate solution and ECM is the numerical solution of ML. However, this similarity is not valid for every model. For example, in Table 3.4 results can differ. On the other hand, ECM does not converge always for Model (11) and (12). In Model (11) and (12), estimators have exploding variances except AMML. Besides, σ is overestimated with unacceptable variances. Model (10) is still comparable to others since it yields the smaller variances and almost unbiased estimators.

To conclude, as sample size increases better results are obtained as expected. LSEs have the highest MSE for all cases, generally. To repeat, it has larger bias than the other estimators. The ϕ is underestimated in all models but as sample size increases it approaches to its assumed value. AMML estimators outperform among the others. Especially, under extreme alternative models only AMML method is robust showing the best performance; while not ignoring the overestimation of σ.

Table 3.1 Monte Carlo Averages and MSEs of LS, ECM, MML, AMML Estimators under Alternative Models (1)-(5)

n			**μ**	**γ**	**ϕ**	**σ**	**μ**	**γ**	**ϕ**	**σ**
			Model (1)				**Model (2)**			
30	LS	Mean	-0.001	1.000	0.418	0.996	-0.006	0.997	0.417	0.991
		MSE	**0.052**	**0.027**	**0.034**	**0.018**	**0.052**	**0.029**	**0.035**	**0.027**
	ECM	Mean	0.000	1.000	0.424	0.949	-0.005	0.999	0.425	0.938
		MSE	**0.048**	**0.021**	**0.033**	**0.019**	**0.047**	**0.021**	**0.033**	**0.026**
	MML	Mean	0.000	1.000	0.425	0.998	-0.005	0.999	0.425	0.990
		MSE	**0.048**	**0.021**	**0.033**	**0.018**	**0.047**	**0.021**	**0.033**	**0.026**
	AMML	Mean	0.000	0.996	0.406	1.012	-0.004	0.995	0.421	0.998
		MSE	**0.048**	**0.023**	**0.034**	**0.018**	**0.048**	**0.021**	**0.034**	**0.023**
50	LS	Mean	0.003	1.001	0.448	0.999	0.007	0.993	0.451	0.994
		MSE	**0.027**	**0.016**	**0.019**	**0.011**	**0.025**	**0.016**	**0.017**	**0.016**
	ECM	Mean	0.003	1.002	0.453	0.973	0.006	0.995	0.456	0.958
		MSE	**0.026**	**0.013**	**0.018**	**0.011**	**0.023**	**0.012**	**0.016**	**0.015**
	MML	Mean	0.003	1.002	0.453	1.003	0.006	0.995	0.457	0.992
		MSE	**0.026**	**0.013**	**0.018**	**0.011**	**0.023**	**0.012**	**0.016**	**0.015**
	AMML	Mean	0.004	0.998	0.444	1.012	0.008	0.993	0.448	0.994
		MSE	**0.026**	**0.013**	**0.019**	**0.011**	**0.023**	**0.012**	**0.017**	**0.013**
100	LS	Mean	0.001	1.006	0.480	1.002	-0.001	1.002	0.477	0.996
		MSE	**0.011**	**0.008**	**0.007**	**0.005**	**0.012**	**0.007**	**0.008**	**0.007**
	ECM	Mean	0.000	1.005	0.482	0.990	-0.002	1.001	0.479	0.974
		MSE	**0.010**	**0.006**	**0.007**	**0.005**	**0.011**	**0.006**	**0.008**	**0.007**
	MML	Mean	0.000	1.005	0.482	1.006	-0.001	0.999	0.479	0.992
		MSE	**0.010**	**0.006**	**0.007**	**0.005**	**0.011**	**0.006**	**0.008**	**0.007**
	AMML	Mean	0.001	1.005	0.479	1.014	-0.000	0.999	0.476	0.994
		MSE	**0.010**	**0.006**	**0.007**	**0.005**	**0.011**	**0.006**	**0.008**	**0.006**

Table 3.1 (continued)

n			μ	γ	ϕ	σ	μ	γ	ϕ	σ
			Model (3)				**Model (4)**			
30	LS	Mean	-0.006	0.998	0.419	0.986	-0.001	0.999	0.423	0.974
		MSE	**0.051**	**0.028**	**0.034**	**0.036**	**0.052**	**0.028**	**0.033**	**0.070**
	ECM	Mean	-0.005	0.999	0.426	0.927	-0.000	0.998	0.432	0.904
		MSE	**0.045**	**0.020**	**0.032**	**0.033**	**0.043**	**0.019**	**0.030**	**0.052**
	MML	Mean	-0.005	0.999	0.426	0.981	-0.000	0.998	0.432	0.965
		MSE	**0.046**	**0.021**	**0.032**	**0.034**	**0.045**	**0.020**	**0.030**	**0.062**
	AMML	Mean	-0.008	0.998	0.412	0.984	-0.001	0.995	0.414	0.953
		MSE	**0.044**	**0.020**	**0.034**	**0.028**	**0.043**	**0.020**	**0.033**	**0.038**
50	LS	Mean	0.005	0.999	0.452	0.990	0.000	1.001	0.449	0.987
		MSE	**0.027**	**0.016**	**0.018**	**0.020**	**0.027**	**0.016**	**0.018**	**0.050**
	ECM	Mean	0.005	1.002	0.458	0.947	-0.001	0.999	0.457	0.923
		MSE	**0.024**	**0.012**	**0.017**	**0.018**	**0.021**	**0.010**	**0.015**	**0.030**
	MML	Mean	0.005	1.002	0.458	0.984	-0.001	0.998	0.456	0.974
		MSE	**0.025**	**0.012**	**0.017**	**0.019**	**0.023**	**0.011**	**0.016**	**0.043**
	AMML	Mean	0.006	0.997	0.450	0.981	-0.001	0.997	0.449	0.949
		MSE	**0.025**	**0.012**	**0.017**	**0.015**	**0.023**	**0.010**	**0.016**	**0.023**
100	LS	Mean	-0.008	1.002	0.481	0.999	-0.000	0.996	0.473	0.986
		MSE	**0.011**	**0.007**	**0.008**	**0.011**	**0.011**	**0.007**	**0.007**	**0.028**
	ECM	Mean	-0.007	1.001	0.484	0.965	-0.004	0.997	0.477	0.928
		MSE	**0.009**	**0.005**	**0.007**	**0.009**	**0.009**	**0.005**	**0.007**	**0.015**
	MML	Mean	-0.007	1.001	0.484	0.989	-0.004	0.997	0.477	0.966
		MSE	**0.009**	**0.005**	**0.007**	**0.010**	**0.010**	**0.005**	**0.007**	**0.022**
	AMML	Mean	-0.005	1.002	0.481	0.982	0.000	0.996	0.475	0.940
		MSE	**0.009**	**0.005**	**0.007**	**0.007**	**0.010**	**0.005**	**0.007**	**0.012**

Table 3.1 (continued)

n			μ	γ	ϕ	σ
		Model (5)				
30	LS	Mean	0.006	0.996	0.416	0.939
		MSE	**0.048**	**0.028**	**0.034**	**0.178**
	ECM	Mean	0.006	0.998	0.426	0.849
		MSE	**0.032**	**0.016**	**0.029**	**0.103**
	MML	Mean	0.006	0.998	0.425	0.924
		MSE	**0.037**	**0.018**	**0.030**	**0.156**
	AMML	Mean	0.007	0.993	0.408	0.880
		MSE	**0.036**	**0.017**	**0.028**	**0.056**
50	LS	Mean	-0.000	0.999	0.455	0.941
		MSE	**0.024**	**0.015**	**0.017**	**0.085**
	ECM	Mean	0.000	1.001	0.463	0.855
		MSE	**0.017**	**0.008**	**0.014**	**0.050**
	MML	Mean	0.000	1.001	0.462	0.920
		MSE	**0.020**	**0.009**	**0.015**	**0.072**
	AMML	Mean	0.001	0.999	0.456	0.875
		MSE	**0.017**	**0.008**	**0.015**	**0.038**
100	LS	Mean	0.002	1.007	0.477	0.958
		MSE	**0.011**	**0.007**	**0.007**	**0.074**
	ECM	Mean	0.002	1.003	0.481	0.861
		MSE	**0.007**	**0.004**	**0.006**	**0.034**
	MML	Mean	0.002	1.003	0.481	0.925
		MSE	**0.008**	**0.004**	**0.006**	**0.059**
	AMML	Mean	0.002	1.002	0.479	0.867
		MSE	**0.007**	**0.003**	**0.006**	**0.028**

Table 3.2 Monte Carlo Averages and MSEs of LS, ECM, MML, AMML Estimators under Alternative Models (6)-(7)

n			μ	γ	ϕ	σ	μ	γ	ϕ	σ
			Model (6)				**Model (7)**			
30	LS	Mean	-0.001	0.954	0.408	0.995	0.002	1.000	0.396	0.880
		MSE	**0.048**	**0.025**	**0.045**	**0.025**	**0.037**	**0.023**	**0.064**	**0.063**
	ECM	Mean	-0.001	0.902	0.414	0.997	0.002	1.001	0.404	0.795
		MSE	**0.044**	**0.028**	**0.043**	**0.019**	**0.028**	**0.014**	**0.058**	**0.071**
	MML	Mean	-0.000	0.998	0.414	0.953	0.002	1.001	0.403	0.864
		MSE	**0.044**	**0.020**	**0.043**	**0.024**	**0.031**	**0.015**	**0.060**	**0.061**
	AMML	Mean	-0.001	0.993	0.395	0.962	0.001	0.997	0.384	0.844
		MSE	**0.045**	**0.021**	**0.043**	**0.022**	**0.028**	**0.015**	**0.061**	**0.055**
50	LS	Mean	0.000	1.007	0.448	0.972	-0.004	0.994	0.432	0.922
		MSE	**0.025**	**0.016**	**0.023**	**0.014**	**0.021**	**0.015**	**0.043**	**0.040**
	ECM	Mean	0.001	1.005	0.453	0.936	-0.003	0.996	0.439	0.837
		MSE	**0.023**	**0.011**	**0.022**	**0.015**	**0.015**	**0.008**	**0.037**	**0.046**
	MML	Mean	0.001	1.005	0.453	0.969	-0.003	0.996	0.439	0.899
		MSE	**0.023**	**0.011**	**0.022**	**0.013**	**0.017**	**0.009**	**0.038**	**0.038**
	AMML	Mean	0.002	1.003	0.444	0.972	-0.004	0.994	0.427	0.860
		MSE	**0.024**	**0.012**	**0.022**	**0.012**	**0.015**	**0.008**	**0.037**	**0.037**
100	LS	Mean	0.007	0.998	0.470	0.985	0.002	1.000	0.452	0.954
		MSE	**0.010**	**0.008**	**0.011**	**0.007**	**0.010**	**0.007**	**0.018**	**0.021**
	ECM	Mean	0.007	0.999	0.474	0.962	0.001	1.000	0.459	0.867
		MSE	**0.010**	**0.006**	**0.010**	**0.007**	**0.008**	**0.004**	**0.017**	**0.027**
	MML	Mean	0.007	0.999	0.474	0.981	0.001	1.001	0.458	0.921
		MSE	**0.010**	**0.006**	**0.010**	**0.007**	**0.008**	**0.004**	**0.017**	**0.021**
	AMML	Mean	0.008	0.998	0.471	0.982	0.001	0.998	0.432	0.871
		MSE	**0.009**	**0.006**	**0.010**	**0.006**	**0.011**	**0.004**	**0.016**	**0.020**

Table 3.3 Monte Carlo Averages and MSEs of LS, ECM, MML, AMML Estimators under Alternative Models (8)-(9)

n			**μ**	**γ**	**ϕ**	**σ**	**μ**	**γ**	**ϕ**	**σ**
			Model (8)				**Model (9)**			
30	LS	Mean	-0.001	0.998	0.419	0.997	0.003	1.003	0.422	0.970
		MSE	**0.051**	**0.026**	**0.034**	**0.029**	**0.049**	**0.030**	**0.031**	**0.082**
	ECM	Mean	-0.001	1.002	0.427	0.942	0.003	1.004	0.434	0.878
		MSE	**0.046**	**0.019**	**0.032**	**0.027**	**0.037**	**0.018**	**0.026**	**0.069**
	MML	Mean	-0.001	1.002	0.428	0.995	0.004	1.004	0.433	0.953
		MSE	**0.047**	**0.019**	**0.032**	**0.027**	**0.041**	**0.019**	**0.027**	**0.075**
	AMML	Mean	-0.004	0.997	0.410	1.001	0.003	0.999	0.417	0.920
		MSE	**0.048**	**0.018**	**0.032**	**0.024**	**0.039**	**0.018**	**0.029**	**0.057**
50	LS	Mean	-0.003	1.002	0.455	0.997	-0.004	1.000	0.449	0.979
		MSE	**0.024**	**0.016**	**0.017**	**0.016**	**0.025**	**0.016**	**0.017**	**0.054**
	ECM	Mean	-0.002	1.003	0.460	0.959	-0.003	0.999	0.459	0.888
		MSE	**0.022**	**0.012**	**0.016**	**0.015**	**0.017**	**0.009**	**0.014**	**0.046**
	MML	Mean	-0.002	1.003	0.460	0.994	-0.003	0.999	0.457	0.954
		MSE	**0.023**	**0.012**	**0.016**	**0.015**	**0.020**	**0.010**	**0.014**	**0.048**
	AMML	Mean	-0.003	1.001	0.453	0.994	-0.003	0.997	0.453	0.908
		MSE	**0.022**	**0.012**	**0.015**	**0.013**	**0.017**	**0.008**	**0.014**	**0.039**
100	LS	Mean	-0.004	1.001	0.472	0.991	0.000	1.003	0.477	0.982
		MSE	**0.011**	**0.007**	**0.008**	**0.009**	**0.011**	**0.007**	**0.007**	**0.029**
	ECM	Mean	-0.004	1.000	0.475	0.967	-0.000	1.000	0.481	0.892
		MSE	**0.010**	**0.006**	**0.007**	**0.008**	**0.008**	**0.004**	**0.005**	**0.030**
	MML	Mean	-0.004	1.000	0.475	0.986	-0.000	1.000	0.481	0.947
		MSE	**0.010**	**0.006**	**0.007**	**0.008**	**0.009**	**0.004**	**0.006**	**0.027**
	AMML	Mean	-0.002	0.999	0.474	0.987	-0.000	0.999	0.479	0.898
		MSE	**0.010**	**0.006**	**0.007**	**0.007**	**0.008**	**0.004**	**0.005**	**0.026**

Table 3.4 Monte Carlo Averages and MSEs of LS, ECM, MML, AMML Estimators under Alternative Models (10)-(12)

n			μ	γ	ϕ	σ	μ	γ	ϕ	σ
			Model (10)				**Model (11)**			
30	LS	Mean	0.006	1.000	0.420	2.532	1.593	3.814	0.425	45.71
		MSE	**0.693**	**0.276**	**0.032**	**11.91**	**3659**	**6069**	**0.045**	**2364**
	ECM	Mean	0.000	0.995	0.435	2.120	1.933	0.789	0.451	49.36
		MSE	**0.215**	**0.101**	**0.027**	**4.095**	**1743**	**1100**	**0.022**	**3294**
	MML	Mean	0.003	0.998	0.432	2.452	6.974	1.992	0.445	80.33
		MSE	**0.426**	**0.141**	**0.028**	**10.29**	**1978**	**1686**	**0.025**	**1003**
	AMML	Mean	-0.004	0.988	0.438	1.781	-0.004	1.011	0.482	3.014
		MSE	**0.146**	**0.061**	**0.024**	**0.803**	**0.235**	**0.150**	**0.008**	**4.971**
50	LS	Mean	-0.005	1.014	0.452	2.566	1.451	2.248	0.459	45.71
		MSE	**0.271**	**0.188**	**0.017**	**6.584**	**3166**	**3382**	**0.013**	**2364**
	ECM	Mean	-0.007	1.003	0.462	2.063	-0.150	1.328	0.478	19.49
		MSE	**0.093**	**0.047**	**0.013**	**1.880**	**62.14**	**77.98**	**0.007**	**3525**
	MML	Mean	-0.004	1.006	0.460	2.448	1.058	1.228	0.470	41.10
		MSE	**0.170**	**0.092**	**0.014**	**5.389**	**1651**	**1635**	**0.009**	**1905**
	AMML	Mean	-0.007	0.997	0.465	1.745	0.002	1.001	0.494	2.949
		MSE	**0.074**	**0.032**	**0.011**	**0.660**	**0.124**	**0.081**	**0.003**	**4.334**
100	LS	Mean	0.012	1.001	0.476	2.733	0.711	1.406	0.474	52.52
		MSE	**0.125**	**0.086**	**0.007**	**6.996**	**2166**	**736.6**	**0.005**	**2133**
	ECM	Mean	0.007	1.002	0.484	2.024	-0.035	1.020	0.488	13.18
		MSE	**0.037**	**0.020**	**0.005**	**1.274**	**0.582**	**0.314**	**0.002**	**5358**
	MML	Mean	0.009	0.998	0.482	2.534	0.402	2.327	0.480	44.48
		MSE	**0.071**	**0.038**	**0.005**	**5.107**	**718.6**	**1185**	**0.004**	**1484**
	AMML	Mean	0.004	1.000	0.485	1.725	-0.004	1.003	0.499	2.868
		MSE	**0.032**	**0.016**	**0.004**	**0.571**	**0.049**	**0.034**	**0.001**	**3.753**

Table 3.4 (continued)

n			μ	γ	ϕ	σ
		Model (12)				
30	LS	Mean	-2.838	-0.584	0.438	51.33
		MSE	**1430**	**350.1**	**0.039**	**5153**
	ECM	Mean	-0.606	0.334	0.451	20.41
		MSE	**108.2**	**64.88**	**0.020**	**9103**
	MML	Mean	-1.880	-0.195	0.444	61.60
		MSE	**6933**	**197.2**	**0.025**	**4099**
	AMML	Mean	-0.030	1.078	0.489	3.572
		MSE	**0.314**	**0.180**	**0.007**	**7.606**
50	LS	Mean	0.412	2.924	0.464	68.38
		MSE	**1190**	**6587**	**0.014**	**5109**
	ECM	Mean	0.161	1.179	0.482	28.09
		MSE	**89.54**	**113.1**	**0.006**	**5950**
	MML	Mean	0.373	2.891	0.475	43.71
		MSE	**5303**	**3339**	**0.008**	**3616**
	AMML	Mean	0.013	1.002	0.501	3.604
		MSE	**0.183**	**0.126**	**0.003**	**7.391**
100	LS	Mean	-0.400	0.482	0.478	32.09
		MSE	**570.1**	**212.5**	**0.005**	**2915**
	ECM	Mean	-0.001	0.973	0.491	14.67
		MSE	**0.891**	**0.394**	**0.002**	**1496**
	MML	Mean	-0.240	0.511	0.482	29.53
		MSE	**182.0**	**129.4**	**0.004**	**2531**
	AMML	Mean	-0.009	0.999	0.499	3.568
		MSE	**0.092**	**0.061**	**0.001**	**6.903**

The simulations are repeated by using an alternative method mentioned in 2.5.3 for estimating the initial estimates of the parameters in AMML method. The results are given for some selected models in Table 3.5. As it is seen from Table 3.5 the results of AMML estimates are similar to the ones where initial estimates of the parameters are calculated from data as in Dönmez (2010). So, it can be concluded that the method of selecting initials do not affect the final AMML estimates significantly.

Table 3.5 Monte Carlo Averages and MSEs of AMML Estimators under Selected Alternative Models

n		**μ**	**γ**	**ϕ**	**σ**
Model (1)					
50	Mean	-0.002	0.998	0.442	1.002
	MSE	**0.023**	**0.014**	**0.020**	**0.010**
100	Mean	-0.000	1.000	0.473	1.002
	MSE	**0.011**	**0.006**	**0.009**	**0.005**
Model (5)					
50	Mean	-0.003	1.001	0.451	0.884
	MSE	**0.019**	**0.009**	**0.016**	**0.041**
100	Mean	-0.003	1.000	0.476	0.878
	MSE	**0.008**	**0.004**	**0.007**	**0.027**
Model (7)					
50	Mean	0.003	0.994	0.404	0.847
	MSE	**0.017**	**0.008**	**0.041**	**0.041**
100	Mean	-0.002	0.995	0.447	0.871
	MSE	**0.007**	**0.004**	**0.019**	**0.023**
Model (8)					
50	Mean	0.004	1.002	0.441	1.009
	MSE	**0.025**	**0.013**	**0.017**	**0.015**
100	Mean	0.002	1.000	0.470	1.012
	MSE	**0.009**	**0.005**	**0.008**	**0.007**
Model (10)					
50	Mean	-0.002	1.000	0.457	1.916
	MSE	**0.135**	**0.042**	**0.012**	**1.025**
100	Mean	0.009	1.002	0.477	1.905
	MSE	**0.052**	**0.018**	**0.005**	**0.094**

CHAPTER 4

HYPOTHESIS TESTING

So far, it is mentioned about parameter estimation and the estimators' comparisons. However, hypothesis testing procedure is also important. In this chapter hypothesis testing is conducted for two crucial parameters, ϕ and γ of autoregressive model by the proposed AMML estimators. Then, the power of the proposed test statistic is compared with the corresponding ones obtained by the other estimation techniques.

4.1 Significance Test of the Model

When $\phi = 0$, Model (2.2) is actually a simple linear regression model. Therefore, it is important to test the null hypothesis $H_0: \phi = 0$. On the other hand, testing $H_0: \gamma = 0$ is also very crutial since model (2.2) reduces to AR (1) model.

Since the likelihood and the modified likelihood equations are asymptotically equivalent (Tiku et al., 2000; Tiku and Akkaya, 2004):

Lemma 1: The conditional distribution of $\hat{\phi}(\mu, \gamma, \sigma)$ is asymptotically normal with mean ϕ and variance $\frac{k\sigma^2}{2p\left[\sum_{i=1}^{n} \beta_i \left(y_{[i]-1} - \gamma x_{[i]-1}\right)^2\right]}$.

Proof: The above result occurs since $\partial \ln L^*/\partial\phi$ assumes the (Kendall and Stuart, 1979; Tiku et al., 1999).

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

Following Lemma 1, we propose the test statistic T_1 to test $H_0: \phi = 0$:

$$T_1 = \sqrt{(1.1)\sum_{i=1}^{n}\hat{\beta}_i(y_{i-1}-\hat{\gamma}x_{i-1})^2}\left(\frac{\hat{\phi}}{\hat{\sigma}}\right); \tag{4.1}$$

where $\hat{\gamma}$, $\hat{\phi}$ and $\hat{\sigma}$ are the AMML estimators, $\hat{\beta}_i$ is the corresponding coefficient under AMML procedure given in equation (2.46) and $2p/k$=1.1 for p=16.5.

Since $\hat{\gamma}$, $\hat{\phi}$ and $\hat{\sigma}$ converges to γ, ϕ and σ respectively, the asymptotic distribution of T_1 is N(0,1).

$H_0: \phi = 0$ is rejected in favor of $H_1: \phi > 0$ for large values of T_1.

T_2 is obtained similarly by the use of the MML estimators. Note that, (y_{i-1}, x_{i-1}) should be replaced with their concomitants $(y_{[i]-1}, x_{[i]-1})$ and $\hat{\beta}_i$ replaced with β_i given in equation (2.30).

The corresponding test statistic with the use of LS estimators is:

$$T_3 = \sqrt{\sum_{i=1}^{n}(y_{i-1}-\tilde{\gamma}x_{i-1})}\left(\frac{\tilde{\phi}}{\tilde{\sigma}}\right). \tag{4.2}$$

Here $\tilde{\gamma}$, $\tilde{\phi}$ and $\tilde{\sigma}$ are the LS estimators of the parameters γ, ϕ and σ, respectively.

Lemma 2: For a given value of ϕ (σ is known), the conditional distribution of $\hat{\gamma}(\phi, \sigma)$ is asymptotically normal with mean γ and variance $\frac{k\sigma^2}{2p\left[\sum_{i=1}^{n}\beta_i\left(u_{[i]}-\bar{u}_{[.]}\right)^2\right]}$,

where $u_{[i]} = x_{[i]} - \phi x_{[i]-1}$, $\bar{u}_{[.]} = \frac{1}{m}\sum_{i=1}^{n}\beta_i u_{[i]}$.

Proof: The result follows from the asymptotic equivalence of $\partial \ln L^*/\partial\gamma$ and $\partial \ln L/\partial\gamma$ (Tiku and Akkaya, 2004; Vaughan and Tiku, 2000). Thus $\partial \ln L^*/\partial\gamma$ assumes the form (Kendall and Stuart, 1979)

$$\frac{\partial \ln L^*}{\partial\gamma} = \frac{2p}{k\sigma^2}\left[\sum_{i=1}^{n}\beta_i\left(u_{[i]}-\bar{u}_{[.]}\right)^2\right](\hat{\gamma}-\gamma).$$

Based on Lemma 2, we first propose a test statistic T_{11} for testing $H_0: \gamma = 0$ under AMML estimation procedure:

$$T_{11} = \sqrt{(1.1)\sum_{i=1}^{n} \hat{\beta}_i \left(u_{[i]} - \bar{u}_{[.]}\right)^2} \left(\frac{\hat{\gamma}}{\hat{\sigma}}\right) \tag{4.3}$$

where 2p/k=1.1 (p=16.5), $\hat{\gamma}$ and $\hat{\sigma}$ are the AMML estimators of γ and σ respectively. $u_i = x_i - \hat{\phi} x_{i-1}$ and $\bar{u} = \sum_{i=1}^{n} \hat{\beta}_i u_i / m$, $\hat{\beta}_i$ is the corresponding coefficient given in equation (2.46). Similar to the previous cases, as n becomes large, $\hat{\phi}$ and $\hat{\sigma}$ converges to ϕ and σ, respectively and the asymptotic null distribution of T_{11} becomes N(0,1).

$H_0: \gamma = 0$ is rejected in favor of $H_1: \gamma > 0$ for large values of T_{11}.

On MML procedure the corresponding test statistic is defined as T_{22}. In this case, ϕ, γ and σ in equation (4.3) are replaced by their MML estimators, (x_i, x_{i-1}) is replaced by their concomitants $(x_{[i]}, x_{[i]-1})$ and $\hat{\beta}_i$ is replaced by β_i given in equation (2.30).

The corresponding test statistic with the use of LS estimators is:

$$T_{33} = \sqrt{\sum_{i=1}^{n} (u_i - \bar{u})^2} \left(\frac{\tilde{\gamma}}{\tilde{\sigma}}\right). \tag{4.4}$$

where $\tilde{\gamma}$, $\tilde{\sigma}$ and $\tilde{\phi}$ are the LS estimators of γ, σ and ϕ, respectively. $u_i = x_i - \tilde{\phi} x_{i-1}$ and $\bar{u} = \sum_{i=1}^{n} u_i / n$.

A test statistic is said to have both criterion and efficiency robustness when its Type I error is not substantially higher than a presumed value and its power is high, at any rate for plausible alternatives (Tiku et al., 1986). For this purpose, the robustness properties of the test statistics based on AMML, MML and LS estimators are investigated with a simulation study based on [100000/n] (integer value) Monte Carlo runs. The simulation programs are written in R language. The error term is having a distribution LTS with p=16.5. Models (1)-(12) in Chapter 3 is used as alternative sample models. The simulated values of Type I errors and the power values of test statistics for n=30, 50 and 100 are given in Tables 4.1- 4.8 for ϕ and γ, respectively. It is seen from Tables 4.1 to Table 4.8 that for the Model (1), the power values of the test statistics based on AMML, MML and LS are similar to each other. As the innovations deviate from normality, the performances of AMML and MML almost close to each other but AMML converges to 1.0 faster as sample size increases especially under extreme alternative sample Models (10)-(12).

Table 4.1 Simulated Power Values of the Test Statistics for Testing ϕ under Alternative Models (1)-(5)

	n=30			**n=50**			**n=100**		
ϕ	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$
Model (1)									
0.00	0.032	0.030	0.030	0.047	0.043	0.040	0.056	0.047	0.047
0.05	0.069	0.066	0.067	0.081	0.077	0.075	0.129	0.109	0.110
0.10	0.111	0.108	0.102	0.150	0.144	0.142	0.256	0.237	0.239
0.20	0.223	0.221	0.210	0.358	0.360	0.353	0.643	0.608	0.601
0.30	0.408	0.427	0.413	0.639	0.632	0.622	0.901	0.890	0.888
0.40	0.639	0.613	0.628	0.849	0.847	0.840	0.990	0.989	0.988
0.50	0.775	0.796	0.790	0.955	0.951	0.951	0.998	0.998	0.998
Model (2)									
0.00	0.036	0.038	0.037	0.047	0.042	0.042	0.058	0.047	0.045
0.05	0.069	0.069	0.066	0.092	0.088	0.085	0.142	0.113	0.119
0.10	0.102	0.114	0.114	0.142	0.134	0.134	0.251	0.219	0.206
0.20	0.221	0.237	0.231	0.371	0.367	0.356	0.631	0.594	0.588
0.30	0.422	0.411	0.398	0.653	0.649	0.643	0.923	0.899	0.897
0.40	0.648	0.654	0.638	0.872	0.867	0.851	0.995	0.992	0.991
0.50	0.794	0.803	0.798	0.956	0.947	0.944	1.000	1.000	1.000
Model (3)									
0.00	0.034	0.034	0.035	0.039	0.036	0.038	0.029	0.024	0.026
0.05	0.063	0.063	0.063	0.077	0.073	0.078	0.121	0.107	0.108
0.10	0.106	0.114	0.111	0.132	0.126	0.127	0.261	0.233	0.234
0.20	0.240	0.253	0.243	0.365	0.353	0.335	0.651	0.613	0.594
0.30	0.411	0.425	0.420	0.659	0.636	0.620	0.932	0.899	0.889
0.40	0.623	0.622	0.623	0.874	0.858	0.844	0.994	0.989	0.984
0.50	0.812	0.823	0.821	0.968	0.964	0.961	0.999	0.999	0.999

Table 4.1 (continued)

	n=30			n=50			n=100		
ϕ	**T_1**	**T_2**	**T_3**	**T_1**	**T_2**	**T_3**	**T_1**	**T_2**	**T_3**
Model (4)									
0.00	0.034	0.036	0.036	0.039	0.039	0.039	0.055	0.051	0.056
0.05	0.059	0.065	0.065	0.085	0.081	0.081	0.124	0.109	0.108
0.10	0.089	0.089	0.088	0.138	0.132	0.129	0.262	0.235	0.219
0.20	0.230	0.239	0.234	0.396	0.366	0.346	0.678	0.634	0.596
0.30	0.438	0.459	0.432	0.676	0.651	0.622	0.934	0.896	0.879
0.40	0.652	0.658	0.636	0.885	0.864	0.848	0.991	0.989	0.984
0.50	0.821	0.830	0.810	0.974	0.964	0.960	0.997	0.998	0.997
Model (5)									
0.00	0.035	0.029	0.031	0.042	0.030	0.030	0.040	0.033	0.039
0.05	0.060	0.055	0.056	0.074	0.067	0.069	0.136	0.104	0.102
0.10	0.099	0.090	0.091	0.146	0.119	0.120	0.296	0.233	0.228
0.20	0.239	0.234	0.224	0.408	0.356	0.331	0.734	0.661	0.623
0.30	0.445	0.436	0.411	0.728	0.670	0.637	0.957	0.924	0.904
0.40	0.688	0.670	0.639	0.919	0.887	0.866	0.996	0.991	0.988
0.50	0.855	0.847	0.828	0.975	0.966	0.959	1.000	0.999	0.998

Table 4.2 Simulated Power Values of the Test Statistics for Testing ϕ under Alternative Models (6)-(7)

	n=30			**n=50**			**n=100**		
ϕ	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$
Model (6)									
0.00	0.034	0.032	0.032	0.050	0.050	0.051	0.063	0.051	0.054
0.05	0.077	0.085	0.083	0.111	0.105	0.107	0.148	0.136	0.137
0.10	0.126	0.130	0.136	0.179	0.175	0.175	0.313	0.291	0.298
0.20	0.246	0.255	0.252	0.363	0.345	0.342	0.621	0.584	0.580
0.30	0.427	0.435	0.426	0.625	0.613	0.606	0.870	0.833	0.827
0.40	0.614	0.627	0.614	0.833	0.816	0.805	0.987	0.979	0.968
0.50	0.766	0.784	0.772	0.943	0.936	0.931	0.998	0.996	0.997
Model (7)									
0.00	0.011	0.011	0.012	0.034	0.034	0.035	0.068	0.066	0.062
0.05	0.150	0.143	0.145	0.182	0.174	0.177	0.244	0.225	0.233
0.10	0.183	0.184	0.188	0.258	0.248	0.252	0.359	0.327	0.339
0.20	0.306	0.306	0.300	0.418	0.389	0.385	0.606	0.546	0.542
0.30	0.476	0.468	0.454	0.627	0.584	0.560	0.824	0.762	0.729
0.40	0.624	0.602	0.578	0.784	0.747	0.720	0.938	0.898	0.877
0.50	0.747	0.731	0.709	0.898	0.871	0.853	0.988	0.975	0.965

Table 4.3 Simulated Power Values of the Test Statistics for Testing ϕ under Alternative Models (8)-(9)

	n=30			**n=50**			**n=100**		
ϕ	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$
Model (8)									
0.00	0.034	0.033	0.031	0.042	0.039	0.039	0.060	0.052	0.053
0.05	0.068	0.068	0.069	0.080	0.077	0.075	0.113	0.096	0.097
0.10	0.102	0.110	0.108	0.141	0.136	0.131	0.265	0.228	0.216
0.20	0.235	0.235	0.231	0.365	0.357	0.346	0.648	0.616	0.608
0.30	0.426	0.428	0.415	0.659	0.651	0.636	0.923	0.902	0.892
0.40	0.638	0.640	0.632	0.852	0.848	0.835	0.993	0.988	0.986
0.50	0.799	0.817	0.804	0.965	0.958	0.951	1.000	0.999	0.999
Model (9)									
0.00	0.028	0.026	0.026	0.040	0.035	0.037	0.042	0.038	0.043
0.05	0.055	0.057	0.055	0.078	0.064	0.065	0.123	0.096	0.098
0.10	0.096	0.097	0.095	0.141	0.121	0.122	0.291	0.221	0.208
0.20	0.211	0.218	0.201	0.425	0.393	0.372	0.721	0.624	0.582
0.30	0.446	0.442	0.413	0.755	0.712	0.662	0.962	0.929	0.906
0.40	0.703	0.693	0.656	0.925	0.889	0.864	0.999	0.997	0.992
0.50	0.855	0.855	0.827	0.981	0.974	0.962	1.000	1.000	0.990

Table 4.4 Simulated Power Values of the Test Statistics for Testing ϕ under Alternative Models (10)-(12)

	n=30			n=50			n=100		
ϕ	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
Model (10)									
0.00	0.018	0.031	0.032	0.033	0.027	0.014	0.025	0.024	0.020
0.05	0.053	0.050	0.053	0.051	0.048	0.023	0.069	0.064	0.061
0.10	0.084	0.083	0.083	0.126	0.125	0.122	0.191	0.186	0.176
0.20	0.204	0.214	0.198	0.352	0.322	0.312	0.705	0.622	0.626
0.30	0.437	0.369	0.392	0.705	0.669	0.653	0.952	0.947	0.921
0.40	0.642	0.690	0.648	0.882	0.885	0.861	0.997	0.995	0.993
0.50	0.840	0.851	0.827	0.971	0.966	0.957	1.000	0.999	0.997
Model (11)									
0.00	0.013	0.014	0.011	0.027	0.022	0.015	0.017	0.016	0.016
0.05	0.044	0.036	0.022	0.037	0.037	0.039	0.046	0.038	0.019
0.10	0.053	0.056	0.060	0.062	0.059	0.041	0.071	0.094	0.088
0.20	0.166	0.145	0.141	0.342	0.309	0.243	0.832	0.817	0.743
0.30	0.465	0.461	0.356	0.864	0.807	0.723	0.995	0.969	0.957
0.40	0.834	0.782	0.710	0.977	0.946	0.927	1.000	0.994	0.989
0.50	0.941	0.905	0.883	0.995	0.976	0.974	0.999	0.993	0.992
Model (12)									
0.00	0.024	0.021	0.018	0.016	0.016	0.011	0.019	0.018	0.016
0.05	0.033	0.034	0.027	0.038	0.035	0.030	0.041	0.041	0.026
0.10	0.058	0.055	0.039	0.068	0.061	0.060	0.112	0.092	0.084
0.20	0.153	0.147	0.137	0.306	0.305	0.224	0.784	0.764	0.700
0.30	0.482	0.455	0.351	0.860	0.831	0.759	0.994	0.977	0.968
0.40	0.840	0.787	0.715	0.972	0.932	0.914	1.000	0.992	0.992
0.50	0.942	0.899	0.881	0.995	0.979	0.973	1.000	0.997	0.997

Table 4.5 Simulated Power Values of the Test Statistics for Testing γ under Alternative Models (1)-(5)

	n=30			n=50			n=100		
γ	T_{11}	T_{22}	T_{33}	T_{11}	T_{22}	T_{33}	T_{11}	T_{22}	T_{33}
Model (1)									
0.00	0.060	0.061	0.078	0.053	0.050	0.055	0.051	0.049	0.055
0.05	0.104	0.104	0.130	0.131	0.127	0.123	0.189	0.181	0.161
0.10	0.183	0.189	0.184	0.234	0.226	0.216	0.38	0.374	0.318
0.20	0.397	0.395	0.364	0.573	0.572	0.513	0.821	0.825	0.762
0.30	0.664	0.662	0.614	0.844	0.839	0.777	0.984	0.984	0.956
0.40	0.853	0.852	0.805	0.964	0.966	0.935	0.998	0.998	0.997
0.50	0.948	0.951	0.920	0.995	0.997	0.991	1.000	1.000	1.000
Model (2)									
0.00	0.064	0.060	0.068	0.046	0.045	0.050	0.048	0.043	0.047
0.05	0.117	0.107	0.121	0.126	0.119	0.117	0.155	0.149	0.158
0.10	0.187	0.185	0.189	0.241	0.233	0.225	0.398	0.376	0.323
0.20	0.407	0.406	0.378	0.571	0.555	0.506	0.836	0.822	0.753
0.30	0.666	0.661	0.599	0.848	0.839	0.781	0.989	0.988	0.967
0.40	0.863	0.862	0.802	0.972	0.970	0.946	0.998	0.999	0.998
0.50	0.951	0.950	0.920	0.996	0.996	0.989	1.000	1.000	1.000
Model (3)									
0.00	0.054	0.051	0.062	0.059	0.058	0.064	0.044	0.041	0.063
0.05	0.111	0.109	0.117	0.126	0.118	0.119	0.178	0.172	0.158
0.10	0.188	0.182	0.186	0.240	0.230	0.208	0.387	0.377	0.324
0.20	0.407	0.400	0.373	0.590	0.568	0.515	0.846	0.819	0.739
0.30	0.693	0.687	0.620	0.857	0.841	0.776	0.985	0.979	0.959
0.40	0.863	0.859	0.811	0.967	0.963	0.929	1.000	1.000	0.997
0.50	0.956	0.942	0.911	0.996	0.996	0.983	1.000	1.000	1.000

Table 4.5 (continued)

	n=30			**n=50**			**n=100**		
γ	**T_{11}**	**T_{22}**	**T_{33}**	**T_{11}**	**T_{22}**	**T_{33}**	**T_{11}**	**T_{22}**	**T_{33}**
Model (4)									
0.00	0.061	0.058	0.070	0.054	0.053	0.055	0.050	0.043	0.048
0.05	0.114	0.110	0.120	0.142	0.136	0.139	0.159	0.155	0.145
0.10	0.198	0.185	0.191	0.250	0.227	0.220	0.416	0.393	0.340
0.20	0.445	0.436	0.405	0.617	0.578	0.530	0.876	0.831	0.741
0.30	0.721	0.702	0.633	0.897	0.859	0.796	0.997	0.995	0.970
0.40	0.891	0.872	0.823	0.983	0.963	0.924	1.000	0.999	0.994
0.50	0.964	0.952	0.918	0.999	0.990	0.980	1.000	1.000	1.000
Model (5)									
0.00	0.050	0.048	0.056	0.042	0.042	0.060	0.063	0.043	0.050
0.05	0.118	0.114	0.126	0.128	0.122	0.123	0.124	0.101	0.107
0.10	0.224	0.216	0.213	0.295	0.263	0.259	0.275	0.261	0.243
0.20	0.525	0.491	0.463	0.720	0.660	0.588	0.711	0.675	0.671
0.30	0.790	0.747	0.693	0.948	0.895	0.829	0.962	0.931	0.917
0.40	0.925	0.900	0.849	0.992	0.968	0.941	0.996	0.992	0.989
0.50	0.975	0.954	0.927	1.000	0.991	0.975	1.000	1.000	0.999

Table 4.6 Simulated Power Values of Test Statistics for Testing γ under Alternative Models (6)-(7)

	n=30			**n=50**			**n=100**		
γ	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$
Model (6)									
0.00	0.057	0.057	0.068	0.049	0.046	0.053	0.056	0.052	0.051
0.05	0.107	0.106	0.112	0.127	0.120	0.125	0.184	0.178	0.171
0.10	0.189	0.186	0.194	0.254	0.243	0.225	0.384	0.367	0.324
0.20	0.441	0.435	0.413	0.588	0.578	0.524	0.851	0.842	0.769
0.30	0.684	0.684	0.639	0.865	0.864	0.790	0.986	0.982	0.961
0.40	0.881	0.881	0.825	0.974	0.971	0.948	1.000	1.000	0.999
0.50	0.961	0.963	0.930	0.997	0.997	0.987	1.000	1.000	1.000
Model (7)									
0.00	0.051	0.056	0.068	0.035	0.034	0.047	0.051	0.048	0.054
0.05	0.120	0.117	0.129	0.134	0.122	0.130	0.187	0.155	0.146
0.10	0.221	0.206	0.209	0.297	0.264	0.245	0.496	0.420	0.363
0.20	0.546	0.502	0.466	0.748	0.679	0.587	0.941	0.893	0.787
0.30	0.832	0.780	0.700	0.955	0.906	0.827	0.999	0.991	0.969
0.40	0.949	0.907	0.847	0.994	0.981	0.952	1.000	1.000	0.998
0.50	0.986	0.971	0.937	0.999	0.996	0.988	1.000	1.000	1.000

Table 4.7 Simulated Power Values of Test Statistics for Testing γ under Alternative Models (8)-(9)

	n=30			**n=50**			**n=100**		
γ	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$
Model (8)									
0.00	0.056	0.057	0.064	0.057	0.056	0.056	0.044	0.040	0.051
0.05	0.110	0.104	0.112	0.127	0.120	0.126	0.154	0.149	0.141
0.10	0.183	0.181	0.184	0.246	0.233	0.226	0.368	0.358	0.311
0.20	0.418	0.411	0.383	0.585	0.576	0.515	0.852	0.841	0.738
0.30	0.667	0.659	0.604	0.863	0.850	0.780	0.984	0.980	0.961
0.40	0.857	0.855	0.791	0.966	0.957	0.923	1.000	1.000	0.998
0.50	0.954	0.953	0.915	0.996	0.996	0.985	1.000	1.000	1.000
Model (9)									
0.00	0.049	0.051	0.070	0.042	0.043	0.061	0.043	0.042	0.054
0.05	0.108	0.107	0.123	0.126	0.119	0.123	0.186	0.154	0.146
0.10	0.203	0.196	0.197	0.278	0.245	0.232	0.459	0.404	0.350
0.20	0.505	0.469	0.438	0.700	0.618	0.553	0.913	0.852	0.738
0.30	0.767	0.711	0.657	0.932	0.880	0.808	0.996	0.984	0.954
0.40	0.911	0.872	0.816	0.987	0.963	0.932	1.000	0.999	0.994
0.50	0.971	0.945	0.911	0.999	0.995	0.977	1.000	1.000	1.000

Table 4.8 Simulated Power Values of Test Statistics for Testing γ under Alternative Models (10)-(12)

	n=30			n=50			n=100		
γ	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$	$\mathbf{T_{11}}$	$\mathbf{T_{22}}$	$\mathbf{T_{33}}$
Model (10)									
0.00	0.040	0.049	0.090	0.045	0.046	0.081	0.052	0.041	0.082
0.05	0.068	0.062	0.121	0.067	0.055	0.099	0.097	0.080	0.121
0.10	0.105	0.096	0.151	0.125	0.105	0.137	0.186	0.125	0.155
0.20	0.196	0.176	0.225	0.284	0.221	0.255	0.471	0.338	0.330
0.30	0.348	0.294	0.338	0.489	0.378	0.376	0.740	0.519	0.466
0.40	0.504	0.426	0.449	0.711	0.534	0.514	0.926	0.717	0.650
0.50	0.650	0.540	0.550	0.839	0.686	0.647	0.988	0.860	0.793
Model (11)									
0.00	0.032	0.033	0.062	0.025	0.024	0.068	0.032	0.014	0.067
0.05	0.049	0.049	0.092	0.053	0.027	0.093	0.072	0.026	0.096
0.10	0.062	0.048	0.098	0.066	0.047	0.099	0.091	0.048	0.111
0.20	0.102	0.069	0.128	0.143	0.054	0.115	0.222	0.047	0.127
0.30	0.161	0.093	0.154	0.251	0.092	0.158	0.422	0.093	0.161
0.40	0.258	0.122	0.179	0.360	0.124	0.177	0.609	0.129	0.170
0.50	0.345	0.263	0.216	0.495	0.262	0.228	0.868	0.262	0.228
Model (12)									
0.00	0.033	0.038	0.041	0.036	0.032	0.030	0.039	0.017	0.083
0.05	0.049	0.037	0.041	0.047	0.026	0.050	0.048	0.020	0.070
0.10	0.060	0.045	0.057	0.066	0.031	0.078	0.083	0.030	0.106
0.20	0.082	0.055	0.079	0.109	0.044	0.092	0.164	0.046	0.109
0.30	0.136	0.083	0.136	0.179	0.069	0.130	0.317	0.111	0.140
0.40	0.181	0.145	0.155	0.270	0.174	0.146	0.472	0.186	0.164
0.50	0.255	0.214	0.179	0.360	0.216	0.187	0.689	0.192	0.180

Figure 4.1-4.4 gives the power curves of the test statistics for ϕ and γ under selected alternative Models (1) and (5).

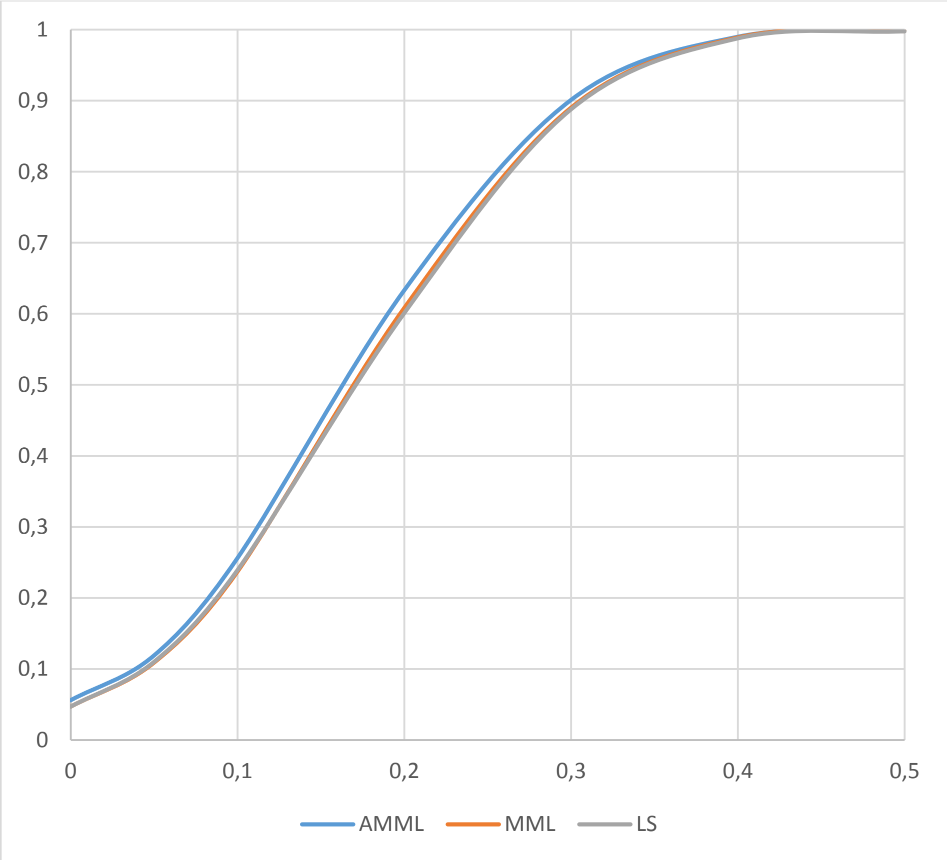

Figure 4.1 Power Curves of the Test Statistics for Model 1; ϕ; n=100

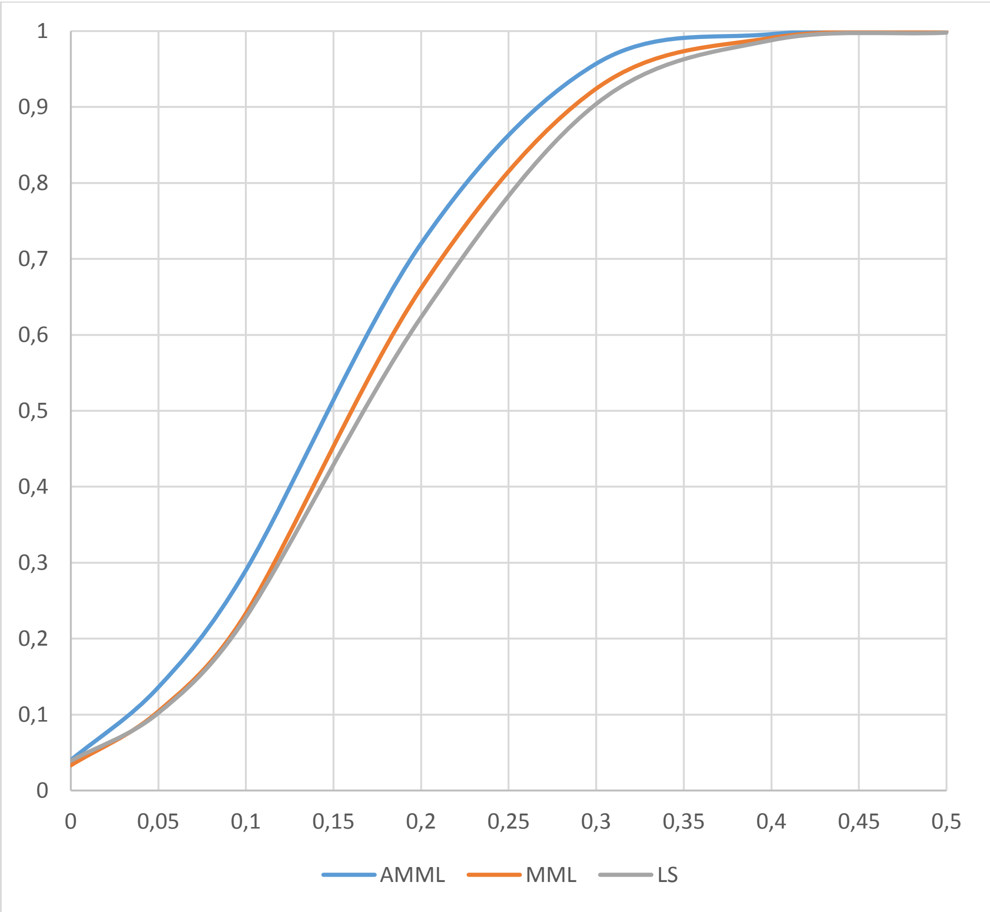

Figure 4.2 Power Curves of the Test Statistics for Model 5; ϕ; n=100

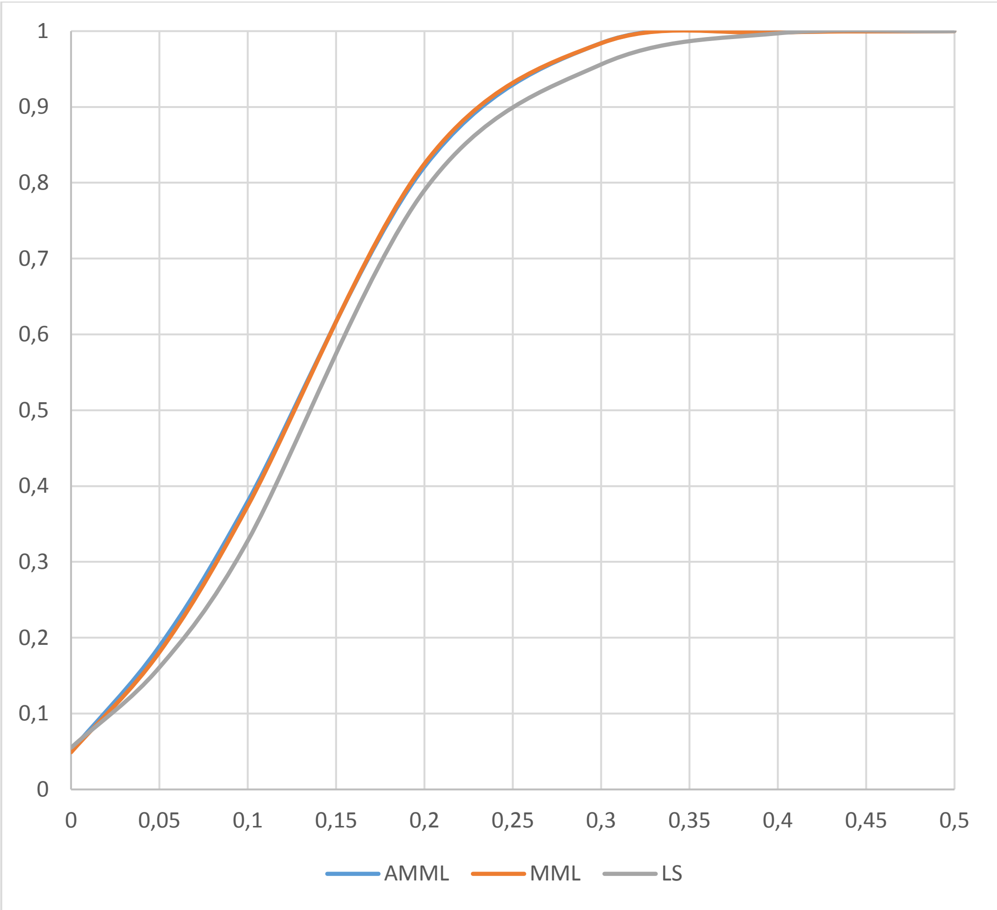

Figure 4.3 Power Curves of the Test Statistics for Model 1; γ; n=100

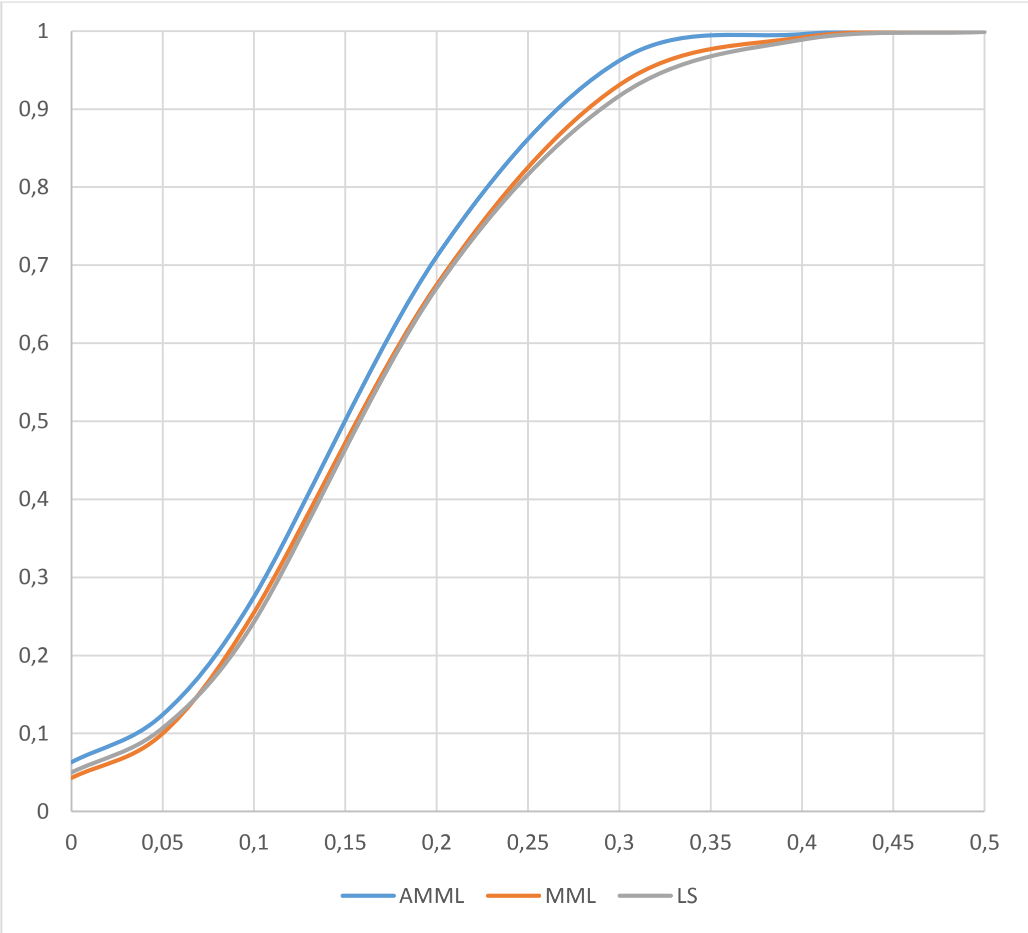

Figure 4.4 Power Curves of the Test Statistics for Model 5; γ; n=100

CHAPTER 5

CONCLUSION

In this thesis, the AR(1) model with a single explanatory variable where the error term has a LTS distribution is considered. Methods for estimating the parameters of autoregressive model are explained and discussed with their advantages and disadvantages. As known, LSEs are not efficient, obtaining MLEs might be problematic with some algorithms such as; EM and ECM, and MML technique assumes the shape parameter to be known. Because of these drawbacks, the idea of Huber M-estimation combined with MML is proposed and named as AMML.

In order to examine the efficiency and robustness properties of the estimators Monte Carlo simulation study is conducted. Estimators are compared with each other and it is seen that LS estimators are less efficient than MML and AMML estimators. In many cases ECM gives similar results to MML. Under extreme distributions like Slash and Cauchy it is observed that only AMML method is robust not ignoring the bias in σ.

Finally, test statistics for testing hypothesis for the parameters in the model are proposed. Their efficiency and robustness properties are investigated via simulation study and test statistics based on LS, MML and AMML are compared with each other with respect to their power. According to the findings, it is seen that under normality results are close to each other but as it deviates from normality, AMML based test statistics become more powerful comparatively.

REFERENCES

Akkaya, A. D. and Bayrak, Ö. T. (2018). A New Estimation Technique for AR(1) Model with Long-Tailed Symmetric Innovations. Time Series Analysis and Forecasting. ITISE 2017. Contributions to Statistics. Springer, Cham.

Andrews, D. F., Bickel, P. J., Hampel, F., Huber, P., Rogers, W., and Tukey, J. W. (1972). Robust Estimation of Location: Survey and Advances. Princeton University press: Princeton, N. J.

Andrews, D. F. (1974). A robust method for multiple linear regression. Technometrics 16, 523-531.

Balci, S. and Akkaya, A. D. (2015). Robust Pairwise Multiple Comparisons under short-tailed symmetric distributions. Journal of Applied Statistics, 42, 2293-2306.

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

Bayrak, Ö. T. and Akkaya, A. D. (2010). Estimating parameters of a multiple autoregressive model by the modified maximum likelihood method. Journal of Computational and Applied Mathematics 233 (8), 1763–1772.

Beaton, A. E. and Tukey, J. W. (1974). The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data. Technometrics, 16, 147-185.

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

Birch, J. B. and Myers, R. H. (1982). Robust analysis of covariance. Biometrics 38, 699-713.

Dempster, A. P., N. M. Laird, and D. B. Rubin. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion.). Journal of the Royal Statistical Society, Series B 39, 1–38.

Dönmez, A. (2010). Adaptive estimation and hypothesis testing methods. PhD thesis, Middle East Technical University.

Dunnett, C. W. (1982). Robust multiple comparisions. Commun. Stat. Theory.Meth. 11, 2611-2629.

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

Gross, A. M. (1976). Confidence interval robustness with long-tailed symmetric distributions. J. Amer. Stat. Assoc. 71, 409-416.

Gross, A. M. (1977). Confidence intervals for bisquare regression estimate. J. Amer. Stat. Assoc. 72, 341-354.

Haghbin, H. and Nematollahi, A. R. (2013). Likelihood-based inference in autoregressive models with scaled t-distributed innovations by means of EM-based algorithms. Communications in Statistics - Simulations and Computation 42 (10), 2239–2252.

Hampel, F. R. (1968). Contributions to the theory of robust estimation. Ph.D. thesis, University of California. Berkeley.

Hampel, F. R. (1971). A General Qualitative Definition of Robustness. Ann. Math. Statist. 42, 1887-1896.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. J. Amer. Statist. Assoc., 62, 1179-1186.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). Robust Statistics. The Approach Based on Influence Functions. New York: John Wiley and Sons.

Huber, P. J. (1964). Robust estimation of a location parameter. Annals Math. Stat. 35, 73-101.

Huber, P. J. (1977). Robust Statistical Procedures. Regional Conference Series in Applied Mathematics No. 27, Soc. Industr. Appl. Math., Philadelphia.

Huber, P. J. (1981). Robust statistics. New York: John Wiley.

Islam, M. Q., Tiku, M. L. (2004). Multiple linear regression model under non-normality. Commun. Stat.- Theory and Methods, 33, 10, 2443-2467.

Kendall, M. G. and Stuart, A. (1979). The Advanced Theory of Statistics, Vol. 3, Mc Millan: New York.

Kramer, W. (1980). Finite sample efficiency of ordinary least squares in the linear regression with autocorrelated errors. J. Amer. Stat. Assoc. 75, 1005-1009.

Meng, X. L., and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika Royal Statistical Society, Series B 44 (2), 226–233.

Nduka, U. C., Iwueze, I. S. and Nwogu, E. C. (2018). Efficient algorithms for robust estimation in autoregressive regression models using Student's t distribution, Communications in Statistics - Simulation and Computation. 47 (1), 206-228.

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.

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

Tan, W. Y. and Lin, V. (1933). Some robust procedures for estimating parameters in an autoregressive model. Sankya B, 55, 415-435.

Tiku, M. L. (1967). Estimating the mean and the standart deviation from a censored normal sample. Biometrika 54, 155-165.

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. Inf. 30, 281-292.

Tiku, M. L., Wong, W. K. and Bian, G. (1999). Time series models with asymmetric innovations. Commun. Stat.- Theory Meth. 28, 1131-1160.

Tiku, M. L., Wong, W. K., Vaughan, D. C. and Bian, G. (2000). Time series models in non-normal situations: symmetric innovations. J. Time Series Analysis 21, 571-596.

Tiku, M. L. and Akkaya, A. D. (2004). Robust Estimation and Hypothesis Testing. New Age International Limited, Publishers: New Delhi.

Tiku, M. L. and Sürücü, B. (2009). MMLEs are as good as M-estimators or better. Statistics and Probability Letters, 79, 984-989.

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

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.

Vinod, H. D. and Shenton, L. R. (1996). Exact moment for autoregressive and random walk models for a zero or stationary initial value. Econometric Theory 12, 481-499.

APPENDIX

Simulation Study Codes in R for Model 1

```
#required packages

library(Metrics)

rm(list=ls())

mu=0

del=1

fi=0.5

sigma=1

n=100

mm=round(100000/n)

t4=numeric()

t=numeric()

dell=numeric()

alf=numeric()

bet=numeric()

em=numeric()

RA=numeric()

RA1=numeric()

e=numeric()

ee=numeric()

B=0

C=0
```

```
D=0
E=0
wt=numeric()
qtem=numeric()
m11=numeric()
xt=numeric()
xt1=numeric()
m12=numeric()
m22=numeric()
M=numeric()
yt=numeric()
yt1=numeric()
m011=numeric()
m012=numeric()
M0=numeric()
del_lsem=numeric()
delerr=numeric()
sigmaerr=numeric()
muerr=numeric()
fierr=numeric()
sigma_lsem=numeric()
fi_lsem=numeric()
mu_lsem=numeric()
mu_emlast=numeric()
del_emlast=numeric()
fi_emlast=numeric()
```

```
sigma_emlast=numeric()
y=numeric()
x=numeric()
y1=numeric()
x1=numeric()
yen=numeric()
zen=numeric()
w=numeric()
ws=numeric()
wind=numeric()
v=numeric()
u=numeric()
ve=numeric()
ue=numeric()
vbar=numeric()
ubar=numeric()
g1=numeric()
g2=numeric()
g=numeric()
pk=numeric()
be=numeric()
ce=numeric()
sigma_gec=numeric()
mu_gec=numeric()
del_gec=numeric()
fi_gec=numeric()
```

```
gam_gec=numeric()

vt=numeric()

rl=numeric()

thet1=numeric()

thet0=numeric()

www=numeric()

dff=numeric()

soes=numeric()

adff=numeric()

amuid=numeric()

delid=numeric()

fid=numeric()

gamid=numeric()

sid=numeric()

tamm=numeric()

dell.a=numeric()

alf.a=numeric()

bet.a=numeric()

em.a=numeric()

en.a=numeric()

mu_ls=numeric()

del_ls=numeric()

fi_ls=numeric()

gam_ls=numeric()

sigma_ls=numeric()

mu_mml=numeric()
```

```
del_mml=numeric()
fi_mml=numeric()
gam_mml=numeric()
sigma_mml=numeric()
mu_amml=numeric()
del_amml=numeric()
fi_amml=numeric()
gam_amml=numeric()
sigma_amml=numeric()
# Calculation of t(i) values
p=16.5
df=2*p-1
k=2*p-3
t4=seq(1:n)/(n+1)
t=qt(t4,df)* sqrt((df-2)/df)
dell=(1+(1/k)*t^2)^2
alf=((2/k)*t^3)/dell
bet=(1-(1/k)*t^2)/dell
if(bet[1]<0)  {
dell=(1+(1/k)*t^2)^2
alf=((1/k)*t^3)/dell
bet=1/dell
}
em=sum(bet)
#Beginning of Monte Carlo
for(j in 1:mm){
```

```
#Data Generation for Model 1

RA=runif(n)

x[2:(n+1)]=(RA[1:n]-0.5)*sqrt(12)

x[2:(n+1)]=x[2:(n+1)]/sqrt(1-fi^2)

RA1=runif(1)

x[1]=(RA1[1]-0.5)*sqrt(12)

x[1]=x[1]/sqrt(1-fi^2)

e=rnorm(n)

ee=rnorm(1)

y[1]=ee[1]/(sqrt(1-fi^2))

# End of Data Generation

# Beginning of LS

for (i in 2:(n+1)){

y[i]=fi*y[i-1]+mu+del*(x[i]-fi*x[i-1])+e[i-1]

}

XiSum=0

Xi2Sum=0

Xi.Sum=0

XiiSum=0

Xi.2Sum=0

Yi.Sum=0

Yi.2Sum=0

Yi.XiSum=0

Yi.Xi.Sum=0

YiSum=0

YiXiSum=0
```

```
YiYi.Sum=0

YiXi.Sum=0

XiSum=sum(x[2:(n+1)])

Xi2Sum=sum(x[2:(n+1)]^2)

Xi.Sum=sum(x[1:n])

XiiSum=sum(x[2:(n+1)]*x[1:n])

Xi.2Sum=sum(x[1:n]^2)

Yi.Sum=sum(y[1:n])

Yi.2Sum=sum(y[1:n]^2)

Yi.XiSum=sum(y[1:n]*x[2:(n+1)])

Yi.Xi.Sum=sum(y[1:n]*x[1:n])

YiSum=sum(y[2:(n+1)])

YiXiSum=sum(y[2:(n+1)]*x[2:(n+1)])

YiYi.Sum=sum(y[2:(n+1)]*y[1:n])

YiXi.Sum=sum(y[2:(n+1)]*x[1:n])

B=matrix(c(n,XiSum,Yi.Sum,Xi.Sum,XiSum,Xi2Sum,Yi.XiSum,XiiSum,Yi.Sum,Yi.XiSum,Yi.2Sum,Yi.Xi.Su
m,Xi.Sum,XiiSum,Yi.Xi.Sum,Xi.2Sum),4,4,byrow=TRUE)

C=matrix(c(YiSum,YiXiSum,YiYi.Sum,YiXi.Sum),4,1)

D=solve(B)

E=D%*%C

mu_ls[j]=E[1]

del_ls[j]=E[2]

fi_ls[j]=E[3]

gam_ls[j]=E[4]

err=numeric()

sumerr=numeric()

for (i in 2:(n+1)){
```

```
err[i-1]=y[i]-fi_ls[j]*y[i-1]-del_ls[j]*(x[i]-fi_ls[j]*x[i-1])-mu_ls[j]
}
sumerr=(sum(err^2))/(n-3)
sigma_ls[j]=sqrt(sumerr)
# End of LS
# Beginning of ECM
sigma_lsem=sigma_ls[j]
del_lsem=del_ls[j]
fi_lsem=fi_ls[j]
mu_lsem=mu_ls[j]
delerr=1
sigmaerr=1
muerr=1
fierr=1
while((abs(delerr)>0.0005)||(abs(sigmaerr)>0.0005)||(abs(muerr)>0.0005)||(abs(fierr)>0.0005)){
wt=(((df+1)*(df-2))/df)*(sigma_lsem^2/((df-2)* sigma_lsem^2+err^2))
qtem=digamma((df+1)/2)-log(df/2)-log(((df-2)* sigma_lsem^2+err^2)/((df-2)* sigma_lsem^2))
m11= sum(wt)
xt=x[-1]
xt1=x[-(n+1)]
m12=sum(wt*(xt-fi_lsem*xt1))
m22=sum(wt*(xt-fi_lsem*xt1)^2)
#Calculation of mu-del-fi-sigma
M=matrix(c(m11,m12,m12,m22),byrow=TRUE,nrow=2)
yt=y[-1]
yt1=y[-(n+1)]
```

```
m011=sum(wt*(yt-fi_lsem*yt1))

m012=sum(wt*(xt-fi_lsem*xt1)*(yt-fi_lsem*yt1))

M0=c(m011,m012)

del_em=solve(M)%*%M0

sigma_em=sqrt((df/(n*(df-2)))*sum(wt*err^2))

fi_em=(sum(wt*(yt1-del_lsem*xt1)*(yt-mu_lsem-del_lsem*xt)))/(sum(wt*(yt1-del_lsem*xt1)^2))

del_lsem=c(mu_lsem,del_lsem)

delerr=del_em[2]-del_lsem[2]

sigmaerr=sigma_em-sigma_lsem

muerr=del_em[1]-mu_lsem

fierr=fi_em-fi_lsem

del_lsem=del_em[2]

sigma_lsem=sigma_em

fi_lsem=fi_em

mu_lsem=del_em[1]

}

del_emlast[j]=del_em[2]

sigma_emlast[j]=sigma_em

mu_emlast[j]=del_em[1]

fi_emlast[j]=fi_em

#End of ECM

#Beginning of MML

fi_gec=fi_ls[j]

del_gec=del_ls[j]

gam_gec=gam_ls[j]

for(s in 1:2) {
```

```
if (s==1){
for(i in 2:(n+1)) {
w[i]=y[i]-fi_gec*y[i-1]-del_gec*x[i]+gam_gec*x[i-1]
}
}
else {
for(i in 2:(n+1)) {
w[i]=(y[i]-fi_gec*y[i-1])-del_gec*(x[i]-fi_gec*x[i-1])
}
}
w=w[-1]
#Finding Concomitants
for(i in 2:(n+1)) {
y1[i]=y[i]
x1[i]=x[i]
yen[i]=y[i-1]
zen[i]=x[i-1]
}
y1=y1[-1]
x1=x1[-1]
yen=yen[-1]
zen=zen[-1]
ws=numeric()
wind=numeric()
ws=sort(w)
wind=order(w)
```

```
y1=y1[wind]
x1=x1[wind]
yen=yen[wind]
zen=zen[wind]
#End of Concomitant Part
#Calculation of Sigma
v=y1-fi_gec*yen
u=x1-fi_gec*zen
ve=sum(bet*v)
ue=sum(bet*u)
vbar=ve/em
ubar=ue/em
g1=sum(bet*v*u)-em*vbar*ubar
g2=sum(bet*(u^2))-em*(ubar^2)
g=g1/g2
be1=numeric()
ce1=numeric()
be1=sum(alf*(v-g*u))
ce1=sum(bet*((v-vbar)^2))-g*sum(bet*((v-vbar)*(u-ubar)))
be=be1*(2*p)/k
ce=ce1*(2*p)/k
sigma_gec=(be+sqrt(be^2+4*n*ce))/(2*sqrt(n*(n-3)))
#Calculation of del
he1=numeric()
he2=numeric()
h=numeric()
```

```
he1=sum(alf*(u-ubar))

he2=sum(bet*((u-ubar)^2))

h=he1/he2

del_gec=g+(h*sigma_gec)

#Calculation of mu

mu_gec=vbar-del_gec*ubar

#Calculation of fi

as1=numeric()

as2=numeric()

as3=numeric()

as4=numeric()

k1=numeric()

L1=numeric()

as1=sum(bet*(y1-del_gec*x1)*(yen-del_gec *zen))

as2=sum(bet*(yen-del_gec*zen)^2)

as3=sum(bet*(yen-del_gec*zen))

as4=sum(alf*(yen-del_gec*zen))

k1=(as1-mu_gec*as3)/as2

L1=as4/as2

fi_gec=k1+L1*sigma_gec

# Calculation of gamma

gam_gec=-fi_gec*del_gec

} # end of s

mu_mml[j]=mu_gec

del_mml[j]=del_gec

fi_mml[j]=fi_gec
```

```
sigma_mml[j]=sigma_gec

gam_mml[j]=gam_gec

#End of MML

#Beginning of AMML

vt=y[1:n]+x[2:(n+1)]-x[1:n]

rl=(y[3:(n+1)]-y[2:n])/(vt[3:(n+1)]-vt[2:n])

thet1=median(rl,na.rm=TRUE)

www=y[2:(n+1)]-thet1*vt[2:(n+1)]

thet0=median(www,na.rm=TRUE)

dff=abs(www-thet0)

adff=dff[1:(n-1)]

soes=1.483*median(adff)

fid=thet1

delid=thet1

amuid=thet0

sid=soes

gamid=thet1

for(z in 1:2) {

if (z==1) {

for(u in 1:n) {

tamm[u]=(y[u+1]-amuid-fid*y[u]-delid*x[u+1]+gamid*x[u])/sid

}

}

else {

for(u in 1:n) {

tamm[u]=(y[u+1]-amuid-fid*y[u]-delid*(x[u+1]-fid*x[u]))/sid
```

```
}
}
dell.a=(1+(1/k)*tamm^2)^2
alf.a=((2/k)*tamm)/dell.a
bet.a=1/dell.a
em.a=sum(bet.a)
en.a=sum(alf.a)
#Calculation of Sigma
v.a=numeric()
u.a=numeric()
ve.a=numeric()
ue.a=numeric()
vbar.a=numeric()
ubar.a=numeric()
v.a=y[2:(n+1)]-fid*y[1:n]
u.a=x[2:(n+1)]-fid*x[1:n]
ve.a=sum(bet.a*v.a)
ue.a=sum(bet.a*u.a)
vbar.a=ve.a/em.a
ubar.a=ue.a/em.a
g1.a=numeric()
g2.a=numeric()
g.a=numeric()
g1.a=sum(bet.a*v.a*u.a)-em.a*vbar.a*ubar.a
g2.a=sum(bet.a*(u.a^2))-em.a*ubar.a^2
g.a=g1.a/g2.a
```

```
be1.a=0

ce1.a=0

be.a=0

ce.a=0

be.a1=sum(alf.a*(v.a-g.a*u.a))

ce.a1=sum(bet.a*((v.a-vbar.a)^2))-g.a*sum(bet.a*((v.a-vbar.a)*(u.a-ubar.a)))

be.a=be.a1*(2*p)/k

ce.a=ce.a1*(2*p)/k

sid=(be.a+sqrt(be.a^2+4*n*ce.a))/(2*sqrt(n*(n-3)))

#Calculation of del

he1.a=numeric()

he2.a=numeric()

he1.a=sum(alf.a*(u.a-ubar.a))

he2.a=sum(bet.a*((u.a-ubar.a)^2))

h.a=he1.a/he2.a

delid=g.a+(h.a*sid)

#Calculation of mu

amuid=vbar.a-(delid*ubar.a)

#Calculation of fi

as1.a=numeric()

as2.a=numeric()

as3.a=numeric()

as4.a=numeric()

k1.a=numeric()

L1.a=numeric()

as1.a=sum(bet.a*(y[2:(n+1)]-delid*x[2:(n+1)])*(y[1:n]-delid*x[1:n]))
```

```
as2.a=sum(bet.a*(y[1:n]-(delid*x[1:n]))^2)

as3.a=sum(bet.a*(y[1:n]-(delid*x[1:n])))

as4.a=sum(alf.a*(y[1:n]-(delid*x[1:n])))

k1.a=(as1.a-amuid*as3.a)/as2.a

L1.a=as4.a/as2.a

fid=k1.a+L1.a*sid

#Calculation of gamma

gamid=-fid*delid

}

mu_amml[j]=amuid

del_amml[j]=delid

gam_amml[j]=gamid

fi_amml[j]=fid

sigma_amml[j]=sid

}

#End of AMML

#End of Monte Carlo

mean(mu_ls)

var(mu_ls)

mse(mu,mu_ls)

mean(del_ls)

var(del_ls)

mse(del,del_ls)

mean(fi_ls)

var(fi_ls)

mse(fi,fi_ls)
```

```
mean(sigma_ls)

var(sigma_ls)

mse(sigma,sigma_ls)

mean(mu_emlast)

var(mu_emlast)

mse(mu,mu_emlast)

mean(del_emlast)

var(del_emlast)

mse(del,del_emlast)

mean(fi_emlast)

var(fi_emlast)

mse(fi,fi_emlast)

mean(sigma_emlast)

var(sigma_emlast)

mse(sigma,sigma_emlast)

mean(mu_mml)

var(mu_mml)

mse(mu,mu_mml)

mean(del_mml)

var(del_mml)

mse(del,del_mml)

mean(fi_mml)

var(fi_mml)

mse(fi,fi_mml)

mean(sigma_mml)

var(sigma_mml)
```

```
mse(sigma,sigma_mml)

mean(mu_amml)

var(mu_amml)

mse(mu,mu_amml)

mean(del_amml)

var(del_amml)

mse(del,del_amml)

mean(fi_amml)

var(fi_amml)

mse(fi,fi_amml)

mean(sigma_amml)

var(sigma_amml)

mse(sigma,sigma_amml)
```