UNIT ROOT PROBLEMS IN TIME SERIES ANALYSIS

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

BY

VİLDA PURUTCUOĞLU

IN PARTIAL FULFILLEMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF STATISTICS

JANUARY 2004

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. H. ÖztaşAyhan
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Moti Lal Tiku
Supervisor

Examining Committee Members

Prof. Dr. Moti Lal Tiku

Prof. Dr. Ali Uzun

Assoc. Prof. Dr. AyşenAkkaya

Assist. Prof. Dr. Nadir Öcal

Dr. Evrim Oral

ABSTRACT

UNIT ROOT PROBLEMS IN TIME SERIES ANALYSIS

Purutcuoğlu, Vilda

Ms., Department of Statistics

Supervisor: Prof. Dr. Moti Lal Tiku

January 2004, 177 pages

In time series models, autoregressive processes are one of the most popular stochastic processes, which are stationary under certain conditions. In this study we consider nonstationary autoregressive models of order one, which have iid random errors. One of the important nonstationary time series models is the unit root process in AR (1), which simply implies that a shock to the system has permanent effect through time. Therefore, testing unit root is a very important problem.

However, under nonstationarity, any estimator of the autoregressive coefficient does not have a known exact distribution and the usual t – statistic is not accurate even if the sample size is very large. Hence,

Wiener process is invoked to obtain the asymptotic distribution of the LSE $\tilde{\phi}_1$ under normality. The first four moments of $\tilde{\phi}_1$ under normality have been worked out for large n.

In 1998, Tiku and Wong proposed the new test statistics R_1 and R_0 whose type I error and power values are calculated by using three – moment chi – square or four – moment F approximations. The test statistics are based on the modified maximum likelihood estimators and the least square estimators, respectively. They evaluated the type I errors and the power of these tests for a family of symmetric distributions (scaled Student's t). In this thesis, we have extended this work to skewed distributions, namely, gamma and generalized logistic.

Keywords: AR (1) Models, Modified Maximum Likelihood Estimator, Unit Root Tests, Three - Moment Chi – Square Approximation, Four – Moment F Approximation, Fisher Information Matrix, Non – normality.

ÖZ

ZAMAN SERİSİ ANALİZLERİNDE BİRİM KÖK PROBLEMLERİ

Purutcuoğlu, Vilda

Yüksek Lisans, İstatistik Bölümü

Tez Yöneticisi: Prof. Dr. Moti Lal Tiku

Ocak 2004, 177 sayfa

Zaman serisi modellerinde, “otoregresyonel” işlemler, belirli şartlar altında durağan olan en popüler rastsal işlemlerden biridir. Bu çalışmada birinci derecede, bağımsız ve benzer dağılımlı hata terimlerine sahip, durağan olmayan otoregresyonel modelleri dikkate alıyoruz. Zaman serilerinde, durağan olmayan önemli modellerden biri, AR (1) modellerindeki birim kök işlemleridir. Bu işlemler basitçe, sisteme gelen bir şokun zaman içerisindeki kalıcı etkisini ifade eder. Bu nedenle birim kökü test etmek önemli bir sorundur.

Buna rağmen, durağan olmayan otoregresyonel katsayının bilinen kesin bir dağılımı yoktur ve genel t – istatistiği, örnek hacim büyük olsa bile

kullanılması doğru değildir. Dolayısıyla katsayının, en küçük kareler tahmin edicisinin $\tilde{\phi}_1$,normal dağılım varsayımı altındaki asimptotik dağılımını elde etmek için Wiener işlemi kullanılır. Bu katsayının normal varsayımı altında ilk dört momenti büyük örnek hacimleri için hesaplanmıştır.

1998'de Tiku ve Wong, birinci tip hataları ve güç değerleri üç – momentli ki – kare veya dört – momentli F tahmini kullanılarak hesaplanan yeni test istatistiği R_1 ve R_0' yu önermişlerdir. Bu istatistikler sırasıyla "uyarlanmış en çok olabilirlik" ve "en küçük kareler" tahmin edicilerine dayalıdır. Tiku ve Wong testlerin birinci tip hatalarını ve güçlerini simetrik dağılım ailesi için hesaplamışlardır. Bu tezde, biz bu çalışmayı çarpık dağılımlar yani gama ve genelleştirilmiş lojistik için genişlettik.

Anahtar Kelimeler: AR (1) Modelleri, Uyarlanmış En Çok Olabilirlik Tahmin Edicileri, En Küçük Kareler Tahmin Edicileri, Birim Kök Testleri, Üç – Momentli Ki – Kare Tahmini, Dört – Momentli F Tahmini, Fisher Bilgi Matriksi, Normal Olmayan Dağılımlar.

ACKNOWLEDGEMENTS

I would like to express my deep respect, love and appreciation to my supervisor Prof. Dr. Moti Lal Tiku for all his contributions, guidance and tolerance throughout my study.

I am very thankful to the members of my jury with my deeper respect, Prof. Dr. Ali Uzun, Assoc. Prof. Dr. Ayşen Akkaya, Assist. Prof. Dr. Nadir Öcal, and Dr. Evrim Oral because of their precious suggestions and interests

I am very grateful with my sincere respect to my teachers Prof. Dr. H. Öztaş Ayhan, Assoc. Prof. Dr. Bilgehan Güven, Assist. Prof. Dr. İnci Batmaz, Assist. Prof. Dr. Birdal Şenoğlu and Inst. S. Ömer Gücelioğlu, and all my other teachers for their supports and considerations during my studies.

I would like to thank to Reser. Assist. B. Emre Ülgen for his great support and tolerance and to Dr. Savaş Sazak for his help and interest.

My deepest appreciation goes to my parents, my mother Yurdagül Purutcuoğlu and my father Ahmet Purutcuoğlu. Throughout their lives, they never avoid any devotion for my older sister and me. I feel proud of having such a family and everyday I thank God as I belong to such a beautiful family.

I would like to thank to my single sister Eda Purutcuoğlu for her invaluable support and love. She is always an excellent older sister for me on every aspect in spite of only one - year difference between us.

Finally to all my friends, I offer sincere thanks for their understanding and patience during my work.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

CHAPTER I

INTRODUCTION AND LITERATURE SURVEY

Time series models are one of the most popular stochastic processes, which are stationary under certain conditions. In practice many of the econometrics data are modelled by means of AR (p) processes.

Most of these cases, error terms are assumed to be normal. However with respect to the real life situation, it does not occur generally; that is, error terms have a non - normal distribution. But LSE (least square estimators), used to estimate unknown parameters in the model, are inefficient when we deal with non-normal cases. Similarly another well - known estimators, which are MLE (maximum likelihood estimators), become also unsuccessful in numerous situations since we can not get always explicit solutions from the likelihood equations and we need iterations to get accurate results. Actually, there are certain iterative methods such as Newton - Raphson or Steepest Ascent in order to estimate parameters under these conditions. But they indicate mostly local maximums instead of global ones. Moreover their successes depend on the shape of the function. For instance if function is concave, Newton- Raphson has poor result. Steepest Ascent method may give more reliable solution than the former; however, it converges very slowly (Hamilton 1994, p. 139).

Thereby estimators based on iterative methods are not optimal. At this point it is obvious that we need a robust estimator while the error terms are not normally distributed.

Hence in this thesis, Modified Maximum Likelihood Estimators (MMLE) are preferred. AR (1) model without intercept term are analysed under different shape parameters of gamma, generalized logistic, and long – tailed symmetric distributions. Efficiency properties of the MML and the LS estimators are compared. Then, similar comparisons are done for AR (1) model with intercept term as an extension of our initial model.

Additionally, one of the other purposes of this study is to test unit root in AR (p) process. Existence of unit root is a vital issue in time series. Because unit root process shows nonstationarity, and nonstationarity implies that a shock to the system has permanent effect through time. This behaviour can be opposite of a stationary AR (p) process since here, all past errors have an impact on y_t, but this impact disappears when time goes on. Therefore deciding whether nonstationary exists or not in the data is a very crucial question. For that reason many unit root tests are applied in econometrics. In these tests, we could not use usual t – statistic even if the sample size is large. This is due to the fact that the autoregressive coefficient does not have an exact distribution and also is not asymptotically normal. Dickey and Fuller, Phillips and Perron, and Abadir developed certain test statistics, which are mainly based on the normal distribution of the errors and the LS estimators of the unknown parameters in the model. But Blough (1990) proved that none of them is better than the others and they have almost the same low power.

Finally from this point of view the asymptotic distribution for unit root process becomes an important topic. Abadir (1995) formulated its asymptotic distribution, and Vinod and Shenton (1996) found the first four moments of the

LSE in AR (1) model without intercept term. Then, Brownian motion was applied to determine the limit as sample size tends to infinity.

Subsequent to these works, Tiku and Wong (1998) proposed new test statistics R_1 (based on the MML estimators) and R_0 (based on the LS estimators) for testing a unit root. They calculated their Pearson coefficients for the family of symmetric distributions (Student - t) ranging from Cauchy to Normal. Tiku, Wong and Bian (1999) did similar works for gamma and generalized logistic distribution for testing $\phi_1 = 1$. They showed that these coefficients are close to the Type III line or the F - region (Pearson and Tiku 1970), for all sample sizes. This enables us to use three - moment chi - square and four – moment F approximations to specify critical values and to evaluate type I error and power when unit root exists in AR (1) models. They indicated that the test based on R_1 is more powerful than the test based on R_0.

Consequently in this study, R_1 and R_0 statistics are used to test unit root for AR (1) models with / without intercept term when the distribution of error terms is long – tailed symmetric, gamma and generalized logistic. For all analyses, the initial value y_0 is taken as zero but the results are essentially the same if y_0 is random with mean zero. All simulations are found from 10000 Monte Carlo runs for $n \leq 100$, and 5000 runs for n>100. Standard deviation σ is taken to 1 without loss of generality owing to the fact that the MML and LS estimators are scale invariant. Also, μ is taken to be equal to zero without loss of generality. So in Chapter I, some basic definitions related to the time series and a brief summary of previous studies about unit root case and its testing procedure are presented. Thus the meaning of certain important terms like stationarity, AR models, unit root, and different types of unit root tests are explained. In Chapter II, a detailed theoretical background and the methodology used in deriving MML estimators are given

besides their comparisons with LS estimators under different distributions and sample sizes. Chapter III represents details relevant to the new unit root tests R_1 and R_0 which are based on the methods introduced in previous chapter. Here, also certain comparisons related to the power and type I error of these test statistics are performed in addition to some information about calculation of approximate probability values. Moreover, in Special Appendix, a few examples, which are based on the real life data, are presented.

1.1. Basic Definitions

Time Series: A time series is a collection of observations indexed by the date of each observation. From a theoretical point of view, a time series is a collection of random variables $\{y_t\}$. Such a collection of random variables ordered in time is a "stochastic process". Thus the time series is a stochastic process. If the set of these observations is continuous, the time series is said to be continuous. If the set is discrete, the time series is said to be discrete.

Stationarity: If neither the mean μ_t nor the autocovariances γ_{jt} depend on the date t, then the process for y_t, is said to be covariance – stationary or weakly stationary:

$$E(y_t) = \mu \text{ for all t,}$$

$$E(y_t - \mu)(y_{t-j} - \mu) = \gamma_j \text{ for all t and any j.}$$

If a process is covariance-stationary, the covariance between y_t and y_{t-j} depends only on j, the length of time separating the observations. Thus for any covariance – stationary process;

$$\gamma_j = \gamma_{-j} \text{ for all integers j.}$$

A process is said to be strictly stationary if for any values of $j_1, j_2, ..., j_n$; the joint distribution of $(y_t, y_{t+j_1}, y_{t+j_2}, ..., y_{t+j_n})$ depends only on the intervals separating the dates $(j_1, j_2, ..., j_n)$ and not on the date itself (t). However since stationarity is defined in terms of the distribution function, it is difficult to verify in practice. For this reason, a definition of stationarity in terms of the moments is preferable. Consequently if a process is strictly stationary with finite second moment, then it must be covariance – stationary.

A process that does not obey the conditions of stationarity explained above is called a nonstationary process.

1.1.1. Autoregressive Models (AR)

Most of the practically occurring series can be represented by a stochastic model called AR model. In this model the current value of the process is expressed as a finite, linear aggregate of previous values of the process and an error term ε_t. An AR process of order p without intercept is represented by the equation,

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t . \qquad (1.1)$$

In terms of the lag operator or backward operator L defined by $L^j y_t = y_{t-j}$ for all j, the AR process can be written as

$$y_t = (\phi_1 L + \phi_2 L^2 + ... + \phi_p L^p) y_t + \varepsilon_t$$
$$(1 - \phi_1 L - \phi_2 L^2 - ..... - \phi_p L^p) y_t = \varepsilon_t . \tag{1.2}$$

Therefore,

$$y_t = \frac{1}{1 - \phi_1 L - \phi_2 L^2 - ..... - \phi_p L^p} \varepsilon_t \tag{1.3}$$

$$= \frac{1}{(1 - \pi_1 L)(1 - \pi_2 L)..... (1 - \pi_p L)} \varepsilon_t$$

where $\pi_1, \pi_2, ..., \pi_p$ are root of the equation

$$Z^p - \phi_1 Z^{p-1} - \phi_2 Z^{p-2} - ..... - \phi_p = 0 . \tag{1.4}$$

The AR (p) model will be stationary if the roots of the characteristic equation

$$y^p - \phi_1 y^{p-1} - ..... - \phi_p = 0 \tag{1.5}$$

are less that one in absolute value; i.e. if they lie within a circle of radius 1. To define a stationary AR (p), autocorrelation function $\rho(k)$, as an alternative

procedure, is used. To do this we multiply both sides of process by y_{t-r}. We take expectations of all terms and divide by $Var(y_t)$, which is assumed finite. This gives

$$\rho(r) = \phi_1\rho(r-1) + \phi_2\rho(r-2) + \ldots\ldots\ldots + \phi_p\rho(r-p). \tag{1.6}$$

Substituting r=1,2,3..p and noting $\rho(r) = \rho(-r)$, equations determining the p parameters, $\phi_1, \phi_2, \ldots\ldots\ldots\ldots, \phi_p$ called the " Yule- Walker " equations, are obtained.

In the case of AR (1) process, model satisfies the following difference equation:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t . \tag{1.7}$$

From the analysis of first – order difference equations that if $|\phi_1| \geq 1$, the consequences of the ε_t's for y accumulate rather than die out over time. Thus there does not exist a covariance – stationary process for y_t with finite variance that satisfies AR (1). In the case when $|\phi_1| < 1$, there is a covariance – stationary process for y_t. When $\phi_1 = 0$, all y_t values become random variables, and when $\phi_1 = 1$, the difference between y_t values is random.

1.1.2. Unit Root

If the autoregressive process has roots outside the unit circle, which is equivalent to $|\phi_1| \geq 1$ in the case of an AR (1) process without intercept, the process

becomes ”nonstationary”.

For the model;

$$y_t = \phi_1 y_{t-1} + \varepsilon_t ; \varepsilon_t \sim \text{iid} ; \text{t} \geq 1$$

without making any assumption about the value of ϕ_1, y_t can be expressed in terms of the fixed and known presample value y_0 by repeated substitution. Thus;

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \tag{1.8}$$

$$= \phi_1 (\phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$= \phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$$

$$= \phi_1^3 y_{t-3} + \phi_1^2 \varepsilon_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$$

$$\vdots$$

$$= \phi_1^t y_0 + \phi_1^{t-1} \varepsilon_1 + \ldots\ldots + \phi_1 \varepsilon_{t-1} + \varepsilon_t .$$

So the mean, variance of y_t and the covariances between y_t and y_{t-p} can be calculated by means of this representation. Let $\varepsilon_t \sim$ iid (0, σ^2)

$$E(y_t) = \phi_1^t y_0 ; t \geq 1,$$

$$V(y_t) = \sigma^2 (1 + \phi_1^2 + \phi_1^4 + ... + \phi_1^{2(t-1)}); t \geq 1, \tag{1.9}$$

$$Cov(y_t, y_{t-p}) = \phi_1^p Var(y_{t-p}); 1 \leq p \leq t-1.$$

Since all these quantities depend on t, this process is nonstationary, when $|\phi_1| \geq 1$.

However if $|\phi_1| < 1$, as $t \to \infty$;

$$E(y_t) \to 0,$$

$$Var(y_t) \to \frac{\sigma^2}{1-\phi_1^2}, \tag{1.10}$$

$$\frac{Cov(y_t, y_{t-p})}{Var(y_t)} \to \phi_1^p.$$

These are the expected values for a stationary AR (1) process.

Thus when $|\phi_1| \geq 1$, the sum, used to define $Var(y_t)$ diverges. So the value of the process exceeds the mean too much and the mean value becomes large unless $y_0 = 0$. On the other hand when $\phi_1 = 1$, this mean is constant at y_0.

The special case of a nonstationary AR (1) process without intercept where $\phi_1 = 1$ is usually called the random walk process.

$$y_t = y_{t-1} + \varepsilon_t \,;\; \varepsilon_t \sim \text{iid} \,;\; t \geq 1. \tag{1.11}$$

$$y_t = y_0 + \varepsilon_1 + \varepsilon_2 + \ldots\ldots\ldots + \varepsilon_t .$$

Here every error term from the start of the process has an effect on the current value of y_t; that is, the errors are persistent. It is the contrary case of a stationary AR (1) process, since under this condition, the impact disappears as we deal with past values of y_t.

There is one big statistical problem in unit root case of the model (1.7). This problem is that the distribution of the least square estimate of the autoregressive parameter ϕ_1 has a non-standard distribution (not the usual normal, t or F) since $H_0 : \phi_1 = 1$ is tested. Moreover it is often asymptotically biased, that is, its limiting distribution is not centered on the true parameter value. Since

$$\tilde{\phi}_1 - \phi_1 = \frac{\sum_{t=1}^{n} y_{t-1}\varepsilon_t}{\sum_{t=1}^{n} y_{t-1}^2} ; t \geq 1 \tag{1.12}$$

where $\tilde{\phi}_1$ is the least square estimate of ϕ_1.

The limiting distribution of ϕ_1 under stationary condition is that $n^{-1}\sum y_{t-1}^2 \xrightarrow{p} Q$ and $\sqrt{n}n^{-1}\sum y_{t-1}\varepsilon_t \xrightarrow{L} N(0, \sigma^2 Q)$ by central limit theorem where $Q = \sigma^2(1-\phi_1^2)^{-1}$. In other words

$$\sqrt{n}(\tilde{\phi}_1 - \phi_1) \xrightarrow{L} N(0,(1-\phi_1^2)) \ ; |\phi_1| < 1 . \tag{1.13}$$

Then if the true value of ϕ_1 is unity, model (1.7) describes a random walk as mentioned before. In order to test the null hypothesis that $\phi_1 = 1$, we need to know the limiting distribution of $\tilde{\phi}_1$ under H_0, in which case y_t is nonstationary.

Let $\phi_1 = 1$, $y_0 = 0$ and ε_t ~ iid $N(0,\sigma^2)$. So y_t is expressed such that $y_t = \varepsilon_1 + \varepsilon_2 + ... + \varepsilon_{t-1}$ and $y_t \sim N(0,\sigma^2 t)$. Moreover since the numerator part of equation (1.12) can be generated as;

$$y_t^2 = (y_{t-1} + \varepsilon_t)^2 = y_{t-1}^2 + \varepsilon_t^2 + 2y_{t-1}\varepsilon_t ,$$

$$y_{t-1}\varepsilon_t = \frac{1}{2}\left\{y_t^2 - y_{t-1}^2 - \varepsilon_t^2\right\} \tag{1.14}$$

when $t = 1,2,3,.....,n$;

$$\sum_{t=1}^{n} y_{t-1}\varepsilon_t = \frac{1}{2}\left\{y_n^2 - y_0^2\right\} - \frac{1}{2}\sum_{t=1}^{n}\varepsilon_t^2 .$$

As $y_0 = 0$;

$$\left(\frac{1}{\sigma^2 n}\right)\sum_{t=1}^{n} y_{t-1}\varepsilon_t = \left(\frac{1}{2}\right)\left(\frac{y_n}{\sigma\sqrt{n}}\right)^2 - \left(\frac{1}{2\sigma^2}\right)\left(\frac{1}{n}\right)\sum_{t=1}^{n}\varepsilon_t^2 . \tag{1.15}$$

Since $\frac{y_n}{\sigma\sqrt{n}} \sim N(0,1)$;

$$\left(\frac{y_n}{\sigma\sqrt{n}}\right)^2 \sim \chi^2_{(1)}, \tag{1.16}$$

$$\left(\frac{1}{n}\right)\sum_{t=1}^{n}\varepsilon_t^2 \xrightarrow{p} \sigma^2, \tag{1.17}$$

$$\left(\frac{1}{\sigma^2 n}\right)\sum_{t=1}^{n}\varepsilon_t^2 \longrightarrow 1. \tag{1.18}$$

Thus the numerator of equation (1.12) is not normal anymore; instead

$$\left(\frac{1}{\sigma^2 n}\right)\sum_{t=1}^{n} y_{t-1}\varepsilon_t \xrightarrow{L} \frac{1}{2}(X-1) \tag{1.19}$$

where $X \sim \chi^2_{(1)}$. On the other hand the denominator of equality (1.12) has a non-standard limiting distribution too. But firstly this quantity has to be divided by n^2 in order to construct a random variable having a convergent distribution.

$$n^{-2}\sum_{t=1}^{n} y_{t-1}^2 \xrightarrow{L} \int_0^1 [W(r)]^2\, dr \tag{1.20}$$

where W(r) is a "Brownian Motion" or " Wiener "process.

A Wiener process is like a continuous random walk defined on the interval $[0,1]$, but has unbounded variation in spite of its continuity. In general, a continuous process W (t), t$\geq$0, is a Wiener process if

1.For all t$\geq$0, E$[W(t)]$ =0,

2.For all fixed t$\geq$0, W (t) is normally distributed, non – degenerate and dates $0 \leq t_1 < t_2 < ..... < t_k \leq 1$, the changes $[W(t_2) - W(t_1)], ..., [W(t_k) - W(t_{k-1})]$ are independent multivariate Gaussian (Normal) with $[W(s) - W(t)] \sim N(0,\ s-t)$,

3. $\Pr\{W(0) = 0\} = 1$.

For that reason, in model (1.11), if we consider the change between y_{t-1} and y_t, and ε_t as the sum of two independent Gaussian variables, ε_t can be written as $\varepsilon_t = e_{1t} + e_{2t}$ where $e_{it} \sim$ iid $N\left(0, \frac{1}{2}\right)$. So if $y_{t-\left(\frac{1}{2}\right)} - y_{t-1} = e_{1t}$ and $y_t - y_{t-\left(\frac{1}{2}\right)} = e_{2t}$;

$$y_t - y_{t-1} = e_{1t} + e_{2t} \sim \text{iid } N(0,1)\ . \tag{1.21}$$

Since the process is defined also at the non - integer dates, we can divide the change between t -1 and t into T separate subperiods and can recognize model (1.11) like the following form:

$$y_t - y_{t-1} = e_{1t} + e_{2t} + ... + e_{Tt} \tag{1.22}$$

with $e_{it} \sim$ iid $N\left(0, \frac{1}{T}\right)$. Therefore it is concluded that in addition to not lose any property of random walk process in equation (1.22), we also succeed to redefine model (1.11) as a Wiener process with unit variance (called standard Brownian Motion) when $T \to \infty$.

Furthermore for the random walk model when $\varepsilon_t \sim$ iid $(0, \sigma^2)$, the following limiting distributions can be derived from the use of Wiener process too.

$$n^{-1/2} \sum_{t=1}^{n} \varepsilon_t \xrightarrow{L} \sigma W(1) . \tag{1.23}$$

$$n^{-1} \sum_{t=1}^{n} y_{t-1} \varepsilon_t \xrightarrow{L} \left(\frac{1}{2}\right) \sigma^2 \left\{ [W(1)]^2 - 1 \right\}. \tag{1.24}$$

$$n^{-3/2} \sum_{t=1}^{n} \varepsilon_{t-1} \xrightarrow{L} \sigma \int_0^1 W(r) dr . \tag{1.25}$$

$$n^{-2} \sum_{t=1}^{n} \varepsilon_{t-1}^2 \xrightarrow{L} \sigma^2 \int_0^1 [W(r)]^2 dr . \tag{1.26}$$

Moreover although W (t) is continuous in t, it cannot be differentiated in a usual way. Because the direction of change at t seems to be completely different from that at t + Δ and also it does not have any analytic expression for this distribution. Thus its percentiles are calculated by using simulations.

However there is one important thing to be emphasized about the "unit

root " distribution of $\tilde{\phi}_1$. Since the numerator is based on the $\chi^2_{(1)}$ random variable, and since χ^2 is not a symmetric distribution, the distribution of $\tilde{\phi}_1$ also is not symmetric. In fact, its distribution is left skewed and has asymptotically downwards bias. For this reason even if the true model is a random walk, estimated AR (1) parameter $\tilde{\phi}_1$ is less than one.

This property is very helpful especially while constructing tests of the null hypothesis of a unit root. Tests based on these statistics are called "Dickey – Fuller "tests.

1.2. Unit Root Tests

As the detection of stationarity in the series is very important in univariate time series models, a large number of tests have been developed to control this property, called unit root tests. In this part the most well – known studies on this subject are considered and summarized.

1.2.1. Dickey – Fuller (DF) Test

The Dickey – Fuller test for unit roots consists of comparing the t – value of ϕ_1 against the tabulated critical value under the assumption that ε_t's are white noise errors and least square estimate of ϕ_1 is used.

$$t = \frac{(\tilde{\phi}_1 - 1)}{\tilde{\sigma}_{\tilde{\phi}_1}} = \frac{(\tilde{\phi}_1 - 1)}{\left\{ s_n^2 / \sum_{t=1}^{n} y_{t-1}^2 \right\}^{1/2}} \tag{1.27}$$

where $\tilde{\sigma}_{\tilde{\phi}_1}$ is the usual LS standard error for the estimated coefficient and s_n^2 indicates the LS estimate of the residual variance.

Although the t – statistic is calculated in the usual way, it does not have a limiting Gaussian distribution when $\phi_1 = 1$. It is composed of the ratio of two Wiener process as mentioned in the previous part (ratio of equations 1. 20 and 1. 24).

$$t = \frac{n^{-1}\sum_{t=1}^{n} y_{t-1}\varepsilon_t}{\left\{n^{-2}\sum_{t=1}^{n} y_{t-1}^2\right\}^{1/2}\left(s_n^2\right)^{1/2}} \tag{1.28}$$

where $s_n^2 = \sum_{t=1}^{n}(y_t - \tilde{\phi}_1 y_{t-1})^2/(n-1)$. So as $n \to \infty$;

$$t \xrightarrow{L} \frac{\left(\frac{1}{2}\right)\sigma^2\left\{[W(1)]^2 - 1\right\}}{\left\{\sigma^2\int_0^1 [W(r)]^2\,dr\right\}^{1/2}\left\{\sigma^2\right\}^{1/2}} = \frac{\left(\frac{1}{2}\right)\left\{[W(1)]^2 - 1\right\}}{\left\{\int_0^1 [W(r)]^2\,dr\right\}^{1/2}}. \tag{1.29}$$

Dickey – Fuller tables for this t distribution for different sample sizes n are found by Monte Carlo simulation.

In testing procedure of DF test, firstly y_{t-1} is subtracted from both sides of the equation

$$y_t = \phi_1 y_{t-1} + \varepsilon_t ,$$

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t \tag{1.30}$$

in which $\Delta y_t = y_t - y_{t-1}$ and $\gamma = \phi_1 - 1$, which is zero under the null hypothesis that $\phi_1 = 1$. Then t value of γ is compared against the critical values. If this t value is too large and positive, the null hypothesis is accepted, and it is concluded that the series contains a unit root. When we add a constant term in the model, the estimated coefficient on y_{t-1} must be farther from unity in order to reject the null hypothesis of a unit root. Because the asymptotic distribution of the new model becomes more strongly skewed than initial model (1.11). For this reason the critical values to compare calculated statistics change as stated in Table 1. I and Table 1. II – Case 2.

But the DF test used for AR (1) process causes autocorrelated error terms if AR (p) process is analysed. Thus DF distribution cannot be valid since it is based on the assumption of the absence of serial correlation in error terms. In this case "augmented" Dickey – Fuller (ADF) test is applied. The test statistic used to this augmented equation has asymptotically the same distribution as the DF test statistic and hence the same significance tables can be used.

1.2.2. Phillips – Perron (PP) Tests

An alternative approach to testing for unit root has been proposed by Phillips (1987) and Phillips and Perron (1988). Instead of the Dickey – Fuller assumptions of independence and homogeneity, the Phillips – Perron test allows the distribution to be weakly independent and heterogeneously distributed. In fact

it is the alternative of ADF test statistic. The ADF test uses the addition of the lagged difference terms of the regressors when the error terms have serial correlation. On the other hand PP test does not use lagged difference terms under correlation and heterogeneity.

The PP test is done by a two - step procedure. At first step DF statistic from an ARIMA (1,0,0) model is calculated and then it is adjusted by the autocovariance of the errors. For AR (1) model without intercept, the test statistic is defined as follows:

$$Z_{\phi} \equiv n(\tilde{\phi}_n - 1) - \frac{1}{2}\left\{n^2 \tilde{\sigma}^2_{\tilde{\phi}_n} \div s_n^2\right\}\left(\tilde{\lambda}_n^2 - \tilde{\gamma}_{0,n}\right) \tag{1.31}$$

and

$$Z_t = \left(\tilde{\gamma}_{0,n} / \tilde{\lambda}_n^2\right)^{1/2}\left(\tilde{\phi}_n - 1\right) / \tilde{\sigma}_{\tilde{\phi}_n} - \frac{1}{2}\left(\tilde{\lambda}_{nT}^2 - \tilde{\gamma}_{0,n}\right)\left(\frac{1}{\tilde{\lambda}_n}\right)\left(n\hat{\sigma}_{\tilde{\phi}_n} \div s_n\right) \tag{1.32}$$

where

$$\tilde{\gamma}_{j,n} = n^{-1}\sum_{t=1}^{n} \tilde{\varepsilon}_t \tilde{\varepsilon}_{t-j}\ ; \tag{1.33}$$

$\tilde{\varepsilon}_t$ = LS sample residual from the estimated regression.

$$\tilde{\lambda}_n^2 = \tilde{\gamma}_{0,n} + 2\sum_{j=1}^{q}\left[1 - j/(q+1)\right]\tilde{\gamma}_{j,n}\ , \tag{1.34}$$

$$s_n^2 = \frac{\sum_{t=1}^{n} \tilde{\varepsilon}_t^{\,2}}{n-k} \tag{1.35}$$

where k is the number of parameters in estimated regression. $\tilde{\sigma}_{\tilde{\phi}_n}$ is the LS standard error for $\tilde{\phi}_1$. Z_ϕ is the Phillips – Perron ϕ_1 test statistic and Z_t is the Phillips – Perron t – test statistic.

In case of limiting distribution, both ADF and PP tests have the same asymptotic distribution. So when estimated regression is

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

and the true process is

$$y_t = y_{t-1} + \varepsilon_t,$$

the LS t – test of ϕ_1=1 is

$$t_n = \frac{\tilde{\phi}_1 - 1}{\tilde{\sigma}_{\tilde{\phi}_1}} \ . \tag{1.36}$$

Thus the limiting distribution of t_n, Z_ϕ and Z_t are defined as follows:

$$t_n \xrightarrow{L} \left(\frac{\lambda^2}{\gamma_0}\right)^{1/2} * \left[\frac{\frac{1}{2}\left\{[W(1)]^2 - 1\right\}}{\left\{\int [W(r)]^2 dr\right\}^{1/2}} + \frac{\frac{1}{2}\left(\lambda^2 - \gamma_0\right)}{\lambda^2 \left\{\int [W(r)]^2 dr\right\}^{1/2}}\right], \tag{1.37}$$

$$Z_\phi \xrightarrow{L} \frac{\left(\frac{1}{2}\right)\left\{[W(1)]^2 - 1\right\}}{\int_0^1 [W(r)]^2 dr}, \tag{1.38}$$

$$Z_t \equiv \left(\frac{\gamma_0}{\lambda^2}\right)^{1/2} t_n - \left\{\frac{1}{2}(\lambda^2 - \gamma_0)/\lambda\right\} * \left\{n\tilde{\sigma}_{\tilde{\phi}_n} \div s_n\right\} \xrightarrow{L} \frac{\frac{1}{2}\left\{[W(1)]^2 - 1\right\}}{\left\{\int [W(r)]^2 dr\right\}^{1/2}} \tag{1.39}$$

where $\lambda \equiv \sigma\Psi(1)$, $\Psi(1) = \sum_{j=0}^{\infty} \Psi_j$, $\varepsilon_t = \Psi(L)u_t$, and $\sum_{j=0}^{\infty} j\left|\Psi_j\right| < \infty$, u_t is iid with mean zero, variance σ^2, and has finite fourth moment.

Accordingly $\Psi(L)$ denotes an infinite – order polynomial in the lag operator and Ψ_j terms in $\Psi(L)$ formula explain the moving average coefficients.

$$\Psi(L) = \Psi_0 + \Psi_1 L + \Psi_2 L^2 + \ldots\ldots\ldots \tag{1.40}$$

$$= \left[1 - \phi_1 L - \phi_2 L^2 - \ldots\ldots\ldots\ldots - \phi_p L^p\right]^{-1}$$

for general AR (p) process. So in our case, it is equal to

$$\Psi(L) = [1 - \phi_1 L]^{-1}. \tag{1.41}$$

As it is seen from the limiting distribution of Z_t, the final term in (1.37) is a correction for serial correlation.

If u_t is serially uncorrelated, then $\Psi_0 = 1$ and $\Psi_j = 0$ for $j = 1,2\ldots$ Thus $\lambda^2 = \sigma^2 [\Psi(1)]^2 = \sigma^2$ and $\gamma_0 = E(\varepsilon_t^2) = \sigma^2$. Therefore

The equation (1.37) includes a special case when ε_t is serially uncorrelated.

By comparing the Z_ϕ with DF critical value (Table 1. I), if it is greater than tabulated number, then the null hypothesis of a unit root is not rejected, otherwise H_0 is rejected, in which case the time series is stationary.

In the same way if we prefer Z_t statistic to test our hypothesis, in this case we compare calculated value with the tabulated ones given in Table 1. II. Similar to the previous test, large value is an indicator of nonstationarity.

These decision rules and test statistics are used also for AR (1) model with constant term except a change in tabulated number (Case - 2 part is used in Table 1. I and Table 1. II).

Both these two unit root tests have very low power based on the Monte Carlo simulations. In fact, this problem was first presented by Schwert (1989). Later D.N. DeJong, J.C. Nankervis, N.E. Savant and C.H. Whiteman (1992) pointed this result. Now it is known that unit root tests do not have power to distinguish between a unit root and near unit root process. Consequently these tests confirm seldom that a series contains a unit root (Gujarati 2003, p. 819).

There are also alternative tests having stationarity as a null hypothesis. The most widely used one is the KPSS test suggested by Kwiatkowski, Phillips, Schmidt, and Shin (1992). However there is no evidence of superiority of this test

to the ADF test.

Besides all of the studies, Karim M. Abadir (1995) made some works about the exact explicit formula for the density and distribution function of the t – ratio under this special case. As the distribution of t is asymmetric, the form of the ratio alters with respect to the sign and value of t. Abadir applied the quadratic forms in the normal variate y_t to define each case explicitly.

$$Q = \frac{\sqrt{2}}{n\sigma^2}\left(\tilde{\phi}_1 - \phi_1\right)\sum_t y_{t-1}^2 , \tag{1.42}$$

$$S = \frac{2}{n^2\sigma^2}\sum_t y_{t-1}^2 \tag{1.43}$$

where $|\phi_1| = 1$.

As $\tilde{\sigma}$ converges to σ; $t = \frac{Q}{S}$. Thus setting $R \equiv \frac{Q}{\phi_1}$, a more general statistic can be constructed as the following way:

$$Z \equiv \frac{t}{\phi_1} \xrightarrow{d} \frac{R}{\sqrt{S}} \ , \ |\phi_1| = 1 . \tag{1.44}$$

He calculated pdf of Z by first deriving the joint limiting distribution of Z and S and then using the transformation theorem for pdf's and finally by inverse Laplace transformation and the integral representation of the parabolic cylinder under different sign of Z.

When z <0;

$$pdf(z) = \left(\sqrt{\frac{2}{\pi x}}\right) e^{-x^2/4} \sum_j \binom{j-\frac{1}{2}}{j} \tag{1.45}$$

$$\text{x} \left[\frac{\left(\sqrt{8x}\right) j^j e^{-2jx^2(2j+1)}}{\left(j+\frac{1}{2}\right)^{j-\frac{1}{2}}} + \sum_l \binom{j}{l} (-x)^l \sum_k \frac{x^k}{k!} \left(\frac{x^2}{2} - \frac{1}{2} - l - k \right) \right] \text{x}\, \Gamma\left(\frac{1}{4} - \frac{l}{2} + \frac{k}{2}; (2j+1)^2 x^2 \right).$$

When $z \geq 0$;

$$pdf(z) = \int_0^\infty \left[\frac{2}{z} \csc h(2zs) \right]^{1/2} \text{x} \exp\left[\frac{s^2}{2} - \frac{1}{4} \left(s + z \coth(2zs)\right)^2 \right] \tag{1.46}$$

$$\text{x} \left[\left(s^2 - \frac{1}{2} \right) D_{-1/2}\left(s + z\coth(2zs)\right) + \left(z\coth(2zs) - s\right) D_{1/2}\left(s + z\cot(2zs)\right) \right] \frac{ds}{\pi s}$$

where D_ϑ (z) is the parabolic cylinder function.

For integer ϑ, this function can be written as

$$D_{-n-1}(z) = \left(\sqrt{2\pi}\right) e^{z^2/4} \int \ldots\ldots\ldots \int \Phi(-z) \left[d(-z) \right]^n \tag{1.47}$$

where $n \in \mathrm{Z}$.

By using these results he gained numerical efficiency of calculating the pdf and cdf, which are mainly used to compute quantiles of the distribution - based on the Monte Carlo simulations -, relative to the other existing methods. Although the power of the t – ratio was not calculated in this study, he compared asymptotic power of unit root tests in his another paper published in 1993. In this paper basically, approximate and exact Wald and Lagrange Multiplier test statistic and White's (1958) Simple Normalized Autocorrelation Coefficient are inspected. Before presenting the comparisons, a brief summary about these tests can be useful.

1.2.3. Asymptotically Equivalent Test Procedure

1.2.3.1 Wald (*W*) Test

The aim of this test is to test the validity of a set of g independent linear restrictions, written as $HQ = h$, where Q is a (n x 1) vector of parameters. Any necessary conditions for the model are assumed so that

$$\sqrt{n}\left(\tilde{Q} - Q\right) \xrightarrow{d} Q^{-1} n \sim N(0, Q^{-1}) \tag{1.48}$$

can be valid where $Q = \lim n^{-1} I(Q)$.

$$\sqrt{n} H\left(\hat{Q} - Q\right) \xrightarrow{d} HQ^{-1} n \sim N(0, HQ^{-1} H') \tag{1.49}$$

under $H_0 : HQ = h$.

$$\sqrt{n}\left(H\hat{Q}-h\right)\xrightarrow{d} HQ^{-1}n \sim N(0, HQ^{-1}H'). \tag{1.50}$$

$$n\left(H\hat{Q}-h\right)'\left[HQ^{-1}H'\right]^{-1}\left(H\hat{Q}-h\right)\xrightarrow{d}\chi^2 \sim \chi_g^2. \tag{1.51}$$

$$\left(H\hat{Q}-h\right)'\left[H\left(n^{-1}Q^{-1}\right)H'\right]^{-1}\left(H\hat{Q}-h\right)\xrightarrow{d}\chi^2 \sim \chi_g^2. \tag{1.52}$$

In practice $n^{-1}Q^{-1}$ is replaced by $I^{-1}(\hat{Q})$. Thus

$$W = \left(H\hat{Q}-h\right)'\left[HI^{-1}H'\right]^{-1}\left(H\hat{Q}-h\right). \tag{1.53}$$

So the Wald statistic under H_0, in large samples, has a chi – square distribution with degree of freedom equal to the number of restrictions (i.e. the number of equations in $H\hat{Q}-h=0$).

The most important property of this test is that it is completely based on the unrestricted ML estimates of parameter estimates and $\hat{Q}$ is the MLE (maximum likelihood estimator) of Q.

If $HQ-h$ is a single restriction like $H_0 : Q = Q_0$ versus $H_1 : Q \neq Q_0$, W will be

$$W = \left[\left(\hat{Q}-Q_0\right)-0\right]\left(Var\left[\left(\hat{Q}-Q_0\right)-0\right]\right)^{-1}\left[\left(\hat{Q}-Q_0\right)-0\right] \tag{1.54}$$

$$= \frac{\left(\hat{Q}-Q_0\right)^2}{Var\left(\hat{Q}\right)}$$

$$= Z^2$$

and has a chi – square distribution with one degree of freedom.

1.2.3.2. Lagrange Multiplier (*LM*) Test

It is also called efficient score or score test. The test procedure is based on the restricted model instead of the unrestricted one. When we maximize the log–likelihood subject to the set of constraints $HQ - h = 0$, Lagrangean function is written as:

$$InL^*(Q) = InL(Q) + \lambda'(HQ - h) \tag{1.55}$$

where λ is the vector of Lagrange multiplier.

So the solution of the constrained maximization problem is the roots of

$$\frac{\partial InL^*}{\partial Q} = \frac{\partial InL(Q)}{\partial Q} + H'\lambda = 0 \tag{1.56}$$

and

$$\frac{\partial InL^*}{\partial \lambda} = HQ - h = 0\,. \tag{1.57}$$

If the restrictions are valid, the difference between maximized value of the likelihood function under restricted and unrestricted model is not significant. So in the *LM* test, it is simply verified whether $H_0 : \lambda = 0$ or not. There is another way to proceed with the *LM* test. At the restricted maximum, the derivatives of the log – likelihood function are

$$\frac{\partial InL(\hat{Q})}{\partial \hat{Q}} = -\hat{C}'\lambda = \hat{g}_R \tag{1.58}$$

where $C = \left[\frac{\partial H\hat{Q}}{\partial \hat{Q}}\right]$.

If restriction is true, $\hat{g}_R$, restricted parameter vector, must be approximately or exactly equal to zero at least within the range of sampling variability.

Thereby, the simple form of the *LM* test is

$$LM = \left(\frac{\partial InL(\hat{Q}_R)}{\partial \hat{Q}_R}\right)' \left[I(\hat{Q}_R)\right]^{-1} \left(\frac{\partial InL(\hat{Q}_R)}{\partial \hat{Q}_R}\right). \tag{1.59}$$

Under the null hypothesis, *LM* has a limiting chi – square distribution with degrees of freedom equal to the number of restrictions and whole terms are derived from the restricted estimates.

1.2.3.3. White's Simple Normalized Autocorrelation Coefficient (*NAC*)

Basic form of this test is that

$$NAC \equiv \left(\hat{\phi}_1 - \phi_1\right) g(n) \tag{1.60}$$

where $g(n) = \left[I(\phi_1)\right]^{1/2}$.

In fact we can calculate the square root of the matrix $\sum$ if it is p x p positive and symmetric matrix. Then there exists an orthogonal matrix Γ with positive diagonal elements such that $\sum = \Gamma D_\lambda \Gamma', \Gamma\Gamma' = \Gamma'\Gamma = T$ where T is an identity matrix and D_λ is a diagonal matrix, diag ($\lambda_1, \lambda_2, ..., \lambda_p$) where $\lambda_i > 0$,

$i = 1,2,..., p$. So we can write $\sum = \Gamma D_{\lambda^{1/2}} \Gamma' \Gamma D_{\lambda^{1/2}} \Gamma' = \sum^{1/2} \sum^{1/2}$ where $D_{\lambda^{1/2}} = diag(\lambda_1^{1/2}, ..., \lambda_p^{1/2})$. Thus each $\sum^{1/2}$ matrix indicates the square root of the matrix $\sum$. While calculating $g(n) = \left[I(\phi_1)\right]^{1/2}$ term, we proceed same way. Because $\left[I(\phi_1)\right]^{1/2}$ matrix indicates the square root of the matrix $\left[I(\phi_1)\right]$.

However, g (n) is taken as the following forms as they are asymptotically equivalent to $\left[I(\alpha)\right]^{1/2}$ ($I(\phi_1)$ is the information matrix).

$$g(n) = \sqrt{\frac{n}{1-\phi_1^{\,2}}} \text{ for } \left|\phi_1\right| < 1. \tag{1.61}$$

$$g(n) = \frac{n}{\sqrt{2}} \text{ for } |\phi_1| = 1. \tag{1.62}$$

$$g(n) = \frac{|\phi_1|^n}{\phi_1^{\,2} - 1} \text{ for } |\phi_1| > 1. \tag{1.63}$$

Under $|\phi_1| < 1$ and $|\phi_1| > 1$ conditions, in order to find the distribution of standardized autocorrelation coefficient, first of all we consider the distribution of $\hat{\phi}_1 - \phi_1$ rather than that of ϕ_1 alone. We have:

$$\hat{\phi}_1 - \phi_1 = \frac{\sum_{t=1}^{n} y_t y_{t-1}}{\sum_{t=1}^{n} y_{t-1}^2} - \phi_1 \tag{1.64}$$

$$= \frac{\sum_{t=1}^{n} y_t y_{t-1} - \phi_1 \sum_{t=1}^{n} y_{t-1}^2}{\sum_{t=1}^{n} y_{t-1}^2} = \frac{Y'AY}{Y'BY}$$

where A and B are the n x n matrices

$$A = -\frac{1}{2}\begin{bmatrix} 2\phi_1 & -1 & 0 & & & & \\ -1 & 2\phi_1 & -1 & & & & \\ 0 & -1 & 2\phi_1 & & & & \\ & & & \cdots\cdots & & & \\ & & & & -1 & 2\phi_1 & -1 \\ & & & & 0 & -1 & 0 \end{bmatrix}, \tag{1.65}$$

$$B = \begin{bmatrix} 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \\ & & & \cdots\cdots & & \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 0 \end{bmatrix}. \tag{1.66}$$

Since the limiting distribution of *NAC* is the distribution of the ratio of two random variables, White proposed the following:

Let X and V be two random variables, P (V >0) = 1. If $Z = X / V$; then

$$W = W_Z = X - zV . \tag{1.67}$$

Thus we have

$$P(Z < z) = P(X / V < z) = P(X - zV < 0) = P(W_Z < 0) . \tag{1.68}$$

Then we need the distribution of W to compute the distribution of Z. Hence, we try to find it by means of moment generating functions. If the joint moment generating function of $Y'AY / g(n)$ and $Y'BY / g^2(n)$ is formulated as;

$$M(U, R) = E\left[\exp Y'AYU / g(n) + Y'BYR / g^2(n)\right] \tag{1.69}$$

$$= m\left[U / g(n), R / g^2(n)\right]$$

and

$$\lim M(U,R) = \exp\left(R + U^2/2\right) \text{ for } |\phi_1| < 1, \tag{1.70}$$

we set $W = Y'AY/g - zY'BY/g^2$ where $g = g(n)$.

Therefore,

$$m(w) = E\left[\exp\{Ww\}\right] \tag{1.71}$$

$$= E\left[\exp\{X - zV\}w\right] = E\left[\exp\{Xw - Vzw\}\right],$$

$$\lim m(w) = \exp(-zw + w^2/2) \text{ for } |\phi_1| < 1. \tag{1.72}$$

From the result, it is obvious that w is a random variable, which is normally distributed with mean – z and variance 1. So;

$$\lim P(W < 0) = (2\pi)^{-1/2} \int_{-\infty}^{0} \exp(-\{t+z\}^2/2)dt \tag{1.73}$$

$$= (2\pi)^{-1/2} \int_{-\infty}^{z} \exp(-t^2/2)dt$$

$$= \lim P(g(n)(\hat{\phi}_1 - \phi_1) < z).$$

For $|\phi_1| > 1$, we obtain limiting distribution again by moment generating function but using a different variable R= a X- b V where X and V are independent

chi – square variables with one degree of freedom. At the end we reach a Cauchy distribution.

As a result $g(n)(\hat{\phi}_1 - \phi_1)$ is asymptotically normal with mean 0 and variance 1 for $|\phi_1| < 1$. Here we do not make any assumption like the normality distributed error terms. The single condition which we put is that all moments of the ε_t's are finite.

When $|\phi_1| > 1$ is tested, $g(n)(\hat{\phi}_1 - \phi_1)$ has the Cauchy distribution (White 1958). But for $\phi_1 = 1$, on the other hand, the limiting distribution of NAC is the same as the distribution of the functional

$$G[y(.)] = \frac{\int_0^1 y(t)dy(t)}{\int_0^1 y^2(t)dt} = \frac{\frac{1}{2}\chi^2(1) - \frac{1}{2}}{\int_0^1 y^2(t)dt} \tag{1.74}$$

on the Wiener process, independent of the distribution of the error terms based on the Donsker's Theorem. The theorem says that the limiting distribution of a function of a sequence of independent random variables, with suitable restrictions on these random variables, depends only on the form of the function and is the same as the distribution of a related functional in a stochastic process.

Now we can return to the comparison of these test statistics in Abadir's paper (1993). Abadir, in this study, generates the data from AR (1) model without intercept where $\varepsilon_t \sim NID(0, \sigma^2)$ and y_0 is a constant. The model to be estimated and tested is

$$y_t = \hat{\phi}_1 y_{t-1} + \varepsilon_t \tag{1.75}$$

by means of the ML estimators;

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{n} y_t y_{t-1}}{\sum_{t=1}^{n} y_{t-1}^2}, \tag{1.76}$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \hat{\phi}_1 y_{t-1}\right)^2}{n}}. \tag{1.77}$$

For the general case, the null hypothesis examined, has the following form:

$$H_0 : \phi_1 = \tilde{\phi}_1. \tag{1.78}$$

Thus under model (1.75) and hypothesis (1.78);

$$AWR \equiv \hat{t}_{\phi_1} \equiv (\hat{\phi}_1 - \tilde{\phi}_1)\left(-\hat{H}_{11}\right)^{1/2} \equiv \hat{R} / \left(\hat{S}\right)^{1/2}, \tag{1.79}$$

$$ALMR \equiv \tilde{t}_{\phi_1} \equiv (\hat{\phi}_1 - \tilde{\phi}_1)\left(\sum_{t}^{n} y_{t-1}^2 / \tilde{\sigma}^2\right)\left(-H_{11}\right)^{-1/2}, \tag{1.80}$$

$$\equiv (\hat{\phi}_1 - \tilde{\phi}_1)\left(\sum_{t}^{n} y_{t-1}^2 / \tilde{\sigma}^2\right)^{1/2} \equiv \tilde{R} / \left(\tilde{S}\right)^{1/2}$$

$$XWR \equiv (\hat{\phi}_1 - \tilde{\phi}_1)(I_{11})^{1/2} \equiv \hat{R} / \hat{S} \,, \tag{1.81}$$

$$NAC \equiv (\hat{\phi}_1 - \tilde{\phi}_1)\tilde{g} \equiv \hat{R} / \tilde{S} \,, \tag{1.82}$$

$$XLMR \equiv (\hat{\phi}_1 - \tilde{\phi}_1)\left(\sum_{t}^{n} y_{t-1}^2 / \tilde{\sigma}^2 \right)(I_{11})^{-1/2} \equiv \tilde{R} \,, \tag{1.83}$$

and

$$R = \left(\sum_{t=1}^{n} y_t y_{t-1} - \phi_1 \sum_{t=1}^{n} y_{t-1}^2 \right) / \left(g\sigma^2 \right) \;\; ; \;\; S = \left(\sum_{t=1}^{n} y_{t-1}^2 \right) / \left(g^2 \sigma^2 \right), \tag{1.84}$$

$$\tilde{R} = \left(\sum_{t=1}^{n} y_t y_{t-1} - \tilde{\phi}_1 \sum_{t=1}^{n} y_{t-1}^2 \right) / \left(\tilde{g}\tilde{\sigma}^2 \right) \;\; ; \;\; \tilde{S} = \left(\sum_{t=1}^{n} y_{t-1}^2 \right) / \left(\tilde{g}^2 \tilde{\sigma}^2 \right), \tag{1.85}$$

$$\hat{R} = \left(\sum_{t=1}^{n} y_t y_{t-1} - \tilde{\phi}_1 \sum_{t=1}^{n} y_{t-1}^2 \right) / \left(\hat{g}\hat{\sigma}^2 \right) \;\; ; \;\; \hat{S} = \left(\sum_{t=1}^{n} y_{t-1}^2 \right) / \left(g^2 \hat{\sigma}^2 \right), \tag{1.86}$$

where $\tilde{g}$ and $\hat{g}$ are defined by replacing ϕ_1, σ by $\tilde{\phi}_1, \tilde{\sigma}$ and $\hat{\phi}_1, \hat{\sigma}$ respectively in equations (1.61), (1.62), and (1.87);

$$g(n) = |\phi_1|^n \left[1 + \left(\phi_1^2 - 1 \right) y_0^2 / \sigma^2 \right]^{1/2} / \left(\phi_1^2 - 1 \right) \;\; ; \;\; |\phi_1| > 1 \,. \tag{1.87}$$

In our abbreviations, we denote:

$AWR \equiv$ Square root of the approximate Wald statistic,

$ALMR \equiv$ Square root of the approximate LM statistic,

$XWR \equiv$ Square root of the exact Wald statistic,

$NAC \equiv$ White's simple normalized autocorrelation coefficient,

$XLMR \equiv$ Square root of the exact LM statistic.

Firstly it is noticed that *NAC* and *XWR* have asymptotically equivalent distribution under the null hypothesis (1.78). The difference between them is that the former applies $\tilde{g}$ (the simple normalized autocorrelation coefficient) but the latter uses the estimated value of the normalization factor $I_{11}^{1/2}$. However, it is important to realize that this equivalence of the two statistics is valid only under the null.

One of the other crucial points of these tests is that while defining exact tests, Abadir worked with Fisher information matrix $I = E(-H)$; on the other hand when he defined approximate ones, he applied directly Hessian matrix H (i.e. derivative of the score). Thus the former tests are based on a stochastic normalization, as the latter are generally based on a deterministic normalization. So in order to calculate limiting distribution under the null, we apply the following theorem:

THEOREM 1: For the model and conditions set in (1.75), and as $n \to \infty$ under the null of (1.78):

a. $(\hat{R}, \tilde{R}, \hat{S}, \tilde{S}) \xrightarrow{d} (R, R, S, S)$,

b. $R \xrightarrow{d} \phi_1 (W(1)^2 - 1)/\sqrt{2}$ $\qquad |\phi_1| = 1$,

$$R \xrightarrow{d} \eta_1 \qquad |\phi_1| < 1,$$

$$R \xrightarrow{d} \eta_1 \eta_2 \qquad |\phi_1| > 1,$$

c. $$S \xrightarrow{d} 2\int_0^1 W(t)^2\, dt \qquad |\phi_1| = 1,$$

$$S \xrightarrow{d} 1 \qquad |\phi_1| < 1,$$

$$S \xrightarrow{d} \chi_{(1)}{}^2(2h) \qquad |\phi_1| > 1,$$

d. $$R/S \xrightarrow{d} \phi_1 (W(1)^2 - 1) / \left\{ \sqrt{8} \int_0^1 W(t)^2\, dt \right\}^{1/2} \qquad |\phi_1| = 1,$$

$$R/S \xrightarrow{d} \eta_1 \qquad |\phi_1| < 1,$$

$$R/S \xrightarrow{d} \eta_1 / \eta_2 \qquad |\phi_1| > 1,$$

e. $$R/\sqrt{S} \xrightarrow{d} \phi_1 (W(1)^2 - 1) / \left\{ 4 \int_0^1 W(t)^2\, dt \right\}^{1/2} \qquad |\phi_1| = 1,$$

$$R/\sqrt{S} \xrightarrow{d} \eta_1 \qquad |\phi_1| \neq 1$$

where W(t) is the standard Wiener process defined on $[0,1]$, $h \equiv y_0^2 (\phi_1^2 - 1)/(2\sigma^2)$, $(\eta_1, \eta_2)' \sim N\left((0, \sqrt{2h})', I_2\right)$, $(\eta_1 / \eta_2) \sim$ C (0,1) when y_0 =0 and C (0,1) is the standard Cauchy.

As a result it is seen that asymmetry is the main property of the unit root case.

COROLLARY 1: For the same model and conditions, as $n \to \infty$ under the null (1.75):

a. $AWR \equiv \hat{t}_{\phi_1} \xrightarrow{d} R / \sqrt{S}$,

b. $ALMR \equiv \tilde{t}_{\phi_1} \xrightarrow{d} R / \sqrt{S}$,

c. $XWR \xrightarrow{d} R / S$,

d. $NAC \xrightarrow{d} R / S$,

e. $XLMR \xrightarrow{d} R$.

So it is obvious that approximate tests and exact tests (c & d) are asymptotically equivalent of each other under the null, for any ϕ_1. But *XWR* (or *NAC*) and *XLMR,* which use Fisher's information matrix, are neither equivalent to each other nor to any of the approximate tests when $|\phi_1| \geq 1$. On the other hand when $|\phi_1| < 1$, that is the root is stable, *S* converges to one. Thus all tests become asymptotically equal.

While concentrating on the limiting distributions under the alternative

$$H_1 : |\phi_1| < 1, \tag{1.88}$$

we get very low powers. The reason is that the finite sample distributions for $|\phi_1| < 1$ look like unit root distributions and these distributions converge very slowly to their limiting normal. Consequently, distinguishing stable from unit roots becomes difficult.

In order to compare the powers of the above - mentioned tests, we apply the following theorem:

THEOREM 2: For the model and conditions set before under the alternative (1.88):

$$\Pi(AWR) > \Pi(ALMR) > \Pi(NAC) > \Pi(XWR) > \Pi\left(XLMR\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) \right) > \Pi(XLMR)$$

for moderately large samples and

$$\Pi(NAC) > \Pi(AWR) > \Pi(ALMR) > \Pi(XWR) > \Pi\left(XLMR\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) \right) > \Pi(XLMR)$$

for large samples, where $\Pi(Z) \equiv$ power function of the test statistic Z.

This theorem in fact supports also Abadir's results, which is primarily based on the Monte Carlo simulations.

From these simulations, it is shown that *AWR* and *ALMR* are both consistent and their powers tend to 1 as n tends to infinity, although the former is

always more powerful than the latter. The reason is that *AWR* uses $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$, but *ALMR* uses $\tilde{\sigma}^2 \xrightarrow{p} 2\sigma^2 / \left(1+\tilde{\phi}_1\phi_1\right) > \sigma^2$ as $\left|\tilde{\phi}_1\phi_1\right| < 1$.

When $n \to \infty$, *NAC* is asymptotically as powerful as *AWR* and *ALMR* and less powerful than both approximate tests for small samples (n<50) due to its long lower tail under the null.

On the other hand *XWR* test is weaker in finite samples even though the three test statistics above have same power asymptotically. Because if *XWR* test and *AWR* are compared, seen that the former' s distribution has a much longer lower tail than the latter' s under the null hypothesis (1.78), although both of them have same distribution under H_1 (Corollary 1).

Furthermore, *XLMR* test has the weakest power among all tests in large samples but it is consistent as $XLMR/\tilde{\phi}_1$ converges to the lower limit of the unit root distribution. But since *XLMR* tends always to the true value whether $\phi_1 = \tilde{\phi}_1$ or not, $XLMR\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right)$ is preferred; that is, we replace $\tilde{\sigma}$ by $\hat{\sigma}$. By this adjustment, the power of the new test increases with n.

Abader also made similar works for

$$H_0 : \left|\phi_1\right| = 1 \tag{1.89}$$

$$H_0 : \left|\phi_1\right| > 1 .$$

So from the results, it is found that the powers of *ALMR*, *XWR* test and *NAC* are asymptotically equal and are only marginally different in finite samples and all these tests are consistent. Moreover *XWR* test is the best of all these three

tests for all values of n. In fact, the reason is clear, because *XWR* uses $I_{11}^{1/2}$ and this value diverges faster under the explosive alternative than $\tilde{g}$ and $H_{11}^{1/2}$.

However *AWR* is not as powerful as others asymptotically any more. The main reason is that the σ value (i.e. $\hat{\sigma}$ or $\tilde{\sigma}$) used in t – ratio affects considerably the power of the test. For finite σ as n increases, $\hat{\sigma}$ is always finite but $\tilde{\sigma}$ is

bounded when $|\phi_1| < 1$ and diverges to ∞ otherwise.

Lastly *XLMR* test is still the weakest among them.

Actually these conclusions confirm the theorem below:

THEOREM 3: For the given model and conditions set, under the hypothesis testing (1.89):

$$\Pi(AWR) > \Pi(XWR) > \Pi(ALMR) > \Pi(NAC) > \Pi\left(XLMR\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) \right)$$

in small sample sizes,

$$\Pi(AWR) > \Pi(XWR) > \Pi(NAC) > \Pi(ALMR) > \Pi\left(XLMR\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) \right)$$

for moderately large sample sizes, and

$$\Pi(XWR) > \Pi(NAC) > \Pi(ALMR) > \Pi(AWR) > \Pi\left(XLMR\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} \right) \right)$$

for large sample sizes.

As a result comparing both $|\phi_1| > 1$ and $|\phi_1| < 1$ cases, commonly it is seen that Wald tests are more powerful than *LM* tests owing to the fact that the latter uses wrong information in the hypothesis. Actually, there is not a single test, which is better than all the others under every alternative. So a suitable test is chosen with respect to the given situation.

In order to get an idea about reliability of these tests when small sample size is taken, powers are fixed at a common level and then how the size of the tests varies as the sample size increases, are observed. Hence the following conclusions are drawn:

- *AWR* is more powerful than *ALMR* under both alternatives and *AWR* is the best test in small samples as expected from Theorems 2 and 3.

- *XWR* test is weaker than others under stable alternatives noting that it diverges slowly like *AWR*; but additionally its null distribution has a long lower tail similar to the *NAC*.

Again, according to the results for small sample sizes, a unique test statistic cannot be defined regardless of the direction of the alternative hypothesis.

The common properties of these tests are that maximum likelihood estimate of ϕ_1 is applicable and limiting distribution of these tests under unit root is defined by Wiener process (Brownian Motion approximation).

However in spite of the usefulness of all these tests and their limiting distributions for different values of ϕ_1, they become impractical for calculations of skewness and kurtosis of the autocorrelation coefficient ϕ_1, which have the following forms;

$$Skewness = \beta_1 = \frac{\mu_3^2}{\mu_2^3} \ , \tag{1.89}$$

$$Kurtosis = \beta_2 = \frac{\mu_4}{\mu_2^2} \tag{1.90}$$

where $\mu_2 = E(x-\mu)^2 = \mu_2' - \mu_1'^2$,

$$\mu_3 = E(x-\mu)^3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 ,$$

$$\mu_4 = E(x-\mu)^4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 .$$

μ_r' s (r=2,3...) are central moments and μ_r''s (r=1,2,.....) are non - central moments.

Knowledge of moments, when they all exist, is equivalent to knowledge of the distribution function in view of the fact that it can be possible theoretically to exhibit all the properties of the distribution in terms of the moments. However, first four moments are sufficient to define the distribution by Pearson System (Pearson 1958). Moreover by the Principle of Moments, we expect that if two distributions have the same moments up to order k, and k is large, these distributions approach identity. Thus identification of lower moments of two distributions indicates approximate equality among them.

For normal distribution, $\beta_1 = 0$, $\beta_2 = 3$, and all higher cumulants are equal to zero.

Shenton and Vinod (1996) computed exact finite sample mean, variance, skewness, and kurtosis of the ML estimator for AR (1) and random walk model without intercept term where the error terms are independent normal variates from $N(0,\sigma^2)$ and $y_0 = 0$ or $y_0 \sim N(0,\sigma^2/(1-\phi_1^2))$. Previous works done about this subject matter (Sawa 1978, Magnus 1986, Magnus and Pesaran 1991) are based on the function of eigenvalues of certain large n x n matrices and they include numerical integration of complex variables.

Even so in the study of Shenton and Vinod, new formulas, consisting of only polynomials and exponentials, are developed. Continuous time approximations based on Wiener process are not used in these calculations. The calculations of related integrals are found by numerical integration especially by q - point Gaussian quadrature methods.

The numerical integration is the approximate computation of an integral using numerical techniques. The numerical computation of an integral is sometimes called "quadrature". The most exact numerical integration techniques use Newton – Cotes formulas (also called quadrature formulas), which approximate a function that is tabulated at a sequence of regularly spaced intervals by various degree polynomials. If the functions are known analytically, instead of being tabulated at equally spaced intervals, the best numerical method of integration is called Gaussian quadrature. By picking the abscissas where the function is evaluated, and giving weights according to the selected weight function, Gaussian quadrature produces the most accurate approximations. Thus this method makes use of orthonormal polynomial such as Legendre, Hermite, Laguerre, and Chebyshev polynomials and their roots to evaluate the integral with an increased degree of precision.

We need larger point quadrature from 12 – point to 96 – point for the approximations of skewness and kurtosis. But this causes relatively less optimal

fits. We lose the accuracy since we deal with polynomials having high power and different sign.

After completing these numerical computations, Shenton and Vinod tabulated all results for different n and ϕ_1 values. Finally it is seen that for $y_0 = 0$, the MSE (as a measure of closeness to the true parameter) decreases as $|\phi_1| \to 1$ for large n and the performance of the ML estimator improves. While n increases, at first β_1 and β_2 depart considerably from their normal values; but they return to

normal ones as long as ${\phi_1}^2 < 1$.

It is also proved that for a given n, β_1 and β_2 increase with $|\phi_1|$ again under ${\phi_1}^2 < 1$ condition. All of these analyses are the evidence of unimodality (i.e. uniqueness of ϕ_1 value for which $\partial InL(\phi_1) / \partial \phi_1 = 0$) in this pattern. Same conclusions are valid for stationary initial value too. Conversely when ${\phi_1}^2 = 1$, despite of the existence of moments, they are very far from values under normality.

Consequently with the help of this study it is seen that in general pattern for both of the initial values, the standard deviation decreases and mean, skewness and kurtosis increase as $|\phi_1|$ increases for each n, or as n increases for each $|\phi_1|$ close to unity. Exact results for large n or $|\phi_1|$ very close to unity can be obtained by using a q – point quadrature with a large q like 320 – point or other specific numerical integration. If relatively lower points are chosen for integration, we obtain close approximations but we may lose some of observations as well.

In 1998, Tiku and Wong used these four moments to test unit root in an AR (1) model. They applied three – moment chi – square and four moment F approximations for the calculation of the new test statistic R that they proposed. By applying these approximations, as opposed to Gaussian quadrature methods

with a large q – points, they used all observations and they did not lose any of them while getting very close estimates of the real values. Moreover, they did not use the normality assumption of error terms in this mentioned test because of the fact that most of the econometric data does not follow the normal distribution and making such an assumption is very restrictive from a practical point of view. Hence, they dealt with the situation when error terms are again iid but have a location – scale family of symmetric distribution ranging from Cauchy to Normal. Moreover as the likelihood equations of this distribution are not intractable, they calculated necessary estimators by the method of modified maximum likelihood instead of ML. Thus, before passing on to a new testing procedure, the question of how these estimates are computed is very important. So, firstly, the methodology of MML estimators will be explained and then the new test statistic will be enunciated in detail.

CHAPTER II

THE METHOD OF MODIFIED

MAXIMUM LIKELIHOOD ESTIMATOR

This method is based on the order statistics of a random sample and is obtained by linearizing the intractable terms in likelihood equations (Tiku 1967,1968; Tiku and Suresh 1992). The reason is that for solving the equations $\partial InL / \partial \phi_1 = 0$ and $\partial InL / \partial \sigma = 0$ to find the values, which maximize *L* or *In L* (*L* is the joint pdf of $y_1, y_2, ..., y_n$), ML estimator of *L* becomes elusive since the equation does not have explicit solutions and may have multiple roots. So, certain iterative methods can be useful. For example Grid Search, Steepest Ascent, Newton – Raphson are some of the most popular iteration techniques. Although these methods give an approximate solution for the likelihood equations, they do not give exact ML estimators for unknown parameters. In fact, they have the possibility of divergence or convergence to local maximum values instead of global ones and can have very slow convergence (Barnett 1966, Tiku 1986, Puthenpura and Sinha 1986, Tiku and Suresh 1992, Vaughan 1992 (a)). Moreover if the data contains outliers, the iterations might never converge (Puthenpura and Sinha 1986).

Therefore, we need linearizations of these likelihood equations. Thus, firstly, all expressions are written in terms of order statistics. Then Taylor series expansions are used to linearize them. The original intractable terms are replaced by their linear approximations and the resulting equations are solved.

2.1. AR (1) Model Without Intercept Term

2.1.1. MMLE for Long – Tailed Symmetric Distribution

For a long - tailed symmetric distribution, the MML estimators are found by using the following steps:

$$f(\varepsilon; p) = \frac{1}{\sigma k^{1/2} \beta\left(\frac{1}{2}; p - \frac{1}{2}\right)} \left\{1 + \frac{\varepsilon^2}{k\sigma^2}\right\}^{-p} ; -\infty < \varepsilon < \infty \qquad (2.1)$$

where p is a shape parameter and $p \geq 2$, $k = 2p - 3$ and $\varepsilon_t = y_t - \phi_1 y_{t-1}$, $\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, $E(\varepsilon_t) = 0$ and $V(\varepsilon_t) = \sigma^2$.

Also $t = \sqrt{\vartheta / k}\left(\varepsilon / \sigma\right)$ has a Student – t distribution with $\vartheta = 2p - 1$ degree of freedom. Firstly the likelihood function L conditional on y_0 and the derivatives of InL are written and equated to zero:

$$L = \prod_{t=1}^{n} f(\varepsilon_t; p) \propto \left(\frac{1}{\sigma}\right)^n \prod_{t=1}^{n} \left\{1 + \frac{\varepsilon_t}{k\sigma^2}\right\}^{-p}, \qquad (2.2)$$

$$\frac{\partial InL}{\partial \phi_1} = \frac{2p}{k\sigma} \sum_{t=1}^{n} g\{z_t\} y_{t-1} = 0, \tag{2.3}$$

and

$$\frac{\partial InL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma} \sum_{t=1}^{n} z_t g\{z_t\} = 0 \tag{2.4}$$

where $z_t = \frac{\varepsilon_t}{\sigma} = \frac{y_t - \phi_1 y_{t-1}}{\sigma}$,

$$g\{z\} = \frac{z}{1 + \frac{1}{k} z^2}.$$

Then, we express these equations in terms of the order statistics of a random sample:

$$\frac{\partial InL}{\partial \phi_1} = \frac{2p}{k\sigma} \sum_{i=1}^{n} g\{z_{(i)}\} y_{(i)-1} = 0, \tag{2.5}$$

$$\frac{\partial InL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma} \sum_{i=1}^{n} z_{(i)} g\{z_{(i)}\} = 0, \tag{2.6}$$

$$z_{(i)} = \frac{y_{(i)} - \phi_1 y_{(i)-1}}{\sigma}, \tag{2.7}$$

and

$$g\{z_{(i)}\} = \frac{z_{(i)}}{1+\frac{1}{k}z_{(i)}^{2}} \; ; \; 1 \le i \le n \, . \tag{2.8}$$

As complete sums are invariant to ordering, they do not change in their values. Here, $y_{(i-1)} = y_{(i)-1}$ and $(y_{(i)}; y_{(i)-1})$ is the pair of (y_t, y_{t-1}) observations which determines $z_{(i)}$.

It is noticed that (2.6) has multiple roots (Vaughan 1992(a)). However, the function $g\{z_{(i)}\}$ is linear in small intervals, and $z_{(i)}$ is located in the vicinity of $t_{(i)}$ at any rate of large n. So we apply Taylor series expansion to linearize $g\{z_{(i)}\}$:

$$g\{z_{(i)}\} \approx g\{t_{(i)}\} + [z_{(i)} - t_{(i)}]\left[\frac{d}{dz}g\{z\}\right]_{z=t_{(i)}} = \alpha_i + \beta_i z_{(i)} ; 1 \le i \le n \tag{2.9}$$

where $t_{(i)} = E\{z_{(i)}\}$ are the expected values of the order statistics of a random sample of size n and are obtained from the equation

$$\int_{-\infty}^{t_{(i)}} f(z)dz = \frac{i}{n+1} ; 1 \le i \le n \, ; \tag{2.10}$$

$$\alpha_i = \frac{(2/k)t_{(i)}^{3}}{[1+(1/k)t_{(i)}^{2}]^2} \text{ and } \beta_i = \frac{1-(1/k)t_{(i)}^{2}}{[1+(1/k)t_{(i)}^{2}]^2} \; ; \; 1 \le i \le n \, . \tag{2.11}$$

When $p = \infty$ (i.e. normal distribution), $\alpha_i = 0$ and $\beta_i = 1$. Moreover since the distribution is symmetric $\alpha_i = -\alpha_{n-i+1}$, $\beta_i = \beta_{n-i+1}$ and $\sum_{i=1}^{n} \alpha_i = 0$.

As a result we obtain the modified likelihood equations as expressed below:

$$\frac{\partial InL}{\partial \phi_1} \approx \frac{\partial InL^*}{\partial \phi_1} = \frac{2p}{k\sigma} \sum_{i=1}^{n} y_{(i)-1} \left\{ \alpha_i + \beta_i z_{(i)} \right\} = 0 , \tag{2.12}$$

$$\frac{\partial InL}{\partial \sigma} \approx \frac{\partial InL^*}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma} \sum_{i=1}^{n} z_{(i)} \left\{ \alpha_i + \beta_i z_{(i)} \right\} = 0 . \tag{2.13}$$

Then, we solve both of them. Finally we get the MML estimators, which are explicit functions of sample observations:

$$\hat{\phi}_1 = K + D\hat{\sigma} \tag{2.14}$$

and

$$\hat{\sigma} = \frac{B + \sqrt{B^2 + 4nC}}{2n} \tag{2.15}$$

where $K = \dfrac{\sum_{i=1}^{n} \beta_i y_{(i)} y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2}$,

$$D = \frac{\sum_{i=1}^{n} \alpha_i y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2} ,$$

$$B = \frac{2p}{k}\sum_{i=1}^{n}\alpha_i\left\{y_{(i)} - Ky_{(i)-1}\right\}^2 ,$$

and

$$C = \frac{2p}{k}\sum_{i=1}^{n}\beta_i\left\{y_{(i)} - Ky_{(i)-1}\right\}^2 = \frac{2p}{k}\left\{\sum_{i=1}^{n}\beta_i y_{(i)}^2 - K\sum_{i=1}^{n}\beta_i y_{(i)} y_{(i)-1}\right\}.$$

Later the divisor n in the original $\hat{\sigma}$ expression is replaced by $\sqrt{n(n-1)}$ to reduce the bias. Furthermore, in a very rare case, especially $p \le 3$, C can assume a negative value. So, $\hat{\sigma}$ has a possibility of negativity. Thus when C<0, the coefficients with α_i and β_i are replaced by 0 and $\frac{1}{\left[1+\left(\frac{1}{k}\right)t_{(i)}^2\right]}$ respectively. This change does not affect the efficiencies of $\hat{\phi}_1$ and $\hat{\sigma}$ in any substantial way but guarantees a real and positive estimate of σ.

The asymptotic variances and covariance of MML estimators are computed from $I_{(\phi_1,\sigma)}{}^{-1}$, where I is the Fisher information matrix (Appendix A):

$$I_{(\phi_1,\sigma)}{}^{-1} = \begin{bmatrix} \dfrac{(p+1)\left(p-\frac{3}{2}\right)(1-\phi_1^2)}{p\left(p-\frac{1}{2}\right)n} & 0 \\ 0 & \dfrac{(p+1)\sigma^2}{\left(p-\frac{1}{2}\right)2n} \end{bmatrix}. \qquad (2.16)$$

It is known that the diagonal elements of inverse information matrix gives us minimum variance bounds for estimating ϕ_1 and σ.

Generally the MML estimators have the following properties:

1. Asymptotically, the MML estimators are identical with the ML estimators, i.e.,

when $n \to \infty$ and

$$\frac{\partial InL}{\partial \phi_1} \cong \frac{\partial InL^*}{\partial \phi_1} \text{ and } \frac{\partial InL}{\partial \sigma} \cong \frac{\partial InL^*}{\partial \sigma} \text{ for all } \phi_1 \text{ and } \sigma . \qquad (2.17)$$

2. Asymptotically, the MML estimators are fully efficient, i.e. they are unbiased, and have minimum variance bounds (Bhattacharya 1985, Vaughan and Tiku 2000).

3. For small samples, the MML estimators are almost fully efficient. In other words they have very little or no bias and their variances are only marginally bigger than the minimum variance bounds.

4. The MML estimators have the invariance property.

5. The MML estimators have closed forms and have the same structure irrespective of the underlying distribution.

6. The computation procedure of the MML estimators is easy.

7. The MML estimators are robust estimators; i.e. they are fully or nearly efficient for an assumed distribution and maintain high efficiency for plausible alternatives to the underlying distribution.

The MMLE are calculated in two iterations. Firstly, ε_i terms are ordered by using the LS estimator $\hat{\Phi}_0$, and $\hat{\phi}_1$ is calculated. Then, $\hat{\Phi}_0$ is replaced by $\hat{\phi}_1$ and a new estimator $\hat{\phi}_1$ calculated. The final values of ϕ_1 and σ give us the MML estimators. Since the LS estimators are based on only minimizing sum of squares of residuals, they are invariant of the assumed distribution. So sum of squares of the errors is

$$S = \sum \varepsilon_t^2 = \sum_{t=1}^{n} \left(y_t - \phi_1 y_{t-1}\right)^2 . \tag{2.18}$$

Now we minimize S with respect to ϕ_1 and obtain LS estimators from the solution of the following normal equation:

$$\frac{\partial S}{\partial \phi_1} = -2\sum_{t=1}^{n} \left(y_t - \phi_1 y_{t-1}\right) = 0 . \tag{2.19}$$

$$\tilde{\phi}_1 = \frac{\sum_{t=1}^{n} y_t y_{t-1}}{\sum_{t=1}^{n} y_{t-1}^2} , \tag{2.20}$$

$$\tilde{\sigma} = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \tilde{\phi}_1 y_{t-1}\right)^2}{n-1}} . \tag{2.21}$$

The least square estimators are exactly the same as the ML or MML estimators with respect to normal samples. However, when our distribution is non - normal, the LS estimators are significantly less efficient. Tiku, Wong et

al. (1996) evaluate relative efficiency of LS and MML estimators by the ratios below:

$$RE(\tilde{\phi}_1) = 100\frac{V(\hat{\phi}_1)}{V(\tilde{\phi}_1)} \tag{2.22}$$

and

$$RE(\tilde{\sigma}) = 100\frac{V(\hat{\sigma})}{V(\tilde{\sigma})}. \tag{2.23}$$

They simulated their means and variances for n=50, 100, 200 and 300, and p =2, 2.5, 3, 3.5, 4 and 5 under mentioned symmetric family and tabulated their values by 10000 Monte Carlo runs for n ≤ 100, and 5000 runs for n >100. They took $\sigma = 1$ without loss of generality. As a result it is seen that the MML estimators are clearly more efficient than LS estimators for all values of n and *p*. Accordingly it is concluded that LS estimators are not robust and their variances and covariances are found in the following way:

$$Var\begin{pmatrix}\tilde{\phi}_1 \\ \tilde{\sigma}\end{pmatrix} = \begin{bmatrix} n & 0 \\ 0 & 2n\left(1+\frac{1}{2}\lambda_4\right)^{-1}\end{bmatrix}^{-1} V(\varepsilon_t) \tag{2.24}$$

where $\lambda_4 = \beta_2 - 3$ and $V(\varepsilon_t) = \sigma^2$.

These non - normal situations are extended to different error distributions, e.g., gamma and generalized logistic.

2.1.2. MMLE for Gamma Distribution

When ε_t's are iid and have a gamma distribution, our density function has the following form:

$$f(\varepsilon;k) = \frac{1}{\sigma^k \Gamma(k)} e^{-\frac{\varepsilon}{\sigma}} \varepsilon^{k-1}; 0 < \varepsilon < \infty \tag{2.25}$$

where k is the shape parameter and is greater than zero, $\varepsilon_t = y_t - \phi_1 y_{t-1}$, $E(\varepsilon) = k\sigma^2$ and $V(\varepsilon) = \sigma^2$.

So, firstly, LS estimators are found by minimizing sum of squares of error terms and then are obtained from normal equations similar to the normality case. But here since $\tilde{\phi}_1$ and $\tilde{\sigma}$ have nonzero mean, we should formulate the least square estimators as

$$\tilde{\phi}_1 = \frac{\sum_{t=1}^{n} y_t (y_{t-1} - \bar{y})}{\sum_{t=1}^{n} (y_{t-1} - \bar{y})^2} \text{ and } \tilde{\sigma} = \sqrt{\frac{\sum_{t=1}^{n} \left\{ y_t - \bar{y} - \tilde{\phi}_1 (y_{t-1} - \bar{y}) \right\}^2}{(n-1)k}}. \tag{2.26}$$

Then the likelihood equations are written in terms of the ordered variates:

$$z_{(i)} = \frac{y_{(i)} - \phi_1 y_{(i)-1}}{\sigma} \quad (1 \le i \le n), \tag{2.27}$$

$$\frac{\partial InL}{\partial \phi_1} = \frac{1}{\sigma} \sum_{i=1}^{n} y_{(i)-1} - \frac{(k-1)}{\sigma} \sum_{i=1}^{n} y_{(i)-1} z_{(i)}^{-1} = 0, \tag{2.28}$$

$$\frac{\partial InL}{\partial \sigma} = -\frac{nk}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n} z_{(i)} = 0\,. \tag{2.29}$$

Solving these equations by iteration has difficulties. Moreover as n becomes large, $z_{(i)}$ converges to zero in which case (2.27) is not defined. Thus the modified likelihood equations are computed. But for second equation, we recast it as follows so that we get a positive estimator $\hat{\sigma}$:

$$\frac{\partial InL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n} z_{(i)} - \frac{(k-1)}{\sigma}\sum_{i=1}^{n} z_{(i)} z_{(i)}{}^{-1} = 0\,. \tag{2.30}$$

Because when outliers enter the time series at very early stages, the solution of (2.28) can assume a negative value (Akkaya and Tiku 2001).

To calculate the MML estimators, $z_{(i)}^{-1}$ is linearized by Taylor series expansion:

$$z_{(i)}^{-1} \cong \alpha_i - \beta_i z_{(i)} \tag{2.31}$$

where $\alpha_i = \dfrac{2}{t_{(i)}}$, $\beta_i = \dfrac{1}{t_{(i)}^2}$ and $t_{(i)} = E\{z_{(i)}\}$.

These terms are incorporated in the modified equations and the MML estimators are obtained as:

$$\hat{\phi}_1 = K - D\hat{\sigma} \tag{2.32}$$

and

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2\sqrt{n(n-1)}} \tag{2.33}$$

where $K = \dfrac{\sum_{i=1}^{n} \beta_i y_{(i)} y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2}$,

$$D = \frac{\sum_{i=1}^{n} \Delta_i y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2},$$

$$B = (k-1)\sum_{i=1}^{n} \Delta_i (y_{(i)} - K y_{(i)-1}),$$

$$C = (k-1)\sum_{i=1}^{n} \beta_i (y_{(i)} - K y_{(i)-1})^2,$$

and

$$\Delta_i = \alpha_i - \frac{1}{k-1}.$$

In order to calculate relative efficiency between the pair $(\tilde{\phi}_1; \hat{\phi}_1)$ and the pair $(\tilde{\sigma}; \hat{\sigma})$, variance - covariance matrix of both estimators is determined.

The asymptotic variances and covariances of MML estimators are computed from $I_{(\phi_1, \sigma)}{}^{-1}$ (Appendix B).

$$I_{(\phi_1,\sigma)}^{-1} = \begin{bmatrix} \dfrac{(k-2)}{nk\left\{\dfrac{1}{1-\phi_1^2}+\dfrac{k}{(1-\phi_1)^2}\right\}} & \dfrac{(1-\phi_1)\sigma}{nk} \\ \dfrac{(1-\phi_1)\sigma}{nk} & \dfrac{\sigma^2}{nk} \end{bmatrix}. \qquad (2.34)$$

For LS estimators, on the other hand, the variances and covariances are difficult to determine theoretically.

Consequently simulated values indicate that for all values of n=20, 30, 50, 100 and 300 and k=3.0, 4.0, 5.0 and 8.0, the MML estimators are more efficient than the least square ones although both of them are unbiased (almost) estimators of unknown parameter ϕ_1 and σ.

2.1.3. MMLE for Generalized Logistic Distribution

If error terms are iid and have a generalized logistic distribution, the pdf has the following form:

$$f(\varepsilon;b) = \frac{b}{\sigma}\frac{\exp(-\varepsilon/\sigma)}{\{1+\exp(-\varepsilon/\sigma)\}^{b+1}} \ ; -\infty < \varepsilon < \infty \qquad (2.35)$$

where b is the shape parameter, $E(\varepsilon_t) = \sigma[\Psi(b) - \Psi(1)]$ and $V(\varepsilon_t) = [\Psi'(b) + \Psi'(1)]\sigma^2$.

For different value of b, probability distribution takes various shapes. So when b =1, (2.34) is the logistic distribution and is symmetric. When b <1,

the distribution is left – skewed (negatively skewed) and when $b > 1$, it is right - skewed (positively skewed).

Since (2.34) has nonzero mean, the LS are defined as

$$\tilde{\phi}_1 = \frac{\sum_{t=1}^{n} y_t (y_{t-1} - \bar{y})}{\sum_{t=1}^{n} (y_{t-1} - \bar{y})^2} \quad \text{and} \quad \tilde{\sigma} = \sqrt{\frac{\sum_{t=1}^{n} \left\{ y_t - \bar{y} - \tilde{\phi}_1 (y_{t-1} - \bar{y}) \right\}^2}{(n-1)\left[\Psi'(b) + \Psi'(1)\right]}} . \qquad (2.36)$$

As before, the likelihood equations are obtained. They are

$$\frac{\partial InL}{\partial \phi_1} = \frac{\sum_{i=1}^{n} y_{)(i)-1}}{\sigma} - \frac{(b+1)}{\sigma} \sum_{i=1}^{n} \frac{e^{-z_{(i)}}}{1+e^{-z_{(i)}}} y_{(i)-1} = 0 \qquad (2.37)$$

and

$$\frac{\partial InL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} z_{(i)} - \frac{(b+1)}{\sigma} \sum_{i=1}^{n} \frac{e^{-z_{(i)}}}{1+e^{-z_{(i)}}} z_{(i)} = 0 \qquad (2.38)$$

where $z_{(i)} = \dfrac{y_{(i)} - \phi_1 y_{(i)-1}}{\sigma}$.

Because of the existence of function $\dfrac{e^{-z}}{1+e^{-z}}$ in likelihood equation, they do not have explicit solutions and have to be solved by iterative methods, which can be problematic (Ross 1990, Tiku et al. 1986). Therefore the method of MML is used to find the estimators (Tiku, Wong, Bian 1999).

As a consequence, $\frac{e^{-z}}{1+e^{-z}}$ function is linearized by using the Taylor series expansion. Thus,

$$\frac{e^{-z}}{1+e^{-z}} \cong \alpha_i - \beta_i z_{(i)} \tag{2.39}$$

where $\alpha_i = \frac{1}{1+e^{t_{(i)}}} + \beta_i t_{(i)}$ or $\alpha_i = \frac{e^{t_{(i)}}\left(1+t_{(i)}+e^{-t_{(i)}}\right)}{\left(1+e^{-t_{(i)}}\right)^2}$,

$$\beta_i = \frac{e^{-t_{(i)}}}{(1+e^{t_{(i)}})} \quad ; \; t_{(i)} = E\{z_{(i)}\}.$$

The approximate values $t_{(i)}$ are generated from the equality:

$$\int_{-\infty}^{t_{(i)}} f(z)dz = \frac{i}{n+1}; 1 \le i \le n, \tag{2.40}$$

like all other mentioned distributions.

Then putting these new terms on the likelihood equations and solving them simultaneously, we obtain the MML estimators of ϕ_1 and σ such that;

$$\hat{\phi}_1 = K + D\hat{\sigma} \tag{2.41}$$

and

$$\hat{\sigma} = \frac{B+\sqrt{B^2+4nC}}{2\sqrt{n(n-1)}} \tag{2.42}$$

where $K = \dfrac{\sum_{i=1}^{n} \beta_i y_{(i)} y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2}$,

$$D = \frac{\sum_{i=1}^{n} \Delta_i y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2},$$

$$B = (b+1)\sum_{i=1}^{n} \Delta_i (y_{(i)} - K y_{(i)-1}),$$

and

$$C = (b+1)\sum_{i=1}^{n} \beta_i (y_{(i)} - K y_{(i)-1})^2 = (b+1)\left\{\sum_{i=1}^{n} \beta_i y_{(i)}{}^2 - K\sum_{i=1}^{n} \beta_i y_{(i)} y_{(i)-1}\right\}.$$

To measure efficiency among LS and MML estimators, we take ratio of their variances. The elements of the variance – covariance matrix for LS estimators do not change completely, only λ_4 term is recalculated for the underlying distribution. On the other hand its form for the MML estimators is given as below (Appendix C):

$$I_{(\phi_1,\sigma)}{}^{-1} = \begin{bmatrix} \dfrac{(b+2)\sigma^2}{nb}\gamma' & \dfrac{(b+2)(1-\phi_1)\sigma}{b[\Psi(b)-\Psi(1)]} \\ \dfrac{(b+2)(1-\phi_1)\sigma}{b[\Psi(b)-\Psi(1)]} & \gamma'' \end{bmatrix} \tag{2.43}$$

where $\gamma' = \left\{ \frac{\Psi'(b) + \Psi(1)}{1-\phi_1^2} + \frac{[\Psi(b) - \Psi(1)]^2}{(1-\phi_1)^2} \right\}^{-1}$,

$$\gamma' = \left[\frac{n}{\sigma^2} + \frac{nb}{(b+2)\sigma^2} \left\{ \Psi'(b+1) + \Psi'(2) + \left(\Psi(b+1) - \Psi(2) \right)^2 \right\} \right]^{-1}.$$

According to the simulation results, it is clear that the MML estimators are efficient than the LS estimators for the distributions considered. Moreover, they have very little bias as compared to the least squares estimators for all values of n and b.

Up to now we were interested in AR (1) model without intercept parameter. We now introduce the intercept parameter also.

2.2. AR (1) Model With Intercept Term

Now we extent our model; that is, we add an intercept term in the autoregressive model and get the following equation:

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t \ ; (0 \le t \le n). \tag{2.44}$$

Then we repeat our estimations based on this new model. Although the structure of MML and LS estimators changes, the procedure of calculations is not altered so much in fact. The main difference comes from the constant term μ.

2.2.1. MMLE for Long – Tailed Symmetric Distribution

Under Student – t distribution (long – tailed symmetric distribution) (2.1), the ordered likelihood equations are:

$$\frac{\partial InL}{\partial \mu} = \frac{2p}{k\sigma} \sum_{i=1}^{n} g\{z_{(i)}\} = 0, \tag{2.45}$$

$$\frac{\partial InL}{\partial \phi_1} = \frac{2p}{k\sigma} \sum_{i=1}^{n} g\{z_{(i)}\} y_{(i)-1} = 0\,, \tag{2.46}$$

$$\frac{\partial InL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma} \sum_{i=1}^{n} z_{(i)} g\{z_{(i)}\} = 0\,, \tag{2.47}$$

$$z_{(i)} = \frac{y_{(i)} - \phi_1 y_{(i)-1} - \mu}{\sigma}\,, \tag{2.48}$$

$$g\{z_{(i)}\} = \frac{z_{(i)}}{1 + \frac{1}{k} z_{(i)}^{\ 2}} \ ;\ 1 \le i \le n\,. \tag{2.49}$$

Since all likelihood equations involve the functions $g\{z_{(i)}\}$, same as the equation obtained under AR (1) without intercept term, ML equations do not give exact solutions. Therefore, we linearize these equations as follows:

$$g\{z_{(i)}\} \approx g\{t_{(i)}\} + \left[z_{(i)} - t_{(i)}\right] \left[\frac{d}{dz} g\{z\}\right]_{z=t_{(i)}} = \alpha_i + \beta_i z_{(i)}; 1 \le i \le n\,, \tag{2.50}$$

$$\alpha_i = \frac{(2/k)t_{(i)}^3}{\left[1+(1/k)t_{(i)}^2\right]^2} \text{ and } \beta_i = \frac{1-(1/k)t_{(i)}^2}{\left[1+(1/k)t_{(i)}^2\right]^2} \; ; \; 1 \le i \le n \tag{2.51}$$

similar to the previous model. By incorporating the linearized $g\{z_{(i)}\}$, modified likelihood equations are obtained. Their simultaneous solutions give the MML estimators (Tiku, Wong, Bian 1999), namely,

$$\hat{\mu} = \overline{\omega}_{(.)}, \tag{2.52}$$

$$\hat{\sigma} = \frac{B+\sqrt{B^2+4nC}}{2n}, \tag{2.53}$$

$$\hat{\phi}_1 = K + D\hat{\sigma} \tag{2.54}$$

where $\omega_{(i)} = y_{(i)} - \phi_1 y_{(i)-1}$,

$$\overline{\omega}_{(.)} = \frac{1}{m}\sum_{i=1}^{n}\beta_i\omega_{(i)} \;\; , \; m = \sum_{i=1}^{n}\beta_i \; ,$$

$$B = \frac{2p}{k}\sum_{i=1}^{n}\alpha_i\omega_{(i)} \; ,$$

$$C = \frac{2p}{k}\sum_{i=1}^{n}\beta_i\left\{\omega_{(i)} - \overline{\omega}_{(.)}\right\}^2 = \frac{2p}{k}\left\{\sum_{i=1}^{n}\beta_i\omega_{(i)}^2 - m\overline{\omega}_{(.)}^2\right\},$$

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

and

$$D = \frac{\sum_{i=1}^{n} \alpha_i y_{(i)-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2}.$$

Then the divisor n in the original $\hat{\sigma}$ expression is replaced by $\sqrt{n(n-2)}$ to reduce the bias. The asymptotic variances and covariance of MML estimators are computed from $I_{(\mu;\phi_1;\sigma)}{}^{-1}$, where I is the Fisher information matrix (Appendix D):

$$I_{(\mu,\phi_1,\sigma)}{}^{-1} = \begin{bmatrix} \dfrac{(p+1)\left(p-\frac{3}{2}\right)\sigma^2}{np\left(p-\frac{1}{2}\right)} & \dfrac{(p+1)\left(p-\frac{3}{2}\right)(1-\phi_1)\sigma^2}{np\left(p-\frac{1}{2}\right)\mu} & 0 \\ \dfrac{(p+1)\left(p-\frac{3}{2}\right)(1-\phi_1)\sigma^2}{np\left(p-\frac{1}{2}\right)\mu} & \dfrac{(p+1)\left(p-\frac{3}{2}\right)\sigma^2}{np\left(p-\frac{1}{2}\right)E\left(y^2\right)} & 0 \\ 0 & 0 & \dfrac{(p+1)\sigma^2}{2n\left(p-\frac{1}{2}\right)} \end{bmatrix} \tag{2.55}$$

where $E\left(y^2\right)=\dfrac{\sigma^2}{\left(1-\phi_1^2\right)}+\dfrac{\mu^2}{\left(1-\phi_1\right)^2}$.

On the other hand, LS estimators are derived by taking derivatives of the following sum of square of residuals with respect to μ and ϕ_1, and then equating them to zero,

$$S=\sum \varepsilon_t^2=\sum_{t=1}^{n}\left(y_t-\phi_1 y_{t-1}-\mu\right)^2 \quad . \tag{2.56}$$

Hence solutions of the normal equations give the LS estimators below:

$$\tilde{\mu}=\frac{\left(\sum_{t=1}^{n} y_t\right)\left(\sum_{t=1}^{n} y_{t-1}^2\right)}{n\sum_{t=1}^{n} y_{t-1}^2-\left(\sum_{t=1}^{n} y_{t-1}\right)^2}-\frac{\left(\sum_{t=1}^{n} y_{t-1}\right)\left(\sum_{t=1}^{n} y_{t-1}y_t\right)}{n\sum_{t=1}^{n} y_{t-1}^2-\left(\sum_{t=1}^{n} y_{t-1}\right)^2}, \tag{2.57}$$

$$\tilde{\phi}_1=\frac{n\left(\sum_{t=1}^{n} y_{t-1}y_t\right)}{n\sum_{t=1}^{n} y_{t-1}^2-\left(\sum_{t=1}^{n} y_{t-1}\right)^2}-\frac{\left(\sum_{t=1}^{n} y_t\right)\left(\sum_{t=1}^{n} y_{t-1}\right)}{n\sum_{t=1}^{n} y_{t-1}^2-\left(\sum_{t=1}^{n} y_{t-1}\right)^2}, \tag{2.58}$$

and

$$\tilde{\sigma}=\sqrt{\frac{\sum_{t=1}^{n}\left(y_t-\tilde{\phi}_1 y_{t-1}-\tilde{\mu}\right)^2}{(n-2)}} . \tag{2.59}$$

Additionally, the variance - covariance matrix of LS estimators is

$$Var\begin{pmatrix}\tilde{\mu}\\ \tilde{\phi}_1\\ \tilde{\sigma}\end{pmatrix}=\begin{bmatrix} n & nE(y) & 0\\ nE(y) & nE(y^2) & 0\\ 0 & 0 & 2n\left(1+\frac{1}{2}\lambda_4\right)^{-1}\end{bmatrix}^{-1} V(\varepsilon_t) \qquad (2.60)$$

where $\lambda_4 = \beta_2 - 3$, $V(\varepsilon_t) = \sigma^2$,

$$E(y) = \frac{\mu}{(1-\phi_1)} \ ,$$

and

$$E\left(y^2\right) = \frac{\sigma^2}{\left(1-\phi_1^2\right)} + \frac{\mu^2}{(1-\phi_1)^2} \ .$$

While comparing the results from MML and LS estimators under various shape parameters (p = 2.0, 2.5, 3.5, 5.0, 10.0) and sample sizes (n= 20, 30, 50, 100, and 300), it is found that MML estimators of μ, ϕ_1 and σ are more efficient than LS estimators in each case. Moreover they have smaller bias than the LS estimators.

2.2.2. MMLE for Gamma Distribution

When ε_t's are iid and have a gamma distribution, the MML estimators are (Akkaya, Tiku 2001)

$$\hat{\mu} = \bar{v}_{(.)} - \frac{\Delta}{m}\hat{\sigma}\,, \tag{2.61}$$

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n}\,, \tag{2.62}$$

$$\hat{\phi}_1 = K - D\hat{\sigma} \tag{2.63}$$

where $v_{(i)} = y_{(i)} - \hat{\phi}_1 y_{(i)-1}$,

$$\bar{v}_{(.)} = \frac{1}{m}\sum_{i=1}^{n}\beta_i v_{(i)} \;\;,\; m = \sum_{i=1}^{n}\beta_i\,,$$

$$\Delta_i = \alpha_i - \frac{1}{k-1}\,,\; \Delta = \sum_{i=1}^{n}\Delta_i\,,$$

$$B = (k-1)\sum_{i=1}^{n}\Delta_i\,(v_{(i)} - \bar{v}_{(i)-1})\,,$$

$$C = (k-1)\sum_{i=1}^{n}\beta_i\,(v_{(i)} - \bar{v}_{(i)-1})^2\,,$$

$$K = \frac{\sum_{i=1}^{n} \beta_i y_{(i)} y_{(i)-1} - \frac{1}{m}\left(\sum_{i=1}^{n} \beta_i y_{(i)-1}\right)\left(\sum_{i=1}^{n} \beta_i y_{(i)}\right)}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2 - \frac{1}{m}\left(\sum_{i=1}^{n} \beta_i y_{(i)-1}\right)^2},$$

$$D = \frac{\sum_{i=1}^{n} \left\{\Delta_i - \left(\frac{\Delta}{m}\right)\beta_i\right\} y_{i-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2 - \frac{1}{m}\left(\sum_{i=1}^{n} \beta_i y_{(i)-1}\right)^2},$$

$$\alpha_i = \frac{2}{t_{(i)}},\ \beta_i = \frac{1}{t_{(i)}^2} \text{ and } t_{(i)} = E\{z_{(i)}\}.$$

The LS estimators are

$$\tilde{\mu} = \bar{v} - \tilde{\sigma} k\,, \tag{2.64}$$

$$\tilde{\phi}_1 = \frac{n\left(\sum_{i=1}^{n} y_{i-1} y_i\right)}{n\sum_{i=1}^{n} y_{i-1}^2 - \left(\sum_{i=1}^{n} y_{i-1}\right)^2} - \frac{\left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} y_{i-1}\right)}{n\sum_{i=1}^{n} y_{i-1}^2 - \left(\sum_{i=1}^{n} y_{i-1}\right)^2}, \tag{2.65}$$

and

$$\tilde{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} \{v_i - \bar{v})\}^2}{(n-2)k}} \tag{2.66}$$

where $v_i = y_i - \tilde{\phi}_1 y_{i-1}$ and $\bar{v} = \dfrac{\sum_{i=1}^{n} v_i}{n}$.

The asymptotic variance – covariance matrix of LSE is calculated by a similar way as stated in equation (2.59). But previous $E(y)$, $E(y^2)$ terms are replaced by $\dfrac{\mu + k\sigma}{1-\phi_1}$ and $\dfrac{k\sigma^2}{(1-\phi_1^2)} + \left(\dfrac{\mu + k\sigma}{1-\phi_1}\right)^2$ respectively and λ_4 value is computed for the underlying distribution.

On the other hand the elements of the variance – covariance matrix of MML estimators are determined by the inverse of Fisher's information matrix (Appendix E).

$$I_{(\mu,\phi_1,\sigma)}^{-1} = \begin{bmatrix} \dfrac{(k-2)\sigma^2}{n} & \dfrac{(k-2)(1-\phi_1)\sigma^2}{n(\mu+k\sigma)} & \dfrac{\sigma^2}{n} \\ \dfrac{(k-2)(1-\phi_1)\sigma^2}{n(\mu+k\sigma)} & \dfrac{(k-2)\sigma^2}{nE(y^2)} & \dfrac{(1-\phi_1)\sigma^2}{n(\mu+k\sigma)} \\ \dfrac{\sigma^2}{n} & \dfrac{(1-\phi_1)\sigma^2}{n(\mu+k\sigma)} & \dfrac{\sigma^2}{nk} \end{bmatrix} \quad (2.67)$$

where $E(y^2) = \dfrac{k\sigma^2}{(1-\phi_1^2)} + \left(\dfrac{\mu + k\sigma}{1-\phi_1}\right)^2$.

Now we can compare efficiencies of both estimators. From the simulations' results, there is a clear evidence that the MML estimators give always smaller variances than the LS estimators. Besides, the MMLE have negligible bias.

2.2.3. MMLE for Generalized Logistic Distribution

When ε_t's are iid and have a generalized logistic distribution (2.34), the MML estimators are obtained as follows:

$$\hat{\mu} = \bar{v}_{(.)} - \frac{\Delta}{m}\hat{\sigma}\ , \tag{2.68}$$

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n}\ , \tag{2.69}$$

and

$$\hat{\phi}_1 = K - D\hat{\sigma} \tag{2.70}$$

where $v_{(i)} = y_{(i)} - \hat{\phi}_1 y_{(i)-1}$,

$$\bar{v}_{(.)} = \frac{1}{m}\sum_{i=1}^{n}\beta_i v_{(i)}\ \ ,\ m = \sum_{i=1}^{n}\beta_i\ ,$$

$$\Delta_i = \alpha_i - \frac{1}{b+1}\ \ ,\ \Delta = \sum_{i=1}^{n}\Delta_i\ ,$$

$$B = (b+1)\sum_{i=1}^{n}\Delta_i(v_{(i)} - \bar{v}_{(.)})\ ,$$

$$C = (b+1)\sum_{i=1}^{n}\beta_i(v_{(i)} - \bar{v}_{(.)})^2\ ,$$

$$K = \frac{\sum_{i=1}^{n} \beta_i y_{(i)} y_{(i)-1} - \frac{1}{m}\left(\sum_{i=1}^{n} \beta_i y_{(i)-1}\right)\left(\sum_{i=1}^{n} \beta_i y_{(i)}\right)}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2 - \frac{1}{m}\left(\sum_{i=1}^{n} \beta_i y_{(i)-1}\right)^2},$$

$$D = \frac{\sum_{i=1}^{n} \left\{\Delta_i - \left(\frac{\Delta}{m}\right)\beta_i\right\} y_{i-1}}{\sum_{i=1}^{n} \beta_i y_{(i)-1}^2 - \frac{1}{m}\left(\sum_{i=1}^{n} \beta_i y_{(i)-1}\right)^2},$$

$$\beta_i = \frac{e^{-t_{(i)}}}{(1+e^{t_{(i)}})} \; ; \; t_{(i)} = E\{z_{(i)}\}.$$

The LS estimators are given by

$$\tilde{\mu} = \bar{v} - \tilde{\sigma}[\Psi(b) - \Psi(1)], \tag{2.71}$$

$$\tilde{\phi}_1 = \frac{n\left(\sum_{i=1}^{n} y_{i-1} y_i\right)}{n\sum_{i=1}^{n} y_{i-1}^2 - \left(\sum_{i=1}^{n} y_{i-1}\right)^2} - \frac{\left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} y_{i-1}\right)}{n\sum_{i=1}^{n} y_{i-1}^2 - \left(\sum_{i=1}^{n} y_{i-1}\right)^2}, \tag{2.72}$$

and

$$\tilde{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} \{v_i - \bar{v})\}^2}{(n-2)[\Psi'(b) + \Psi'(1)]}} \tag{2.73}$$

where $v_i = y_i - \tilde{\phi}_1 y_{i-1}$ and $\bar{v} = \dfrac{\sum_{i=1}^{n} v_i}{n}$.

As applied before, the $Var\begin{pmatrix}\tilde{\mu}\\ \tilde{\phi}_1\\ \tilde{\sigma}\end{pmatrix}$ matrix in equation (2.59) is redefined by new $E(y)$ and $E(y^2)$ terms stated below (2.73). Except these terms and recomputed λ_4 value, others elements of the matrix remain same. The variance - covariance matrix of MML estimators is (Appendix F).

$$I_{(\mu,\phi_1,\sigma)}^{-1} = \begin{bmatrix} \dfrac{(b+2)\sigma^2}{nb} & \dfrac{(b+2)\sigma^2}{nb\left\{\dfrac{\mu+\sigma[\Psi(b)-\Psi(1)]}{1-\phi_1}\right\}} & \dfrac{(b+2)\sigma^2}{nb[\Psi(b)-\Psi(1)]} \\ \dfrac{(b+2)\sigma^2}{nb\left\{\dfrac{\mu+\sigma[\Psi(b)-\Psi(1)]}{1-\phi_1}\right\}} & \dfrac{(b+2)\sigma^2}{nbE(y^2)} & \dfrac{(b+2)\sigma^2}{nb[\Psi(b)-\Psi(1)]E(y)} \\ \dfrac{(b+2)\sigma^2}{nb[\Psi(b)-\Psi(1)]} & \dfrac{(b+2)\sigma^2}{nb[\Psi(b)-\Psi(1)]E(y)} & \dfrac{1}{\left\{\dfrac{n}{\sigma^2}+\dfrac{nbH}{(b+2)\sigma^2}\right\}} \end{bmatrix} \tag{2.74}$$

where $E(y) = \dfrac{\mu + [\Psi(b) - \Psi(1)]\sigma}{1-\phi_1}$,

$$E(y^2) = \frac{[\Psi'(b) + \Psi'(1)]\sigma^2}{1-\phi_1^2} + \left(\frac{\mu + [\Psi(b) - \Psi(1)]\sigma}{1-\phi_1}\right)^2,$$

and

$$H = \Psi'(b+1) + \Psi'(2) + \left[\Psi(b+1) - \Psi(2)\right]^2 .$$

While comparing the results obtained by simulations, the MML estimators are clearly more efficient than the LS estimators for both small and large sample sizes and different shape parameters (b =0.5, 3.0, 4.0, 6.0, 8.0). The MML estimators also have negligible bias.

CHAPTER III

R_1 AND R_0 UNIT ROOT TESTS

Now we consider the R_1 and R_0 unit root tests that were proposed by Tiku and Wong (1998), as mentioned in Chapter I. We define the statistics as

$$R_1 = \sqrt{n}(2 - \hat{\phi}_1) \tag{3.1}$$

and

$$R_0 = \sqrt{n}(2 - \tilde{\phi}_1) \tag{3.2}$$

where $\hat{\phi}_1$ is the MML estimate of ϕ_1, and $\tilde{\phi}_1$ is the LS estimate of ϕ_1.

Moreover $E(R_1) \approx \sqrt{n}$ and $E(R_0) \approx \sqrt{n}$ (Tiku, Wong 1998).

While testing

$$H_0 : \phi_1 = 1 \text{ against } H_1 : \phi_1 < 1, \tag{3.3}$$

large values of R_1 and R_0 lead to the rejection of H_0 .

It is very difficult to work out the distribution of R_0 and R_1 or its moments. To find them the Monte Carlo simulations are applied. Again for $n \leq 100$, 10000 runs; for n>100, 5000 runs are done, and also β_1 and β_2 values are tabulated besides their means and variances. Thus it is seen that these test statistics have positive skewness, and are located very close to the Type III line or are inside the F – region (Pearson and Tiku 1970, Figure I) for all values of n and shape parameters. Therefore, we use the following approximations.

Three – Moment Chi – Square Approximation: (Tiku 1966; Tiku and Wong 1998). Let μ_1' and μ_2 denote the mean and variance of a random variable X and $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ and $\beta_2 = \frac{\mu_4}{\mu_2^2}$ its third and fourth standard cumulants respectively. If the Pearson coefficients β_1 and β_2 satisfy the condition

$$\left|\beta_2 - (3 + 1.5\beta_1)\right| \leq 0.5, \tag{3.4}$$

the distribution of

$$\chi^2 = \frac{X + a}{b} \tag{3.5}$$

is approximately a central chi – square with ϑ degrees of freedom. Then a, b and ϑ are determined such that the first three moments on both sides agree. Equating these moments, we obtain

$$\vartheta = \frac{8}{\beta_1}, \; b = \sqrt{\frac{\mu_2}{2\vartheta}}, \text{ and } a = b\vartheta - \mu_1' \; . \tag{3.6}$$

As long as β_2 does not differ from $3+1.5\beta_1$ by more than 0.5, the three – moment chi – square approximation provides successful approximation for the percentage points besides the probability integral of X, except may be for the extreme left tail. Thus, distributions belonging to the Type III region are approximated by this method.

Four – moment F Approximation: (Tiku and Yip 1978; Tiku and Wong 1998) A four – moment approximation based on the F – distribution is obtained by equating the first four moments on both sides of the equation (similar to the three - moment chi – square approximation)

$$F = \frac{X+g}{h} \tag{3.7}$$

where F has the central F - distribution with $(\vartheta_1;\vartheta_2)$ degrees of freedom. The values of $\vartheta_1, \vartheta_2, g$ and h are found by equating the first four moments on both sides of F above;

$$\vartheta_2 = 2\left[3+\frac{\beta_2+3}{\beta_2-(3+1.5\beta_1)}\right], \tag{3.8}$$

$$\vartheta_1 = \frac{1}{2}(\vartheta_2-2)\left[-1+\sqrt{1+\frac{32(\vartheta_2-4)/(\vartheta_2-6)^2}{\beta_1-32(\vartheta_2-4)/(\vartheta_2-6)^2}}\right], \tag{3.9}$$

$$h = \sqrt{\frac{\vartheta_1(\vartheta_2-2)^2(\vartheta_2-4)}{2\vartheta_2^2(\vartheta_1+\vartheta_2-2)}\mu_2}\ , \tag{3.10}$$

and

$$g = \frac{\vartheta_2}{\vartheta_2 - 2} h - \mu_1'. \tag{3.11}$$

For the equation (3.7) to be valid (that is, $\vartheta_1 > 0$ and $\vartheta_2 > 0$), the $(\beta_1; \beta_2)$ values of X should satisfy

$$\beta_1 > \frac{32(\vartheta_2 - 4)}{(\vartheta_2 - 6)^2} \text{ and } \beta_2 > 3 + 1.5\beta_1. \tag{3.12}$$

The inequalities (3.12) define the F – region in the $(\beta_1; \beta_2)$ plane bounded by the χ^2- line $(\beta_2 = 3 + 1.5\beta_1)$ and the reciprocal of the χ^2- line (Pearson and Tiku 1970,Figure I). So if $(\beta_1; \beta_2)$ - points of X lie within the F – region, the four – moment F approximation gives us accurate values for the probability integral and the percentage points of X. 100 (1-α) % point of X is approximately $hF_{1-\alpha}(\vartheta_1; \vartheta_2) - g$ where $F_{1-\alpha}(\vartheta_1; \vartheta_2)$ is the 100 (1-α) % point of the central F distribution with $(\vartheta_1; \vartheta_2)$ degrees of freedom.

Consequently as the distribution of R_1 and R_0 may not be known but whose first moments satisfy the inequalities (3.12) for all values of shape parameters and n, and (3.4) for $p \geq$ 10, the four – moment F approximation and three – moment chi – square approximations give at least three (generally four) decimal place accuracy for the true distribution respectively.

To illustrate the success of these approximations using the formulas of MMLE and LSE presented in Chapter II (3.13), simulations run and the following calculations of type I error and power are made respectively:

$$P(R_1 \geq d_1 / \phi_1 = 1) \text{ and } P(R_0 \geq d_0 / \phi_1 = 1) \tag{3.13}$$

and

$$P(R_1 \geq d_1 / H_1) \text{ and } P(R_0 \geq d_0 / H_1) \tag{3.14}$$

where d_1 and d_0 are the 95% points of equations (3.5) and (3.7).

3.1. AR (1) Model Without Intercept Term

From the conclusions of Student t - distribution, it is seen that although none of the values are close to the $\beta_1 = 0$ and $\beta_2 = 3$, that is normal distribution's case (Table 3. I – A), these simulated values give almost same results which are tabulated in Shenton and Vinod's paper (1996, Table 3).

Moreover while comparing type I error and power values of R_0 and R_1, it is noticed that the probability of type I error for both test statistics are calculated by mainly four – moment F approximation, although a few of them are found by using the three–moment chi – square approximation (Table 3. I – B). Then if we focus on their powers, it is seen that R_1 test statistic (based on the MML estimators) has always considerable high power as compared to the R_0 test statistic (based on the LS estimators) for all values of n=20, 30, 50, 100 and 300 and p =2.0, 2.5, 3.5, 5.0 and 10.0 (Table 3. I – C).

Later this unit root test is extended to different distributions of error terms such as gamma and generalized logistic. Accordingly, for each distribution R_0 and R_1 are calculated and their powers and type I errors are simulated in order to see whether the two approximations mentioned above are applicable or not. For all calculations, same computation procedure is followed, initial value of y is taken to zero and same number of runs is applied by simulation (10000 runs for n ≤ 100, 5000 runs for n>100).

Firstly, when dealing with the error terms, which are iid and have a gamma distribution, it is shown that the MML estimate of ϕ_1 is considerably

more efficient than the LS estimate. Moreover, skewness and kurtosis of estimates do not indicate great departure from normal values and both estimators are (almost) unbiased under the unit root (Table 3. II – A).

Additionally from the simulations, it is shown that the three – moment chi - square approximations work very well by computing type I errors of the test statistics (Table 3. II – B). Finally again based on the relevant results, it is concluded that R_1 is accurately more powerful than R_0 especially for small sample sizes. As n increases, both tests reach high power rapidly (Table 3. II – C).

Then the other error terms, which are iid and have a generalized logistic distribution, are considered. From the results, it is determined that the MML estimators are more efficient than the LS estimators. Moreover the bias decreases as n increases for LS estimators but their values are not as close as the true ϕ_1=1 and σ=1 based on the MML estimators. If we check third and fourth moments of ϕ_1, it is noticed even though all skewness and kurtosis are greater than normal values for small n and b<6.0, β_1 and β_2 of $\hat{\phi}_1$ tend to normal ones when sample size and shape parameter increase. But we can not claim same thing for the LSE $\tilde{\phi}_1$ (Table 3. III - A). Furthermore similar to the previous distribution, the three – moment chi – square approximations perform successively in most of the calculations of type I error and generally R_1 has smaller error – probability with respect to R_0 under various b and n cases (Table 3. III - B). R_0 can reach this high level if we increase sample sizes (Table 3. III – C).

Then in order to see whether our conclusions are applicable or not if the initial value y_0 is distributed as normally with mean zero and variance $1/\left(1-\phi_1^2\right)$ for different ϕ_1 - value instead of zero, we recast our simulations when error terms have long – tailed symmetric distribution and n=50. From the

results it is proved that neither the efficiency property nor the unbiasedness of MML estimators disappear when initial values are chosen in a stationary random variable (Table 3. I – D). While calculating type I error both approximations are utilized but mainly four – moment F ones is more common. From these tabulated values it is notified that the superiority of R_1 against R_0 is valid since although both of them have almost same type I error (Table 3. I – E), R_1 indicates additionally higher power regarding to the R_0 (Table 3. I – F).

Consequently, it is seen that both gamma and generalized logistic cases fit three – moment chi – square approximation in most of the different sample sizes and shape parameters even though for long - tailed symmetric distribution generally four – moment F approximation gives accurate results. Apart from these conclusions there is an evidence that R_1 is more successful while testing unit root versus near unit root cases under whole specified error terms, shape parameters and sample sizes. Accordingly, it is also proved that the MML estimators are always efficient and unbiased under all these three main types of distribution (long – tailed symmetric, gamma, generalized logistic) and whole these findings do not change if the initial value y_0 is taken as either zero or a random variable.

3.2. AR (1) Model With Intercept Term

Now we extend our model by adding an intercept term and making Monte Carlo simulations based on the same number of runs that we made in previous cases. We obtain the following conclusions for these distributions:

- If the underling model is long – tailed symmetric family of distribution, it is found that the MML estimators are always more efficient and have relatively smaller bias than the LS estimators. Besides, though β_1 and β_2

values of $\hat{\phi}_1$ and $\tilde{\phi}_1$ are close to each other, they are very different from normal cases (Table 3. IV – A). On the other hand, type I error of R_1 test is smaller than its alternative and both of them satisfy mainly three – moment chi – square approximation. But as n is small (n=20 and n=30 while $p \geq 3.5$), tests maintain four –moment F approximation too(Table 3. IV-B). Considering their powers, it is observed that both R_1 and R_0 tests have almost same values of power especially up to n=50. But when sample size is greater than 50 and ϕ_1 is far away from the unit root situation, the power of R_1 becomes larger (Table 3. IV – C).

- While focusing on error terms having gamma distribution, we get again same conclusions such that the MML estimators are efficient and both estimators are (almost) unbiased specifically for large n and k (shape parameter). Their third and fourth moments are close to zero and three respectively (Table 3. V – A). Accordingly, if we examine their type I errors, we conclude that R_1's performance is better than R_0's under each simulated situations. Also both tests are completely approximated by three-moment chi – square method as they belong to the Type III region (Table 3. V – B). Lastly as seen in Table 3. V – C, R_1 is more powerful than R_0 mainly near unit root cases (ϕ_1=0.99, ϕ_1=0.95). In spite of the superiority of R_1 for small sample sizes like 20 or 30, R_0 catches its power when we extend the situation from these near root values to ϕ_1=0.80, ϕ_1=0.70 and so on.

- The last situation, which we analyse under various sample sizes and estimated parameters, is related to the generalized logistic distribution. As the MML estimators give more accurate results than the LS ones; that is they are efficient and almost unbiased for every specification (Table 3. VI –

A), the R_1 test, based on these estimators, are more powerful than R_0 test which depends on the LS estimators (Table 3.VI–C).Despite this fact both have nearly same type I errors and their probability approximation belong to the Type III region. Thus three - moment chi – square approximation is preferable in these cases (Table 3. VI – B).

As a result it is concluded that similar to the AR (1) without intercept term, the parameters of this new model also are estimated by the MML estimators more efficiently. Therefore, unit root test based on these estimators is powerful than its alternatives and especially three – moment chi – square approximation is very helpful while determining the probability values of these test statistics.

3.3. Outliers and Mixtures

3.3.1. Outlier Model

To investigate the performances of the MMLE and LSE, and accordingly relevant R unit root tests when the errors have r (r=0.5+0.1n) number of outliers, we generate data from AR (1) model without intercept term as n=50, and AR (1) model with intercept term as n=100. Both errors are chosen from the long – tailed symmetric distribution (LTS) but have mean zero, variance firstly taken as $4\sigma^2$ and then $16\sigma^2$ to obtained an outlier model. Later their simulated values are compared with respect to the first four – moment of ϕ_1, type I error and powers of R, similar to the other cases.

So based on the tabulated results of the two models, the MML estimators are more efficient than LS estimators, and have smaller bias with respect to LSE under ϕ_1=1 and p =2.5, 3.0, 5.0, 10.0 (Table 3. VII – A, Table

3. VIII - A). The only difference among models is that the intercepted model has higher efficiency regarding to the pure random walk model. Both estimators have non - normal skewness and kurtosis, as it should be. On the other hand R_1 test has smaller type I error generally (Table 3. VII – B, Table 3. VIII – B). In the same way both three – moment chi – square approximation and four – moment F approximation are still applicable.

Finally while checking their powers we observe that in pure AR (1) model, R_1 has high power with regard to the R_0 (Table 3. VII – C) starting from near unit root cases. However in intercepted model the power of R_1 is small for n=100 and less, and we deal with near unit root cases. But its power increases when ϕ_1 is 0.85 or small for selected sample size, and maintains high power in all cases if the sample size increases (Table3. VIII – C).

3.3.2. Mixture Model

While constructing a mixture model, we select π (the probability of being an outlier) value as 0.10 and combine two LTS distributions having different variances like the following form:

$$\pi LTS(p, k\sigma) + (1-\pi) LTS(p, \sigma) \quad ; \; k = 2,4. \tag{3.15}$$

From the simulations, it is seen that under both AR (1) with / without constant term, the MMLE are more efficient and have smaller bias than the LSE, similar to the other results (Table 3. IX – A, Table 3. X – A). Furthermore R_1 and R_0 tests have almost same type I error (Table 3. IX – B, Table 3. X – B), and different from the outlier model, all probabilities can be computed by four – moment F approximation.

At the end, based on the comparison of their powers, R_1 performs

higher power than R_0 (Table 3. IX – C, Table 3. X – C) under all specified situations and reaches 1 quicker than the other test in both AR (1).

If we extend our model as a contaminated one having the structure below:

$$\pi U(-0.5,0.5)+(1-\pi)LTS(p,\sigma), \tag{3.16}$$

we notice that our conclusions related to the mixture models are completely valid also here (Table 3. XI - A, B, C and Table 3. XII - A, B, C).

Consequently it is noticed that even we have an outlier, a mixture or a contamination model, the MMLE results are more efficient and have smaller bias compared to the LSE. R_1 test shows higher power considering R_0 test and as n increases, the power of R_1 increases too faster than the other.

CHAPTER IV

SUMMARY AND CONCLUSION

In time series data, stationarity means that the means, variances and covariances of the process are constant through time. This might not be true in certain situations. In particular, nonstationarity might be caused by the existence of a unit root.

Testing unit root is, therefore, a very important problem. From practical point of view, it implies that an innovation (shock) has permanent effect through time; that is, the effect of an innovation applied to the current value does not disappear as time passes. In this study we are interested in only simple AR (1) models.

Many unit root tests have been proposed in the literature for this model. The main difficulty with these tests is that as the exact distribution of the sample autocorrelation coefficient is intractable, the usual t – statistic is not applicable either. For that reason attempts have been made to find at least the asymptotic distribution of the sample autocorrelation coefficient $\tilde{\phi}_1$ under the assumption of normality

Dickey – Fuller (1979) and Phillips – Perron (1988) used the Wiener process to find the asymptotic distribution of $\tilde{\phi}_1$ when $\phi_1 = 1$. This process is like a continuous random walk defined on the interval $[0.1]$, but has unbounded variation.

Abadir (1995) used the quadratic forms of the usual t^2 – ratio of the autoregressive coefficient under nonstationarity. He utilized certain transformation techniques and integral representation of the so - called parabolic cyclinder.

In 1996, Shenton and Vinod evaluated the first four – moments of the maximum likelihood estimator for AR (1) and random walk model without intercept term by using q – point Gaussian quadrature methods assuming normality of course.

The common property of all these works up to now is that they are based on the errors assumed to be independent normal variates. The initial value is taken as $y_0 = 0$ or $y_0 \sim N\left(0, \sigma^2 / \left(1 - \phi_1^{\,2}\right)\right)$ when $|\phi_1| < 1$. However, in practice, non – normal distributions occur so often. Also, samples often contain outliers. In such situations, the least square estimators (which are used generally in time series) lose their efficiency and tests based on them lose power. As a result we need to use a robust estimator, which is fully efficient (or nearly so) for an assumed distribution and maintains high efficiency for plausible alternatives to the assumed distribution. In our study, therefore, we use modified maximum likelihood estimators, which are known to be efficient and robust. Such estimators originated with Tiku (1967).

For various type of distributions; e. g. long - tailed symmetric, gamma and generalized logistic distributions, the unknown parameters μ, ϕ_1 and σ in AR (1) models are estimated by the MML and LS methods and calculated by using 10000 Monte Carlo runs for n ≤ 100 and 5000 runs for $n > 100$. Without loss of generality, the standard deviation σ and the constant term μ (intercept) are taken as 1 and 0, respectively. It is seen that the MML results are always considerably more efficient than the LS results and have mostly negligible bias for different shape parameters in the underlying distribution.

Furthermore, the results of the LSE and the MMLE based on the outlier , mixture and contamination AR (1) models are compared. Here, it is shown that the MMLE again are more efficient although the two estimates are numerically close to one – another. We made this comparison with real life data also and get similar results (Special Appendix – Applications).

Tiku and Wong (1998) proposed new test statistics R_1 (based on the MML estimators) and R_0 (based on the LS estimators) for testing the unit root. They evaluated their Pearson coefficients of skewness and kurtosis for a family of symmetric distributions. In this thesis, we extend their results to skew distributions, e.g., gamma and generalized logistic. We show that the R_1 test is considerably more powerful than the R_0 test. The null distributions ($\phi_1 = 1$) of the R_1 and R_0 statistics are approximated by three – moment chi – square and four – moment F distributions. The chi – square and F distributions provide remarkable accurate approximations.

REFERENCES

1. Abadir, K. M. (1993), "On the Asymptotic Power of Unit Root Tests", Econometric Theory, 9, 189 - 221.

2. Abadir, K. M. (1995), "The Limiting Distribution of the t Ratio Under a Unit Root", Econometric Theory, 11, 775 – 793.

3. Akkaya, A. D., Tiku, M. L. (2001), "Corrigendum: Time Series Models with Asymmetric Innovations", Communication in Statistics – Theory and Methods, 30 (10), 2227 – 2230.

4. Akkaya, A. D., Tiku, M. L. (2001), "Estimating Parameters in Autoregressive Models in Non - normal Situations; Asymmetric Innovations", Communication in Statistics – Theory and Method, 30 (3), 517 – 536.

5. Barnett, V. (1966), "Order Statistics of the Location of the Cauchy Distribution", Journal of American Statistical Association, 61, 1205 – 1218.

6. Bhattacharya, G. K. (1985), "The Asymptotic of the Maximum Likelihood and Related Estimators Based on Type II Censored Data", Journal of American Statistical Association, 80, 398 – 404.

7. Cryer, J. (1986), Time Series Analysis, PWS – KENT Publishing, Boston.

8. DeJong, D. N., Nankervis, J. C., Savin, N. E., Whiteman, C. H. (1992), "The Power Problems of Unit Root Tests in Time Series with Autoregressive Errors", Journal of Econometrics, 53, 323 – 343.

9. Dickey, D. A., Fuller, W. A. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", Journal of American Statistical Association, 74, 427 – 431.

10. Dickey, D. A., Pantula, S. G. (1987), "Determining the Order of Differencing in Autoregressive Processes", Journal of Business and Economic Statistics, 5, 453 – 461.

11. Elderson, W. P., Johnson, N. L. (1969), Systems of the Frequency Curves, Cambridge University Press, Aberdeen.

12. Enders, W. (1995), Applied Econometric Time Series, John Wiley & Sons, Inc, Canada.

13. Greene, W. H. (1997), Econometric Analysis, Prentice – Hall, Inc, New Jersey.

14. Gujarati, D. N. (2003), Basic Econometrics, McGraw – Hall Companies, Inc., New York.

15. Hamilton, J. D. (1994), Time Series Analysis, Princeton University Press, New Jersey.

16. Hamilton, L. C. (1992), Regressions With Graphics, California: Brooks / Cole Publishing Company, California.

17. Harvey, A. C. (1993), Time Series Models, P & R Typesetters Ltd, Salisbury, Wiltshire.

18. Kendall, M. G., Stuart A. (1977), The Advanced Theory of Statistics, Volume I, Charles Griffin & Company Limited, London.

19. Kwiatkowski, D., Phillips, P. C. B., Schmidth, P., Shin, Y. (1992), "Testing the Null Hypothesis of Stationary Against the Alternatives of a Unit Root", Journal of Econometrics, 54, 159 – 178.

20. Maddala, G. S. (2001), Introduction to Econometrics, John Wiley & Sons Ltd, Canada.

21. Öcal, N. (1992), "Unit Root Tests and Their Application to Certain Macroeconometric Series of Turkey, M. S. Thesis.

22. Pearson, E. S. (1958), "Table of Percentage Points of Pearson Curves, for Given $\sqrt{\beta_1}$ and β_2, Expressed in Standard Measure", Biometrika, 50, 459 – 498.

23. Pearson, E. S., Tiku, M. L. (1970), "Some Notes on the Relationship Between the Distributions of Central and Non – central F", Biometrika, 57, 175 – 179.

24. Phillips, P. C. B., Perron, P. (1988), "Testing for Unit Root in Time Series Regression", Biometrika, 75, 335 – 346.

25. Puthenpura, S., Sinha, N. K. (1986), "Modified Maximum Likelihood Method for the Robust Estimation of System Parameters from Very Noisy Data", Automatica, 22, 231 – 235.

26. Ross, J. S. (1990), Non-linear Estimation, New York, Springer.

27. Sawa, T. (1978), 'The Exact Moments of the Least Squares Estimators for the Autoregressive and Random Walk Models for a Zero or Stationary Initial Value", Econometric Theory, 12, 481 – 499.

28. Srivastava M. S. (2002), Methods of Multivariate Statistics, John Wiley & Sons, Inc., New York.

29. Stewart, J., Gill, L. (1998), Econometrics, Biddles Ltd, Oxford

30. Şenoğlu, B., Tiku, M. L. (2002), "Linear Contrasts in Experimental Design with Non – Identical Error Distribution", Biometrical Journal, 44 (3), 359 - 374

31. Tiku, M. L., (1966), 'Usefulness of Three – Moment χ^2 and t Approximations", Journal of Indian Society of Agricultural Statistics", 18.

32. Tiku, M. L. (1967), 'Estimating the Mean and Standard Deviation from a Censored Normal Samples", Biometrika, 54, 155 – 165.

33. Tiku, M. L. (1968), 'Estimating the Parameters of Normal and Logistic Distributions from Censored Samples", Austral Journal of Statisitcs, 10, 64 – 74.

34. Tiku, M. L., Yip, D. Y. N. (1978), A Four – moment Approximation Based on the F Distribution", 20 (3), 257 – 261.

35. Tiku, M. L., Tan, W. Y., Balakrishnan, N. (1986), Robust Inference, New York, Marcel Dekker.

36. Tiku, M. L., Suresh, R. P. (1992), "A New Method of Estimation for Location and Scale Parameters", Journal of Statistical Planning and Inference", 30, 281 – 292.

37. Tiku, M. L., Wong, W. K., Vaughan, D. C., Bian, G. (1996), "Time Series Models in Non – normal Situations: Symmetric Innovations", Journal of Time Series Analysis, 21 (5), 571 – 596.

38. Tiku, M. L., Wong, W. K. (1998), "Testing for a Unit Root in an AR (1) Model Using Three and Four Moment Approximations: Symmetric Distributions", Communication in Statistics – Simulations, 27 (1), 185 – 198.

39. Tiku, M. L., Wong, W. K., Bian G. (1999), "Estimating Parameters in Autoregressive Models in Non – normal Situations: Symmetric Innovations", Communication in Statistics – Theory and Methods, 28 (2), 315 – 341.

40. Tiku, M. L., Wong, W. K., Bian, G. (1999), "Time Series Models With Asymmetric Innovations", Communication in Statistics – Theory and Methods, 28 (6), 1331 – 1360.

41. Türker, Ö. (1999), "Structural Time Series Modelling of Inflation Behaviour in Turkey", M. S. Thesis.

42. Vaughan, D. C. (1992a), "On the Tiku – Suresh Method of Estimation", Communication in Statistics – Theory and Methods, 21, 451 – 469.

43. Vinod, H. D., Shenton, L. R. (1996), "Exact Moments for Autoregressive and Random Walk Models for a Zero or Stationary Initial Value", Econometric Theory, 12, 481 – 499.

44. White, J. (1958), 'The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case", Annals of Mathematical Statistics, 29, 1188 – 1197.

SPECIAL APPENDIX

APPLICATIONS

In order to compare the MML and the LS estimates computed from real life data, we select the following examples. In all cases the data are read across each row before moving to the next row. Before making estimation, we construct Q – Q plots to identify a plausible distribution; we plot the ordered error values $\varepsilon_{(i)}$ $(1 \le i \le n)$ of a random sample of size n against the quantiles $t_{(i)}$ of normal distribution $N(0,1)$ defined as follows:

$$\int_{-\infty}^{t_{(i)}} f(z)dz = \frac{i}{n+1}\,;\ (1 \le i \le n) \qquad \text{(A. 1)}$$

If we obtain a straight - line (or closest to such), it indicates that the underlying distribution is normal or near – normal. In the absence of a straight – line, a Q – Q plot provides information about the underlying distribution, whether it is symmetric long – tailed, skewed, symmetric with outliers, etc. see Hamilton (1992, p.16).

Example 1: The following data belongs to the monthly AA railroad bond yields from January 1968 through June 1976 (Source: Cryer, J. D. (1986), Time Series Analysis, PWS – KENT Publishing Company, Boston).

Table A. 1. 1. Monthly AA Railroad Bond Yields from January 1968 through June 1976 (Source: Cryer, J. D. 1986)

639	643	640	653	667	667	663	654	649	651	659	672
670	675	692	702	706	710	722	729	740	755	763	788
818	826	821	819	827	848	881	879	878	878	868	856
844	824	820	819	813	815	822	818	815	792	769	775
771	773	780	779	774	772	775	770	766	771	773	772
767	775	777	777	776	779	787	790	791	792	802	799
792	780	790	799	810	814	828	862	874	892	872	869
870	859	857	870	867	856	854	862	861	855	846	847
845	838	828	823	814	812						

We assume an AR (1) model with an intercept term such that $y_i = \mu + \phi_1 y_{i-1} + \varepsilon_i$; $(1 \le i \le n)$. Then we use the LSE $\tilde{\mu}$ and $\tilde{\phi}_1$ instead of the unknown parameter to find $\tilde{\varepsilon}_i$. Next we plot ascending the $\tilde{\varepsilon}_{(i)}$ against $t_{(i)}$ and get the following Q – Q plot. Realize that

$$\tilde{e}_{(1)} \le \tilde{e}_{(2)} \le \ldots\ldots\ldots\ldots\ldots\ldots \le \tilde{e}_{(n)}\ ;\ (n = 102). \qquad \text{(A. 2)}$$

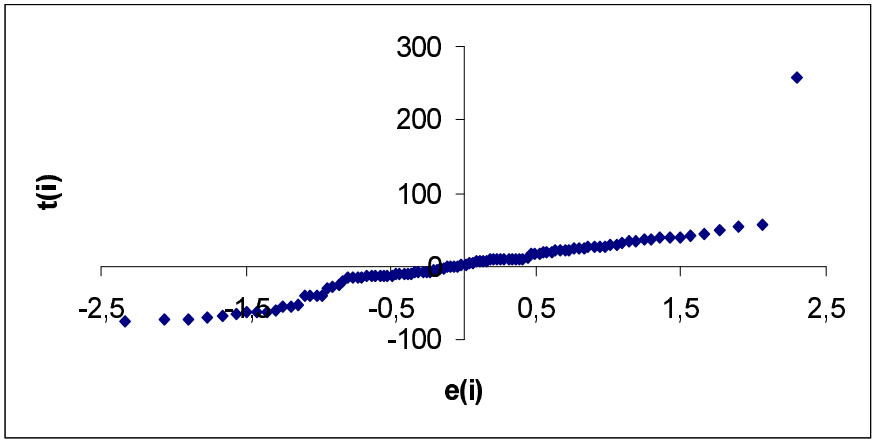

Figure A. 1. 1. Q – Q Plot for the Data Given in Example 1

From the graph, it is seen that the observations deviate from a straight - line at the tails and it is a property of long – tailed distributions. But before determining a plausible value of the shape parameter by maximazing log - likelihood function, we omit the initial observation, as it is an outlier.

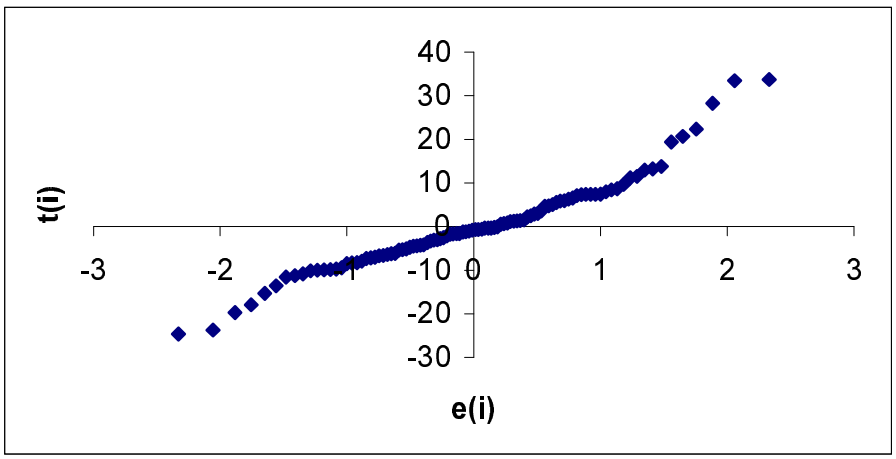

Figure A. 1. 2. Q – Q Plot for the Data Given in Example 1 After Omitting Outlier

The values of $(1/n)InL$ are given below,

$$L = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{k}\,\beta\left(\frac{1}{2}, p-\frac{1}{2}\right)} \left(1+\frac{\varepsilon_i^{\,2}}{k\sigma^2}\right)^{-p} ;\ k = 2p-3\ \ (p \geq 2) \qquad \text{(A. 3)}$$

Table A. 1. 2. The Values of $(1/n)InL$ for the Data Given in Table A. 1. 1. After Omitting Outlier

	$p = 2.0$	$p = 2.5$	$p = 3.0$	$p = 3.5$	$p = 5.0$	$p = 10.0$
$(1/n)InL$	-3.6814	-3.6807 *	-3.683	-3.687	-3.696	-3.712

* The maximum of log – likelihood value

As the maximum of log - likelihood is attained when $p = 2.5$, we select this value of the shape parameter to compute the MML and LS estimates:

parameter	*MML*	*LS*
μ	29.615	30.336
ϕ_1	0.964	0.964
σ	10.548	10.217

(A.4)

It is found that in both estimates, a unit root $\phi_1 = 1$ is indicated. Thus in order to get a stationary model, we take first difference of the series. Then

we obtain the following new estimates of ϕ_1:

$$\hat{\phi}_1 = 0.424 \text{ and } \tilde{\phi}_1 = 0.469 \qquad \text{(A. 5)}$$

It is interesting to see that the estimates are close to one – another in both cases. However, the relative efficiencies of the LSE are $RE(\tilde{\phi}_1)$ =70 percent and $RE(\tilde{\sigma})$ =0 (because of infinite variance of the LSE $\tilde{\sigma}$). The relative efficiencies are found by using the following equations:

$$RE(\tilde{\phi}_1) = 100 \frac{(p+1)\left(p - \frac{3}{2}\right)}{p\left(p - \frac{1}{2}\right)} \qquad \text{(A. 6)}$$

and

$$RE(\tilde{\sigma}) = 100 \frac{(p+1)}{\left(p - \frac{1}{2}\right)\left(1 + \frac{1}{2}\lambda_4\right)} \; ; \qquad \text{(A. 7)}$$

where $\lambda_4 = \beta_2 - 3$; $\beta_2 = 3\frac{(2p-3)}{(2p-5)}$.

Therefore, we conclude that the two estimates are close to one – another but the LSE are considerably less efficient.

Example 2: The following data belongs to the monthly U.S. air passenger miles from January 1960 through December 1977 (n=216) (Source:

Cryer, J. D. (1986), Time Series Analysis, PWS – KENT Publishing Company, Boston).

Table A. 2. 1. Monthly U.S. Air Passenger Miles from January 1960 through December 1977 (Source: Cryer, J. D. 1986)

2.42	2.14	2.28	2.50	2.44	2.72	2.71	2.74	2.55	2.49	2.13	2.28
2.35	1.82	2.40	2.46	2.38	2.83	2.68	2.81	2.54	2.54	2.37	2.54
2.62	2.34	2.68	2.75	2.66	2.96	2.66	2.93	2.70	2.65	2.46	2.59
2.75	2.45	2.85	2.99	2.89	3.43	3.25	3.59	3.12	3.16	2.86	3.22
3.24	2.95	3.32	3.29	3.32	3.91	3.80	4.02	3.53	3.61	3.22	3.67
3.75	3.25	3.70	3.98	3.88	4.47	4.60	4.90	4.20	4.20	3.80	4.50
4.40	4.00	4.70	5.10	4.90	5.70	3.90	4.20	5.10	5.00	4.70	5.50
5.30	4.60	5.90	5.50	5.40	6.70	6.80	7.40	6.00	5.80	5.50	6.40
6.20	5.70	6.40	6.70	6.30	7.80	7.60	8.60	6.60	6.50	6.00	7.60
7.00	6.00	7.10	7.40	7.20	8.40	8.50	9.40	7.10	7.00	6.60	8.00
10.45	8.81	10.61	9.97	10.69	12.40	13.38	14.31	10.90	9.98	9.20	10.94
10.53	9.06	10.17	11.17	10.84	12.09	13.66	14.06	11.14	11.10	10.00	11.98
11.74	10.27	12.05	12.27	12.03	13.95	15.10	15.65	12.47	12.29	11.52	13.08
12.50	11.05	12.94	13.24	13.16	14.95	16.00	16.98	13.15	12.88	11.99	13.13
12.99	11.69	13.78	13.70	13.57	15.12	15.55	16.73	12.68	12.65	11.18	13.27
12.64	11.01	13.30	12.19	12.91	14.90	16.10	17.30	12.90	13.36	12.26	13.93
13.94	12.75	14.19	14.67	14.66	16.21	17.72	18.15	14.19	14.33	12.99	15.19
15.09	12.94	15.46	15.39	15.34	17.02	18.85	19.49	15.61	16.16	14.84	17.04

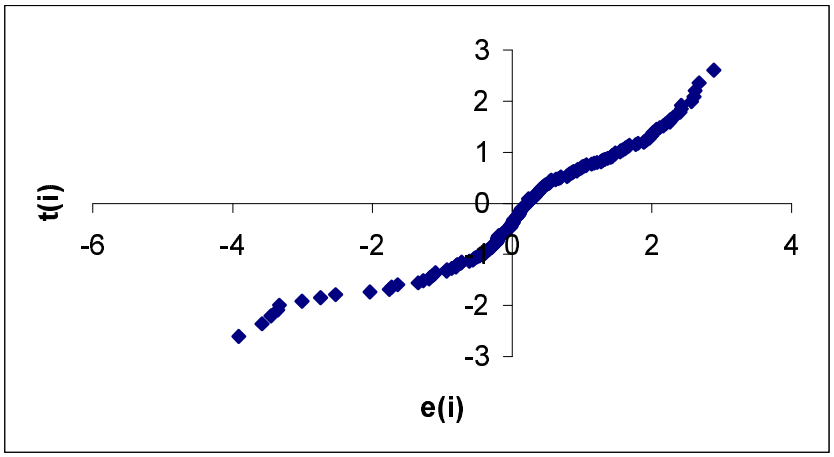

Figure A. 2. 1. Q – Q Plot for the Data Given in Example 2

From the graph, it is seen that the observations deviate from the straight - line especially towards the left tail. This indicates a negatively skewed distribution. A distribution from the generalized logistic family can be suitable. In order to specify a plausible value of the shape parameter, log - likelihood function is maximized:

$$L = \prod_{i=1}^{n} \frac{b}{\sigma} \frac{\exp(-\varepsilon_i / \sigma)}{\{1 + \exp(-\varepsilon_i / \sigma)\}^{b+1}}; \; b > 0 \qquad \text{(A. 8)}$$

The values of $(1/n)InL$ are given below (Table A. 2. 2.),

Table A. 2. 2. The Values of $(1/n)InL$ for the Data Given in Example 2

	$b=0.1$	$b=0.2$	$b=0.5$	$b=1.0$	$b=2.0$
$(1/n)InL$	2.994 *	1.678	-0.120	-1.528	-2.966

*The maximum of log – likelihood value.

The maximum of log - likelihood is attained when $b=0.1$. The resulting estimates are given below:

parameter	*MML*	*LS*
μ	0.648	1.453
ϕ_1	1.101	0.972
σ	0.176	0.114

(A. 9)

It is clear that again we have a random walk model. Therefore we take first difference of the series to make it stationary. Thus we get the following new stationary ϕ_1 values:

$$\hat{\phi}_1 = \text{-0.279 and } \tilde{\phi}_1 = \text{-0.202} \tag{A. 10}$$

Here also both estimates are reasonably close to each other. The relative efficiencies of the LSE are

$$RE(\tilde{\phi}_1) = 100\frac{(b+2)}{b[\Psi'(b)+\Psi'(1)]} \tag{A. 11}$$

and

$$RE(\tilde{\sigma}) = 100 \frac{2}{\left(1+\frac{bH}{b+2}\right)\left(1+\frac{1}{2}\lambda_4\right)} \quad . \qquad \text{(A. 12)}$$

where $\lambda_4 = \beta_2 - 3$ and $H = \Psi'(b+1) + \Psi'(2) + [\Psi(b+1) - \Psi(2)]^2$.

The $RE(\tilde{\phi}_1)$ is 20.37 percent. On the other hand while computing λ_4, we need β_2 for the underlying distribution. At b=0.5, the skewness and kurtosis of the distribution are calculated as 0.731 and 5.400 respectively (Şenoğlu, Tiku 2002). Accordingly the kurtosis at b=0.1 is greater than the kurtosis at b=0.5. Therefore the relative efficiency at this value should be less than the relative efficiency at b=0.5. For b=0.5, $RE(\tilde{\sigma})$= 67.55. It is concluded that at b=0.1, $RE(\tilde{\sigma})$< 67.55 percent.

We conclude that the MMLE are enormously more efficient than the LSE although both of them are numerically close to one – another.

Example 3: The following data belongs to the Portland, Oregon average monthly bus ridership (/100) from January 1973 through June 1973 (n=114) (Source: Cryer, J. D. (1986), Time Series Analysis, PWS – KENT Publishing Company, Boston).

Table A. 3. 1. Portland, Oregon Average Monthly Bus Ridership (/100) from January 1973 through June 1973 (Source: Cryer, J. D. 1986)

640	646	639	654	630	622	617	613	661	695	690	707
817	839	810	789	760	724	704	691	745	803	780	761
857	907	873	910	900	880	867	854	928	1064	1103	1026
1102	1080	1034	1083	1078	1020	984	952	1033	1114	1160	1058
1209	1200	1130	1182	1152	1116	1098	1044	1142	1222	1234	1155
1286	1281	1224	1280	1228	1181	1156	1124	1152	1205	1260	1188
1212	1269	1246	1299	1284	1345	1341	1308	1448	1454	1467	1431
1510	1558	1536	1523	1492	1437	1365	1310	1441	1450	1424	1360
1429	1440	1414	1424	1408	1337	1258	1214	1326	1417	1417	1329
1461	1425	1419	1432	1394	1327						

Firstly Q – Q plot of the data is observed.

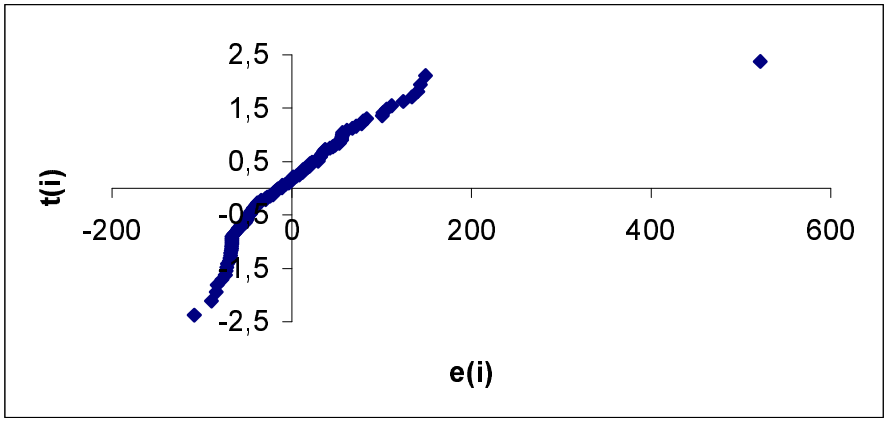

Figure A. 3. 1. Q – Q Plot for the Data Given in Example 3

One residual is clearly an outlier. We set it aside. The Q – Q plot of the remaining n=113 ordered residuals $\tilde{e}_{(i)}$ indicates a normal distribution, may be a symmetric long – tailed distribution. Proceeding as in Example 1, we have the following values of $(1/n)InL$ (Table A. 3. 2.):

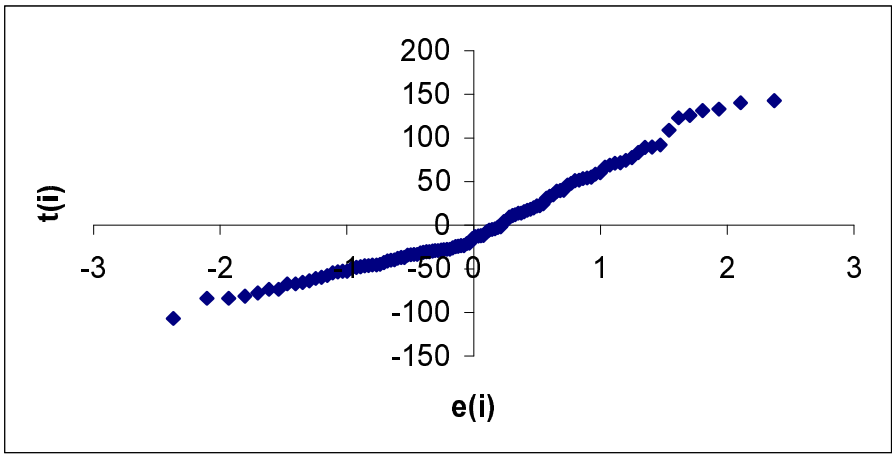

Figure A. 3. 2. Q – Q Plot for the Data Given in Example 3 After Omitting Outlier

Table A. 3. 2. The Values of $(1/n)InL$ for the Data Given in Table A. 3. 1. After Omitting Outlier

	$p=3.5$	$p=5.0$	$p=7.0$	$p=10.0$	$p=15.0$	$p=20.0$
$(1/n)InL$	-5.462	-5.456	-5.452	-5.450	-5.449	-5.448*

*The maximum of log – likelihood value.

The value $p = 20.0$ maximizes InL. The corresponding estimates are

parameter	*MML*	*LS*
μ	43.779	44.726
ϕ_1	0.966	0.966
σ	57.058	56.647

(A. 13)

As ϕ_1 values are almost equal to 1, the first difference of the data is taken to attain a stationary model. Consequently these new autocorrelation coefficients are calculated as:

$$\hat{\phi}_1 = \text{-0.050 and } \tilde{\phi}_1 = \text{-0.060} \qquad \text{(A. 14)}$$

The relative efficiencies of $\tilde{\phi}_1$ and $\tilde{\sigma}$ are $RE(\tilde{\phi}_1)$=99.62 and $RE(\tilde{\sigma})$=99 respectively. In fact these are expected results because $p = 20.0$ implies that the underlying distribution is very close to normal.

APPENDIX A

When the error terms have location – scale distribution of symmetric family, Fisher information matrix is calculated as follows:

$$\frac{1}{n}I_{11} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1^2}\right\} = \frac{2p}{k\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{1-\frac{1}{k}z_t^2}{\left[1+\frac{1}{k}z_t^2\right]^2}y_{t-1}^2\right\}$$

$$= \frac{2p}{k\sigma^2}\frac{\sigma^2}{1-\phi_1^2}E\left\{\frac{1-\frac{1}{k}z^2}{\left(1+\frac{1}{k}z^2\right)^2}\right\}$$

$$= \frac{2p}{k}\frac{1}{1-\phi_1^2}\int_{-\infty}^{\infty}\left\{2\left(1+\frac{1}{k}z^2\right)^{-2}-\left(1+\frac{1}{k}z^2\right)^{-1}\right\}f(z)dz$$

$$= \frac{2p}{k}\frac{1}{1-\phi_1^2}\left(\frac{2\left(p+\frac{1}{2}\right)\left(p-\frac{1}{2}\right)}{p(p+1)}-\frac{\left(p-\frac{1}{2}\right)}{p}\right)$$

$$= \frac{2p}{k}\frac{1}{1-\phi_1^2}\frac{\left(p-\frac{1}{2}\right)}{(p+1)}$$

realizing that

$$\int_{-\infty}^{\infty}\left\{1+\frac{1}{k}z^2\right\}^{-j}dz=\sqrt{k}\,\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(j-\frac{1}{2}\right)}{\Gamma(j)}$$

for $j \geq 1$ and

$$\frac{\partial^2 InL}{\partial\phi_1^2}=-\frac{2p}{k\sigma^2}\sum_{t=1}^{n}\frac{1-\frac{1}{k}z_t^2}{\left(1+\frac{1}{k}z_t^2\right)^2}y_{t-1}^2$$

$$z_t=\frac{y_t-\phi_1 y_{t-1}}{\sigma}.$$

$$\frac{1}{n}I_{12}=\frac{1}{n}I_{21}=-\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1\partial\sigma}\right\}=\frac{2p}{k\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{\frac{2}{k}z_t}{\left(1+\frac{1}{k}z_t^2\right)^2}y_{t-1}\right\}$$

$$=\frac{2p}{k}\frac{1}{1-\phi_1^2}E\left\{\frac{\frac{2}{k}z}{\left(1+\frac{1}{k}z^2\right)^2}\right\}$$

$$=\frac{2p}{k}\frac{1}{1-\phi_1^2}\int_{-\infty}^{\infty}\frac{2}{k}\left[\frac{z}{\left(1+\frac{z^2}{k}\right)^2}\right]f(z)dz$$

$$= \frac{2p}{k}\frac{1}{1-\phi_1^2}\left[\left(1+\frac{z^2}{k}\right)^{-p-1}\right]_{-\infty}^{\infty} = 0$$

since y_{t-1} and z_t are independent and

$$\frac{\partial^2 Inl}{\partial\phi_1\partial\sigma} = -\frac{2p}{k\sigma^2}\sum_{t=1}^{n}\frac{\frac{2}{k}z_t}{\left(1+\frac{z_t^2}{k}\right)^2}y_{t-1}\ .$$

$$\frac{1}{n}I_{22} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\sigma^2}\right\} = -\frac{1}{\sigma^2}+\frac{2p}{k\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{z_t^2}{\left(1+\frac{z_t^2}{k}\right)^2}\left(3+\frac{z_t^2}{k}\right)\right\}$$

$$= -\frac{1}{\sigma^2}+\frac{2p}{k\sigma^2}E\left\{\frac{z^2}{\left(1+\frac{z^2}{k}\right)^2}\left(3+\frac{z^2}{k}\right)\right\}$$

$$= -\frac{1}{\sigma^2}+\frac{2p}{k\sigma^2}\int_{-\infty}^{\infty}\left\{\left(1+\frac{z^2}{k}\right)^{-1}-2\left(1+\frac{z^2}{k}\right)^{-2}\right\}f(z)dz$$

$$= -\frac{1}{\sigma^2}+\frac{2p}{k\sigma^2}\left[1+\frac{\left(p-\frac{1}{2}\right)}{p}-2\frac{\left(p+\frac{1}{2}\right)\left(p-\frac{1}{2}\right)}{p(p+1)}\right]$$

$$= \frac{2\left(p - \frac{1}{2}\right)}{(p+1)\sigma^2}$$

where $$\frac{\partial^2 InL}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{2p}{k\sigma^2} \sum_{t=1}^{n} \frac{z_t^2}{\left(1 + \frac{z_t^2}{k}\right)^2} \left(3 + \frac{z_t^2}{k}\right).$$

APPENDIX B

When the error terms in AR (1) model without intercept have gamma distribution, Fisher information matrix is computed as the following way:

$$\frac{1}{n}I_{11} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1^2}\right\} = \frac{(k-1)}{\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{z_t^2}y_{t-1}^2\right\}$$

$$= \frac{(k-1)}{\sigma^2}E\left\{\frac{y^2}{z^2}\right\}$$

$$= \frac{(k-1)}{\sigma^2}\sigma^2 k\left(\frac{1}{1-\phi_1^2} + \frac{k}{(1-\phi_1)^2}\right)\frac{1}{(k-2)(k-1)}$$

$$= \frac{k}{(k-2)}\left(\frac{1}{1-\phi_1^2} + \frac{k}{(1-\phi_1)^2}\right)$$

where $\frac{\partial^2 InL}{\partial\phi_1^2} = -\frac{(k-1)}{\sigma^2}\sum_{t=1}^{n}\frac{1}{z_t^2}y_{t-1}^2$,

$$E(y^2) = V(y) + [E(y)]^2,$$

$$= \frac{\sigma^2 k}{1-\phi_1^2} + \frac{k^2\sigma^2}{(1-\phi_1)^2}$$

$$= \sigma^2 k\left(\frac{1}{1-\phi_1^2} + \frac{k}{(1-\phi_1)^2}\right).$$

$$\frac{1}{n} I_{12} = \frac{1}{n} I_{21} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\phi_1 \partial\sigma}\right\} = \frac{1}{\sigma} E\left\{\frac{1}{n}\sum_{t=1}^{n} y_{t-1}\right\}$$

$$= \frac{1}{\sigma} E(y) = \frac{k}{(1-\phi_1)\sigma}$$

where $$\frac{\partial^2 InL}{\partial\phi_1 \partial\sigma} = -\frac{\sum_{t=1}^{n} y_{t-1}}{\sigma},$$

$$E(y) = \frac{k}{1-\phi_1}.$$

$$\frac{1}{n} I_{22} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\sigma^2}\right\} = -\frac{k}{\sigma^2} + 2E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{y_t - \phi_1 y_{t-1}}{\sigma^3}\right\}$$

$$= -\frac{k}{\sigma^2} + 2\frac{E(z)}{\sigma^2} = -\frac{k}{\sigma^2} + 2\frac{k}{\sigma^2} = \frac{k}{\sigma^2}$$

where $$\frac{\partial^2 InL}{\partial\sigma^2} = \frac{nk}{\sigma^2} - \frac{2}{\sigma^2}\sum_{t=1}^{n} \frac{y_t - \phi_1 y_{t-1}}{\sigma}.$$

APPENDIX C

While the error terms belong to the generalized logistic distribution, Fisher information matrix is found that;

$$\frac{1}{n}I_{11} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1^2}\right\} = \frac{b+1}{\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{e^{-z_t}}{(1+e^{-z_t})^2}y_{t-1}^2\right\}$$

$$= \frac{b+1}{\sigma^2}E\left(y^2\right)E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\}$$

$$= \frac{b}{(b+2)\sigma^2}\left\{\frac{\Psi'(b)+\Psi'(1)}{1-\varphi_1^2} + \frac{\left[\Psi(b)-\Psi(1)\right]^2}{(1-\phi_1)^2}\right\}$$

where $E\left(y^2\right) = \left\{\frac{\Psi'(b)+\Psi'(1)}{1-\phi_1^2} + \frac{\left[\Psi(b)-\Psi(1)\right]^2}{(1-\phi_1)^2}\right\}\sigma^2$,

$$\frac{\partial^2 InL}{\partial\phi_1^2} = -\frac{b+1}{\sigma^2}\sum_{t=1}^{n}y_{t-1}^2\frac{e^{-z_t}}{(1+e^{-z_t})^2}.$$

$$\frac{1}{n}I_{12} = \frac{1}{n}I_{21} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1\partial\sigma}\right\}$$

$$= \frac{1}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} y_{t-1}\right\} + \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{e^{-z_t} z_t}{(1+e^{-z_t})^2} y_{t-1}\right\}$$

$$- \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{e^{-z_t}}{1+e^{-z_t}} y_{t-1}\right\}$$

$$= E(y)\left\{\frac{1}{\sigma^2} + \frac{b}{(b+2)\sigma^2} - \frac{b+1}{\sigma^2}\frac{1}{b+1}\right\}$$

$$= E(y)\frac{b}{(b+2)\sigma^2}$$

$$= \frac{[\Psi(b) - \Psi(1)]\sigma}{1-\phi_1}\frac{b}{(b+2)\sigma^2}$$

$$= \frac{\Psi(b) - \Psi(1)}{1-\phi_1}\frac{b}{(b+2)\sigma}$$

where $$\frac{\partial^2 InL}{\partial\phi_1\partial\sigma} = -\frac{\sum_{t=1}^{n} y_{t-1}}{\sigma^2} + \frac{(b+1)}{\sigma^2}\sum_{t=1}^{n} y_{t-1}\frac{e^{-z_t} - e^{-z_t} z_t + e^{-2z_t}}{(1+e^{-z_t})^2},$$

$$E(y) = \frac{[\Psi(b) - \Psi(1)]\sigma}{1-\phi_1}.$$

$$\frac{1}{n} I_{22} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\sigma^2}\right\} = \frac{1}{\sigma^2} - \frac{2}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} z_t\right\} + \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{(e^{-z_t} z_t^2 - 2e^{-z_t} z_t)(1+e^{-z_t}) + e^{-2z_t} z_t^2}{(1+e^{-z_t})^2}\right\}$$

$$=\frac{1}{\sigma^2}-\frac{2}{\sigma^2}E(z)+\frac{(b+1)}{\sigma^2}E\left\{\frac{z^2e^{-z}}{(1+e^{-z})^2}\right\}-2\frac{(b+1)}{\sigma^2}E\left\{\frac{ze^{-z}}{1+e^{-z}}\right\}$$

$$=\frac{1}{\sigma^2}-2\frac{\Psi(b)-\Psi(1)}{\sigma^2}+\frac{(b+1)}{\sigma^2}\frac{b}{(b+1)(b+2)}\left\{\Psi'(b+1)+\Psi'(2)+\left[\Psi(b+1)-\Psi(2)\right]^2\right\}$$

$$-2\frac{(b+1)}{\sigma^2}\frac{1}{(b+1)}\left[\Psi(b)-\Psi(2)\right]$$

$$=\frac{1}{\sigma^2}+\frac{b}{(b+2)\sigma^2}\left\{\Psi'(b+1)+\Psi'(2)+\left[\Psi(b+1)-\Psi(2)\right]^2\right\}$$

where

$$\frac{\partial^2 InL}{\partial\sigma^2}=-\frac{n}{\sigma^2}+2\frac{\sum_{t=1}^{n}z_t}{\sigma^2}-\frac{(b+1)}{\sigma^2}\sum_{t=1}^{n}\frac{\left(e^{-z_t}z_t^2-2e^{-z_t}z_t\right)(1+e^{-z_t})+e^{-2z_t}z_t^2}{(1+e^{-z_t})^2}.$$

APPENDIX D

When the error terms comes from in AR (1) model with intercept term and have a long – tailed symmetric distribution, Fisher information matrix is constructed as the following way:

$$\frac{1}{n} I_{11} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial \mu^2}\right\} = \frac{2p}{k\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{1-\frac{z_t^2}{k}}{\left[1+\frac{z_t^2}{k}\right]^2}\right\}$$

$$= \frac{2p}{k\sigma^2} E\left\{\frac{1-\frac{z^2}{k}}{\left[1+\frac{z^2}{k}\right]^2}\right\}$$

$$= \frac{2p}{k\sigma^2}\int_{-\infty}^{\infty}\left\{2\left(1+\frac{1}{k}z^2\right)^{-2} - \left(1+\frac{1}{k}z^2\right)^{-1}\right\} f(z)dz$$

$$= \frac{2p}{k\sigma^2}\left(\frac{2\left(p+\frac{1}{2}\right)\left(p-\frac{1}{2}\right)}{p(p+1)} - \frac{\left(p-\frac{1}{2}\right)}{p}\right)$$

$$= \frac{2p}{k\sigma^2} \frac{\left(p - \frac{1}{2}\right)}{(p+1)} = \frac{p\left(p - \frac{1}{2}\right)}{\left(p - \frac{3}{2}\right)(p+1)\sigma^2}$$

where *k=2p-3* realizing that

$$\int_{-\infty}^{\infty} \left\{1 + \frac{1}{k} z^2\right\}^{-j} dz = \sqrt{k} \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(j - \frac{1}{2}\right)}{\Gamma(j)}$$

for $j \geq 1$ as it is given in Appendix A, and

$$\frac{\partial^2 InL}{\partial \phi_1^2} = -\frac{2p}{k\sigma^2} \sum_{t=1}^{n} \frac{1 - \frac{1}{k} z_t^2}{\left(1 + \frac{1}{k} z_t^2\right)^2}$$

$$z_t = \frac{y_t - \phi_1 y_{t-1} - \mu}{\sigma}.$$

$$\frac{1}{n} I_{12} = \frac{1}{n} I_{21} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial \mu \partial \phi_1}\right\} = \frac{2p}{k\sigma^2} E\left\{\frac{1}{n} \sum_{t=1}^{n} \frac{1 - \frac{1}{k} z_t^2}{\left[1 + \frac{1}{k} z_t^2\right]^2} y_{t-1}\right\}$$

$$= \frac{2p}{k\sigma^2} \frac{\mu}{1 - \phi_1} E\left\{\frac{1 - \frac{1}{k} z^2}{\left(1 + \frac{1}{k} z^2\right)^2}\right\}$$

$$= \frac{\mu}{(1-\phi_1)} \frac{2p}{k\sigma^2} \frac{\left(p-\frac{1}{2}\right)}{(p+1)} = \frac{\mu}{(1-\phi_1)} \frac{p\left(p-\frac{1}{2}\right)}{\left(p-\frac{3}{2}\right)(p+1)\sigma^2}$$

since

$$\frac{\partial^2 InL}{\partial\mu\partial\phi_1} = -\frac{2p}{k\sigma^2} \sum_{t=1}^{n} \frac{1-\frac{1}{k} z_t^2}{\left(1+\frac{1}{k} z_t^2\right)^2} y_{t-1} \quad ,$$

$$\frac{2p}{k\sigma^2} E\left\{\frac{1-\frac{1}{k} z^2}{\left(1+\frac{1}{k} z^2\right)^2}\right\} = \frac{np\left(p-\frac{1}{2}\right)}{\left(p-\frac{3}{2}\right)(p+1)\sigma^2}$$ as we calculate in I_{11} case , and

$$E(y) = \frac{\mu}{(1-\phi_1)} .$$

$$\frac{1}{n} I_{13} = \frac{1}{n} I_{31} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\mu\partial\sigma}\right\} = \frac{2p}{k\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{2z_t}{\left[1+\frac{z_t^2}{k}\right]^2}\right\}$$

$$= \frac{2p}{k} E\left\{\frac{2z}{\left(1+\frac{1}{k} z^2\right)^2}\right\} = 2p \int_{-\infty}^{\infty} \frac{2}{k} \left[\frac{z}{\left(1+\frac{z^2}{k}\right)^2}\right] f(z)dz$$

$$= -\frac{p}{(1+p)}\left[\left(1+\frac{z^2}{k}\right)^{-p-1}\right]_{-\infty}^{\infty} = 0$$

since μ and σ are independent, and

$$\frac{\partial^2 Inl}{\partial\mu\partial\sigma} = -\frac{2p}{k\sigma^2}\sum_{t=1}^{n}\frac{2z_t}{\left(1+\frac{z_t^2}{k}\right)^2}.$$

$$\frac{1}{n}I_{22} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1^2}\right\} = \frac{2p}{k\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{1-\frac{1}{k}z_t^2}{\left[1+\frac{1}{k}z_t^2\right]^2}y_{t-1}^2\right\}$$

$$= \frac{2p}{k\sigma^2}\left(\frac{\sigma^2}{1-\phi_1^2}+\frac{\mu^2}{(1-\phi_1)^2}\right)E\left\{\frac{1-\frac{1}{k}z^2}{\left(1+\frac{1}{k}z^2\right)^2}\right\}$$

$$= \left(\frac{\sigma^2}{1-\phi_1^2}+\frac{\mu^2}{(1-\phi_1)^2}\right)\frac{2p}{k}\frac{\left(p-\frac{1}{2}\right)}{(p+1)}$$

$$= \left(\frac{\sigma^2}{1-\phi_1^2}+\frac{\mu^2}{(1-\phi_1)^2}\right)\frac{p\left(p-\frac{1}{2}\right)}{\left(p-\frac{3}{2}\right)(p+1)\sigma^2}.$$

Because,

$$E(y^2) = V(y) + [E(y)]^2$$

$$= \frac{\sigma^2}{1-\phi_1^2} + \frac{\mu^2}{(1-\phi_1)^2}$$

and

$$\frac{2p}{k\sigma^2} E\left\{ \frac{1-\frac{1}{k}z^2}{\left(1+\frac{1}{k}z^2\right)^2} \right\} = \frac{np\left(p-\frac{1}{2}\right)}{\left(p-\frac{3}{2}\right)(p+1)\sigma^2} .$$

$$\frac{1}{n} I_{23} = \frac{1}{n} I_{32} = -\frac{1}{n} E\left\{ \frac{\partial^2 InL}{\partial\phi_1 \partial\sigma} \right\} = \frac{2p}{k\sigma^2} E\left\{ \frac{1}{n}\sum_{t=1}^{n} \frac{2z_t}{\left(1+\frac{1}{k}z_t^2\right)^2} y_{t-1}^2 \right\}$$

$$= \left(\frac{\sigma^2}{1-\phi_1^2} + \frac{\mu^2}{(1-\phi_1)^2} \right) \frac{4p}{k\sigma^2} E\left\{ \frac{z^2}{\left(1+\frac{1}{k}z^2\right)^2} \right\} = 0$$

where $$\frac{\partial^2 Inl}{\partial\phi_1 \partial\sigma} = -\frac{2p}{k\sigma^2} \sum_{t=1}^{n} \frac{2z_i}{\left(1+\frac{z_t^2}{k}\right)^2} y_{t-1}^2 ,$$

$E\left\{ \frac{z^2}{\left(1+\frac{1}{k}z^2\right)^2} \right\} = 0$ as it is computed in I_{13} calculation.

$$\frac{1}{n}I_{33} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\sigma^2}\right\} = -\frac{1}{\sigma^2} + \frac{2p}{k\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{z_t^2}{\left(1+\frac{z_t^2}{k}\right)^2}\left(3+\frac{z_t^2}{k}\right)\right\}$$

$$= -\frac{1}{\sigma^2} + \frac{2p}{k\sigma^2}E\left\{\frac{z^2}{\left(1+\frac{z^2}{k}\right)^2}\left(3+\frac{z^2}{k}\right)\right\}$$

$$= -\frac{1}{\sigma^2} + \frac{2p}{k\sigma^2}\int_{-\infty}^{\infty}\left\{\left(1+\frac{z^2}{k}\right)^{-1} - 2\left(1+\frac{z^2}{k}\right)^{-2}\right\}f(z)dz$$

$$= -\frac{1}{\sigma^2} + \frac{2p}{k\sigma^2}\left[1+\frac{\left(p-\frac{1}{2}\right)}{p} - 2\frac{\left(p+\frac{1}{2}\right)\left(p-\frac{1}{2}\right)}{p(p+1)}\right]$$

$$= \frac{2\left(p-\frac{1}{2}\right)}{(p+1)}\sigma^2 \ .$$

Because,

$$\frac{\partial^2 InL}{\partial\sigma^2} = \frac{n}{\sigma^2} - \frac{2p}{k\sigma^2}\sum_{t=1}^{n}\frac{z_t^2}{\left(1+\frac{z_t^2}{k}\right)^2}\left(3+\frac{z_t^2}{k}\right).$$

APPENDIX E

For the AR (1) model with intercept term, where the error terms are iid, and have the gamma distribution, the elements of the Fisher information matrix are as follows:

$$\frac{1}{n} I_{11} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial \mu^2}\right\} = \frac{1}{n} I_{11} = \frac{(k-1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{z_t^2}\right\}$$

$$= \frac{(k-1)}{\sigma^2} E\left\{\frac{1}{z^2}\right\}$$

$$= \frac{(k-1)}{\sigma^2}\frac{1}{\Gamma(k)}\int_0^\infty e^{-z} z^{(k-2)-1} dz$$

$$= \frac{1}{(k-2)\sigma^2}$$

since

$$\frac{\partial^2 InL}{\partial \phi_1^2} = -\frac{(k-1)}{\sigma^2}\sum_{t=1}^{n}\frac{1}{z_t^2}.$$

$$\frac{1}{n} I_{12} = \frac{1}{n} I_{21} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\mu\partial\phi_1}\right\} = \frac{(k-1)n}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{z_t^2} y_{t-1}\right\}$$

$$= \frac{(k-1)}{\sigma^2} E\left\{\frac{y}{z^2}\right\}$$

$$= \frac{(k-1)}{\sigma^2}\frac{(\mu + k\sigma)}{\sigma(1-\phi_1)\Gamma(k)}\int_0^\infty e^{-z} z^{(k-2)-1} dz$$

$$= \frac{(k-1)}{\sigma^2}\frac{(\mu + k\sigma)}{\sigma(1-\phi_1)(k-1)(k-2)\Gamma(k-2)}\Gamma(k-2)\sigma$$

$$= \frac{(\mu + k\sigma)}{(k-2)(1-\phi_1)\sigma^2}$$

where $$\frac{\partial^2 InL}{\partial\mu\partial\phi_1} = -\frac{(k-1)}{\sigma^2}\sum_{t=1}^{n}\frac{1}{z_t^2} y_{t-1},$$

$$E(y) = \frac{\mu + k\sigma}{1-\phi_1}.$$

$$\frac{1}{n} I_{13} = \frac{1}{n} I_{31} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\mu\partial\sigma}\right\} = \frac{1}{n}\frac{n}{\sigma^2} = \frac{1}{\sigma^2}$$

since

$$\frac{\partial^2 InL}{\partial\mu\partial\sigma} = -\frac{n}{\sigma^2},$$

and

$$\frac{(k-1)}{\sigma}\sum_{t=1}^{n}\frac{1}{z_t}=\frac{(k-1)}{\sigma}\sum_{t=1}^{n}\frac{\sigma}{y_t-\phi_1 y_{t-1}-\mu}$$ is a constant term with respect to σ.

$$\frac{1}{n}I_{22}=-\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\phi_1^2}\right\}=\frac{(k-1)}{\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{z_t^2}y_{t-1}^2\right\}$$

$$=\frac{(k-1)}{\sigma^2}E\left\{\frac{y^2}{z^2}\right\}$$

$$=\frac{(k-1)}{\sigma^2}\left(\frac{k\sigma^2}{1-\phi_1^2}+\frac{(\mu+k\sigma)^2}{(1-\phi_1)^2}\right)\frac{1}{\Gamma(k)\sigma}\Gamma(k-2)\sigma$$

$$=\frac{(k-1)}{\sigma^2}\left(\frac{k\sigma^2}{1-\phi_1^2}+\frac{(\mu+k\sigma)^2}{(1-\phi_1)^2}\right)\frac{1}{(k-1)(k-2)\Gamma(k-2)\sigma}\Gamma(k-2)\sigma$$

$$=\frac{1}{(k-1)\sigma^2}\left(\frac{k\sigma^2}{1-\phi_1^2}+\frac{(\mu+k\sigma)^2}{(1-\phi_1)^2}\right)$$

since

$$\frac{\partial^2 InL}{\partial\phi_1^2}=-\frac{(k-1)}{\sigma^2}\sum_{t=1}^{n}\frac{1}{z_t^2}y_{t-1}^2\ ,$$

and

$$E(y^2) = V(y) + [E(y)]^2$$

$$= \frac{k\sigma^2}{1-\phi_1^2} + \frac{(\mu + k\sigma)^2}{(1-\phi_1)^2}.$$

$$\frac{1}{n} I_{23} = \frac{1}{n} I_{32} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\phi_1 \partial\sigma}\right\} = \frac{1}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} y_{t-1}\right\}$$

$$= \frac{1}{\sigma^2} E(y) = \frac{\mu + k\sigma}{(1-\phi_1)\sigma^2}$$

where $$\frac{\partial^2 InL}{\partial\phi_1 \partial\sigma} = -\frac{\sum_{t=1}^{n} y_{t-1}}{\sigma^2}.$$

$$\frac{1}{n} I_{33} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\sigma^2}\right\} = -\frac{1}{\sigma^2} - \frac{(k-1)}{\sigma^2} + \frac{2}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} z_t\right\}$$

$$= -\frac{1}{\sigma^2} - \frac{(k-1)}{\sigma^2} + \frac{2}{\sigma^2} E(z)$$

$$= -\frac{1}{\sigma^2} - \frac{(k-1)}{\sigma^2} + \frac{2}{\sigma^2}\frac{1}{\Gamma(k)\sigma}\int_0^\infty e^{-z} z^{(k+1)-1} dz$$

$$= -\frac{1}{\sigma^2} - \frac{(k-1)}{\sigma^2} + \frac{2}{\sigma^2}\frac{1}{\Gamma(k)\sigma}\Gamma(k+1)\sigma$$

$$= -\frac{1}{\sigma^2} - \frac{(k-1)}{\sigma^2} + \frac{2k}{\sigma^2} = \frac{k}{\sigma^2}$$

where $$\frac{\partial^2 InL}{\partial \sigma^2} = \frac{n}{\sigma^2} + \frac{n(k-1)}{\sigma^2} - \frac{2}{\sigma^2}\sum_{t=1}^{n} z_t \ .$$

APPENDIX F

For the AR (1) model with intercept term where the error terms are iid and come from the generalized logistic distribution, the elements of the Fisher information matrix are written as the following form:

$$\frac{1}{n} I_{11} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial \mu^2}\right\} = \frac{b+1}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{e^{-z_t}}{(1+e^{-z_t})^2}\right\}$$

$$= \frac{b+1}{\sigma^2} E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\}$$

$$= \frac{b(b+1)}{\sigma^2} \int_0^1 \frac{t}{t-1}(1-t)^{b+1} dt \text{ where } t = \frac{e^{-z}}{1+e^{-z}}$$

$$= \frac{b(b+1)}{\sigma^2} \frac{\Gamma(b+1)\Gamma(2)}{\Gamma(b+3)}$$

$$= \frac{b}{(b+2)\sigma^2}$$

where $$\frac{\partial^2 InL}{\partial \mu^2} = -\frac{b+1}{\sigma^2}\left\{\sum_{t=1}^{n} \frac{e^{-z_t}}{(1+e^{-z_t})^2}\right\}.$$

$$\frac{1}{n} I_{12} = \frac{1}{n} I_{21} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\mu\partial\phi_1}\right\} = \frac{b+1}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{e^{-z_t}}{(1+e^{-z_t})^2} y_{t-1}\right\}$$

$$= \frac{b+1}{\sigma^2} E(y) E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\}$$

$$= \frac{\mu + \sigma\left[\Psi(b) - \Psi(1)\right]}{1-\phi_1} \frac{b+1}{\sigma^2} E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\}$$

$$= \frac{\mu + \sigma\left[\Psi(b) - \Psi(1)\right]}{1-\phi_1} \frac{b}{(b+2)\sigma^2}$$

since

$$\frac{\partial^2 InL}{\partial\mu\partial\phi_1} = -\frac{b+1}{\sigma^2}\left\{\sum_{t=1}^{n} \frac{e^{-z_t}}{(1+e^{-z_t})^2} y_{t-1}\right\}$$

and

$$\frac{b+1}{\sigma^2} E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\} = \frac{b}{(b+2)\sigma^2}$$ as it is calculated in I_{11}.

$$\frac{1}{n} I_{13} = \frac{1}{n} I_{31} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial\mu\partial\sigma}\right\}$$

$$= \frac{1}{\sigma^2} - \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{e^{-z_i}}{1+e^{-z_i}}\right\} + \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} z_t \frac{e^{-z_t}}{\left(1+e^{-z_t}\right)^2}\right\}$$

$$= \frac{1}{\sigma^2} - \frac{(b+1)}{\sigma^2} E\left\{ \frac{e^{-z}}{1+e_{-z}} \right\} + \frac{(b+1)}{\sigma^2} E\left\{ z \frac{e^{-z}}{(1+e^{-z})^2} \right\}$$

$$= \frac{1}{\sigma^2} - \frac{(b+1)}{\sigma^2} \frac{1}{(b+1)} + \frac{(b+1)}{\sigma^2} \frac{b}{(b+1)(b+2)} \left[\Psi(b+1) - \Psi(2) \right]$$

$$= \frac{1}{\sigma^2} \frac{b}{(b+2)} \left[\Psi(b+1) - \Psi(2) \right]$$

since

$$\frac{\partial^2 InL}{\partial \mu \partial \sigma} = -\frac{n}{\sigma^2} + \frac{(b+1)}{\sigma^2} \left\{ \sum_{t=1}^{n} \frac{e^{-z_t}}{1+e^{-z_t}} \right\} - \frac{(b+1)}{\sigma^2} \left\{ \sum_{t=1}^{n} z_t \frac{e^{-z_t}}{\left(1+e^{-z_t}\right)^2} \right\},$$

$$E\left\{ \frac{e^{-z}}{1+e_{-z}} \right\} = \frac{1}{(b+1)},$$

$$E\left\{ z \frac{e^{-z}}{(1+e^{-z})^2} \right\} = \frac{b}{(b+1)(b+2)} \left[\Psi(b+1) - \Psi(2) \right],$$

and

$$\Psi(b) = \frac{\Gamma'(b)}{\Gamma(b)}.$$

$$\frac{1}{n} I_{22} = -\frac{1}{n} E\left\{ \frac{\partial^2 InL}{\partial \phi_1^2} \right\} = \frac{b+1}{\sigma^2} E\left\{ \frac{1}{n} \sum_{t=1}^{n} \frac{e^{-z_i}}{(1+e^{-z_i})^2} y_{t-1}^2 \right\}$$

$$= \frac{b+1}{\sigma^2} E\left(y^2\right) E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\}$$

$$= \left\{\frac{[\Psi'(b)+\Psi'(1)]\sigma^2}{1-\phi_1^2} + \frac{(\mu+[\Psi(b)-\Psi(1)]\sigma)^2}{(1-\phi_1)^2}\right\} \frac{b+1}{\sigma^2} E\left\{\frac{e^{-z}}{(1+e^{-z})^2}\right\}$$

$$= \frac{b}{(b+2)\sigma^2} \left\{\frac{[\Psi'(b)+\Psi'(1)]\sigma^2}{1-\phi_1^2} + \frac{(\mu+[\Psi(b)-\Psi(1)]\sigma)^2}{(1-\phi_1)^2}\right\}$$

where $$\frac{\partial^2 InL}{\partial \phi_1{}^2} = -\frac{b+1}{\sigma^2}\left\{\sum_{t=1}^{n} \frac{e^{-z_t}}{(1+e^{-z_t})^2} y_{t-1}^2\right\},$$

$$E(y^2) = V(y) + [E(y)]^2$$

$$= \frac{[\Psi'(b)+\Psi'(1)]\sigma^2}{1-\phi_1^2} + \left(\frac{\mu+[\Psi(b)-\Psi(1)\sigma]}{1-\phi_1}\right),$$

and

$$E\left\{\frac{e^{-z}}{1+e_{-z}}\right\} = \frac{1}{(b+1)}$$ as we stated before..

$$\frac{1}{n} I_{23} = \frac{1}{n} I_{32} = -\frac{1}{n} E\left\{\frac{\partial^2 InL}{\partial \phi_1 \partial \sigma}\right\}$$

$$= \frac{1}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} y_{t-1}\right\} + \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=1}^{n} \frac{e^{-z_t} z_t}{(1+e^{-z_t})^2} y_{t-1}\right\} - \frac{(b+1)}{\sigma^2} E\left\{\frac{1}{n}\sum_{t=2}^{n} \frac{e^{-z_t}}{1+e^{-z_t}} y_{t-1}\right\}$$

$$= E(y)\left\{\frac{1}{\sigma^2} + \frac{b(b+1)}{(b+1)(b+2)\sigma^2}\left[\Psi(b+1) - \Psi(2)\right] - \frac{b+1}{\sigma^2}\frac{1}{b+1}\right\}$$

$$= \frac{\mu + \left[\Psi(b) - \Psi(1)\right]\sigma}{1-\phi_1}\frac{b}{(b+2)\sigma^2}\left[\Psi(b+1) - \Psi(2)\right]$$

$$= \frac{\mu + \left[\Psi(b) - \Psi(1)\right]\sigma}{1-\phi_1}\frac{b}{(b+2)\sigma^2}\left[\Psi(b) - \Psi(1)\right]$$

since

$$\frac{\partial^2 InL}{\partial\phi_1\partial\sigma} = \frac{1}{\sigma^2}\left\{\frac{1}{n}\sum_{t=1}^{n} y_{t-1}\right\} + \frac{(b+1)}{\sigma^2}\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{e^{-z_t}z_t}{(1+e^{-z_t})^2}y_{t-1}\right\} - \frac{(b+1)}{\sigma^2}\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{e^{-z_{ti}}}{1+e^{-z_{ti}}}y_{t-1}\right\},$$

$$\left[\Psi(b+1) - \Psi(2)\right] = \frac{\Gamma'(b+1)}{\Gamma(b+1)} - \frac{\Gamma'(2)}{\Gamma(2)} = \left[\Psi(b) - \Psi(1)\right].$$

$$\frac{1}{n}I_{33} = -\frac{1}{n}E\left\{\frac{\partial^2 InL}{\partial\sigma^2}\right\}$$

$$= \frac{1}{\sigma^2} - \frac{2}{\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n} z_t\right\} + \frac{(b+1)}{\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{z_t^2 e^{-z_t}}{\left(1+e^{-z_t}\right)^2}\right\} - \frac{2(b+1)}{\sigma^2}E\left\{\frac{1}{n}\sum_{t=1}^{n}\frac{z_t e^{-z_t}}{(1+e^{-z_t})}\right\}$$

$$= \frac{1}{\sigma^2} - \frac{2}{\sigma^2}E\{z\} + \frac{(b+1)}{\sigma^2}E\left\{\frac{z^2 e^{-z}}{(1+e^{-z})^2}\right\} - \frac{2(b+1)}{\sigma^2}E\left\{\frac{ze^{-z}}{(1+e^{-z})}\right\}$$

$$= \frac{1}{\sigma^2} - \frac{2}{\sigma^2}\left[\Psi(b) - \Psi(1)\right] + \frac{(b+1)}{\sigma^2}\frac{b}{(b+1)(b+2)}\left\{\Psi'(b+1) + \Psi'(2) + \left[\Psi(b+1) - \Psi(2)\right]^2\right\}$$

$$- \frac{2(b+1)}{\sigma^2}\frac{1}{(b+1)}\left[\Psi(b) - \Psi(2)\right]$$

$$= \frac{1}{\sigma^2} + \frac{b}{\sigma^2(b+2)}\left\{\Psi'(b+1) + \Psi'(2) + \left[\Psi(b+1) - \Psi(2)\right]^2\right\}$$

where

$$\frac{\partial^2 InL}{\partial\sigma^2} = \frac{1}{\sigma^2} - \frac{2}{\sigma^2}\left\{\sum_{t=1}^{n} z_t\right\} + \frac{(b+1)}{\sigma^2}\left\{\sum_{t=1}^{n}\frac{z_t^2 e^{-z_t}}{\left(1+e^{-z_t}\right)^2}\right\} - \frac{2(b+1)}{\sigma^2}\left\{\sum_{t=1}^{n}\frac{z_t e^{-z_t}}{(1+e^{-z_t})}\right\},$$

$$E\left\{\frac{z^2 e^{-z}}{(1+e^{-z})^2}\right\} = \frac{b}{(b+1)(b+2)}\left\{\Psi'(b+1) + \Psi'(2) + \left[\Psi(b+1) - \Psi(2)\right]^2\right\}.$$

Moreover,

$E\left\{\frac{ze^{-z}}{1+e^{-z}}\right\} = \frac{1}{(b+1)}\left[\Psi(b) - \Psi(2)\right]$ as it is defined in previous cases.

Table 1. I. Critical Values for the Phillips – Perron Z_{ϕ_1} Test and for the Dickey – Fuller Test Based on Estimated LS Autoregressive Coefficient:

Sample Size n	**Probability that $n(\hat{\phi}_1 - 1)$ is less than entry**							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
	Case 1							
25	−11.9	−9.3	−7.3	−5.3	1.01	1.40	1.79	2.28
50	−12.9	−9.9	−7.7	−5.5	0.97	1.35	1.70	2.16
100	−13.3	−10.2	−7.9	−5.6	0.95	1.31	1.65	2.09
250	−13.6	−10.3	−8.0	−5.7	0.93	1.28	1.62	2.04
500	−13.7	−10.4	−8.0	−5.7	0.93	1.28	1.61	2.04
∞	−13.8	−10.5	−8.1	−5.7	0.93	1.28	1.60	2.03
	Case 2							
25	−17.2	−14.6	−12.5	−10.2	−0.76	0.01	0.65	1.40
50	−18.9	−15.7	−13.3	−10.7	−0.81	−0.07	0.53	1.22
100	−19.8	−16.3	−13.7	−11.0	−0.83	−0.10	0.47	1.14
250	−20.3	−16.6	−14.0	−11.2	−0.84	−0.12	0.43	1.09
500	−20.5	−16.8	−14.0	−11.2	−0.84	−0.13	0.42	1.06
∞	−20.7	−16.9	−14.1	−11.3	−0.85	−0.13	0.41	1.04

The probability shown at the head of the column is the area in the left – hand tail (Source: Fuller, W. A. (1976), Introduction to Statistical Time Series, Wiley, New York, p. 371).

Table 1. II. Critical Values for the Phillips – Perron Z_t Test and for the Dickey – Fuller Test Based on Estimated LS t Statistic

Sample Size n	**Probability that $(\hat{\phi}_1 - 1)/\hat{\sigma}_{\hat{\phi}_1}$ is less than entry**							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
	Case 1							
25	−2.66	−2.26	−1.95	−1.60	0.92	1.33	1.70	2.16
50	−2.62	−2.25	−1.95	−1.61	0.91	1.31	1.66	2.08
100	−2.60	−2.24	−1.95	−1.61	0.90	1.29	1.64	2.03
250	−2.58	−2.23	−1.95	−1.62	0.89	1.29	1.63	2.01
500	−2.58	−2.23	−1.95	−1.62	0.89	1.28	1.62	2.00
∞	−2.58	−2.23	−1.95	−1.62	0.89	1.28	1.62	2.00
	Case 2							
25	−3.75	−3.33	−3.00	−2.63	−0.37	0.00	0.34	0.72
50	−3.58	−3.22	−2.93	−2.60	−0.40	−0.03	0.29	0.66
100	−3.51	−3.17	−2.89	−2.58	−0.42	−0.05	0.26	0.63
250	−3.46	−3.14	−2.88	−2.57	−0.42	−0.06	0.24	0.62
500	−3.44	−3.13	−2.87	−2.57	−0.43	−0.07	0.24	0.61
∞	−3.43	−3.12	−2.86	−2.57	−0.44	−0.07	0.23	0.60

The probability shown at the head of the column is the area in the left – hand tail (Source: Fuller, W. A. (1976), Introduction to Statistical Time Series, Wiley, New York, p. 373).

Table 3. I - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student – t) – AR (1) Model Without Intercept Term

ϕ_1=1.0	n=20			n=30			n=50			n=100			n=300		
	MML	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
p=2.0															
Mean	0.947	0.929		0.966	0.950		0.981	0.969		0.991	0.984		0.997	0.994	
Stnd dev.	0.121	0.137	78	0.080	0.095	71	0.046	0.058	63	0.022	0.030	54	0.007	0.011	40
Skewness	3.990	2.730		4.379	3.865		5.119	4.135		5.847	4.649		5.255	4.699	
Kurtosis	9.479	8.187		9.925	10.352		11.238	9.890		12.773	10.465		11.164	10.118	
p=2.5															
Mean	0.938	0.926		0.959	0.948		0.976	0.967		0.988	0.983		0.996	0.994	
Stnd dev.	0.127	0.137	86	0.088	0.096	84	0.052	0.060	75	0.026	0.031	70	0.008	0.010	64
Skewness	2.989	2.581		4.234	3.207		4.826	4.186		4.992	4.403		6.007	5.794	
Kurtosis	7.357	7.158		10.037	7.980		10.681	9.622		11.214	10.212		13.608	13.098	
p=3.5															
Mean	0.927	0.924		0.949	0.946		0.970	0.967		0.986	0.983		0.995	0.994	
Stnd dev.	0.135	0.139	94	0.095	0.098	94	0.055	0.058	90	0.027	0.030	81	0.009	0.010	81
Skewness	2.913	2.638		4.264	3.717		3.804	3.435		4.629	4.339		5.020	4.655	
Kurtosis	7.332	6.741		10.046	9.030		8.710	8.055		10.331	9.566		10.932	9.846	
p=5.0															
Mean	0.923	0.922		0.948	0.946		0.967	0.966		0.984	0.983		0.995	0.994	
Stnd dev.	0.141	0.143	97	0.096	0.098	96	0.059	0.060	97	0.030	0.031	94	0.010	0.010	100
Skewness	3.074	2.969		3.851	3.815		4.131	4.057		4.069	4.130		4.200	4.270	
Kurtosis	7.777	7.558		8.720	8.727		9.506	9.172		8.900	9.026		9.154	9.359	
p=10.0															
Mean	0.922	0.922		0.946	0.946		0.967	0.967		0.983	0.982		0.994	0.994	
Stnd dev.	0.139	0.139	100	0.098	0.098	99	0.059	0.059	100	0.031	0.031	100	0.011	0.011	100
Skewness	3.025	2.994		3.389	3.380		3.629	3.603		4.276	4.317		6.110	6.159	
Kurtosis	7.401	7.358		7.929	7.935		8.701	8.622		9.692	9.858		13.602	13.741	

Table 3. I - B: Type I Error (ε_t ~ Student – t) – AR (1) Model Without Intercept Term

	p=2.5		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
n=20	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	5.925*	6.045	5.925	5.994	6.012	6.050*	6.046	6.061	6.024*	6.024*
Type I Error	0.052	0.053	0.053	0.055	0.049	0.051	0.051	0.051	0.054	0.053
n=30										
Critical Value	6.635	6.787	6.706	6.798*	6.779	6.847	6.809*	6.831*	6.827*	6.830*
Type I Error	0.051	0.051	0.049	0.051	0.051	0.051	0.052	0.053	0.055	0.054
n=50										
Critical Value	7.978	8.141	8.024	8.144	8.094	8.181	8.119*	8.150*	8.123*	8.129*
Type I Error	0.051	0.050	0.050	0.050	0.049	0.049	0.051	0.052	0.052	0.052
n=100										
Critical Value	10.612	10.770	10.666	10.775	10.711	10.786	10.744	10.776	10.784	10.790
Type I Error	0.049	0.048	0.050	0.052	0.049	0.049	0.052	0.052	0.051	0.049
n=300										
Critical Value	17.670	17.784	17.683	17.757	17.722	17.775	17.753	17.780	17.777	17.781
Type I Error	0.048	0.049	0.051	0.050	0.051	0.047	0.051	0.050	0.049	0.050

Note: For (*) values, Three- Moment Chi- Square Approximation, is valid; other cases satisfy Four- Moment F Approximation

Table 3. I - C: Values of the Power (ε_t ~ Student – t) – AR (1) Model Without Intercept Term

	ϕ_1 =1.0		ϕ_1 =0.99		ϕ_1 =0.95		ϕ_1 =0.90		ϕ_1 =0.85		ϕ_1 =0.80		ϕ_1 =0.70	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=20														
p=2.5	0.052	0.053	0.053	0.056	0.079	0.078	0.125	0.121	0.202	0.194	0.274	0.259	0.494	0.454
p=3.0	0.053	0.055	0.058	0.058	0.084	0.085	0.125	0.124	0.179	0.174	0.278	0.265	0.476	0.445
p=3.5	0.049	0.051	0.056	0.056	0.082	0.084	0.125	0.124	0.190	0.188	0.268	0.263	0.468	0.453
p=5.0	0.051	0.051	0.055	0.055	0.082	0.083	0.128	0.127	0.190	0.191	0.244	0.244	0.444	0.435
p=10.0	0.054	0.053	0.056	0.056	0.081	0.081	0.121	0.120	0.181	0.180	0.259	0.257	0.454	0.452
n=30														
p=2.5	0.051	0.051	0.056	0.057	0.095	0.095	0.191	0.186	0.332	0.296	0.512	0.448	0.820	0.724
p=3.0	0.049	0.051	0.056	0.057	0.099	0.097	0.177	0.171	0.306	0.285	0.483	0.437	0.795	0.745
p=3.5	0.051	0.051	0.058	0.056	0.098	0.097	0.182	0.179	0.310	0.297	0.472	0.443	0.772	0.730
p=5.0	0.052	0.053	0.061	0.060	0.104	0.101	0.187	0.184	0.293	0.291	0.454	0.443	0.757	0.732
p=10.0	0.055	0.054	0.058	0.058	0.102	0.102	0.184	0.184	0.291	0.290	0.447	0.444	0.741	0.736
n=50														
p=2.5	0.051	0.050	0.064	0.062	0.168	0.149	0.412	0.342	0.690	0.571	0.885	0.794	0.995	0.983
p=3.0	0.050	0.050	0.062	0.062	0.158	0.150	0.376	0.334	0.635	0.564	0.856	0.787	0.991	0.981
p=3.5	0.049	0.049	0.065	0.064	0.143	0.138	0.350	0.322	0.628	0.571	0.840	0.792	0.988	0.977
p=5.0	0.051	0.052	0.063	0.063	0.148	0.142	0.351	0.338	0.584	0.561	0.815	0.790	0.982	0.977
p=10.0	0.052	0.052	0.063	0.063	0.153	0.152	0.319	0.317	0.556	0.552	0.802	0.799	0.980	0.979

Table 3. I - C: Values of the Power (ε_t ~ Student – t) – AR (1) Model Without Intercept Term (Continue)

	ϕ_1 =1.0		ϕ_1 =0.99		ϕ_1 =0.95		ϕ_1 =0.90		ϕ_1 =0.85		ϕ_1 =0.80		ϕ_1 =0.70	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=100														
p=2.5	0.049	0.048	0.084	0.080	0.379	0.307	0.881	0.761	0.994	0.974	1.000	0.999	1.000	1.000
p=3.0	0.050	0.052	0.086	0.082	0.353	0.303	0.862	0.777	0.993	0.979	1.000	0.999	1.000	1.000
p=3.5	0.049	0.049	0.082	0.080	0.348	0.313	0.815	0.754	0.988	0.975	1.000	0.999	1.000	1.000
p=5.0	0.052	0.052	0.078	0.076	0.314	0.300	0.801	0.768	0.981	0.972	0.999	0.999	1.000	1.000
p=10.0	0.051	0.049	0.081	0.082	0.323	0.317	0.771	0.763	0.977	0.976	0.999	0.998	1.000	1.000
n=300														
p=2.5	0.048	0.049	0.221	0.174	0.997	0.974	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=3.0	0.051	0.050	0.172	0.154	1.000	0.968	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=5.0	0.051	0.050	0.180	0.172	0.983	0.971	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=10.0	0.049	0.050	0.159	0.158	0.978	0.975	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. I - D: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student–t and $y_0 \sim N(0,1/(1-\phi_1^2))$) AR(1) Model Without Intercept Term

	p=2.5			**p=3.0**			**p=3.5**			**p=5.0**			**p=10.0**		
n=50	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
ϕ_1=0.5			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	0.976	0.968		0.974	0.968		0.972	0.968		0.969	0.968		0.967	0.967	97
Stnddev.	0.051	0.059	75	0.053	0.058	84	0.055	0.059	87	0.056	0.057	97	0.058	0.059	
Skew.	5.391	5.338		4.831	4.476		5.425	4.821		4.203	3.925		5.014	4.911	
Kurtosis	11.778	11.944		11.246	10.567		12.803	11.112		9.335	8.763		11.670	11.391	
ϕ_1=0.9															
Mean	0.977	0.970		0.975	0.969		0.972	0.969		0.970	0.969		0.969	0.968	
Stnddev.	0.049	0.056	77	0.051	0.056	83	0.052	0.056	86	0.054	0.056	93	0.056	0.057	97
Skew	4.924	4.510		4.761	4.191		4.191	4.189		4.258	4.283		4.189	4.218	
Kurtosis	10.999	10.401		10.810	9.836		9.172	9.193		9.728	9.908		9.461	9.577	

Table 3. I - E: Type I Error ($\varepsilon_t \sim$ Student–t and $y_0 \sim N(0,1/(1-\phi_1^2))$) - AR (1) Model Without Intercept Term

n=50	**p=2.5**		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
ϕ_1=0.5	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	7.939	8.103*	7.991	8.096	8.082	8.165	8.083*	8.111*	8.118	8.125
Type I Error	0.0518	0.054	0.050	0.051	0.048	0.050	0.053	0.051	0.048	0.049
ϕ_1=0.9										
Critical Value	7.907	8.062*	7.988	8.101	7.998	8.070	8.038*	8.067*	8.066	8.072
Type I Error	0.048	0.050	0.050	0.049	0.049	0.049	0.050	0.049	0.049	0.049

Note: For (*) values, Three- Moment Chi- Square Approximation, is valid; other cases satisfy Four - Moment F Approximation

Table 3. I - F: Values of the Power ($\varepsilon_t \sim$ Student–t and $y_0 \sim N(0,1/(1-\phi_1^2))$) - AR (1) Model Without Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
n=50	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
ϕ_1=0.5														
p=2.5	0.052	0.054	0.064	0.063	0.162	0.149	0.401	0.330	0.718	0.602	0.898	0.805	0.996	0.986
p=3.0	0.050	0.051	0.063	0.066	0.157	0.141	0.344	0.344	0.660	0.590	0.859	0.791	0.993	0.985
p=3.5	0.048	0.050	0.066	0.063	0.159	0.147	0.345	0.313	0.637	0.584	0.854	0.801	0.992	0.985
p=5.0	0.053	0.051	0.063	0.065	0.162	0.155	0.352	0.335	0.606	0.581	0.826	0.801	0.986	0.983
p=10.0	0.048	0.049	0.065	0.064	0.150	0.149	0.343	0.341	0.584	0.578	0.803	0.799	0.982	0.980
ϕ_1=0.9														
p=2.5	0.048	0.050	0.067	0.066	0.177	0.153	0.449	0.372	0.763	0.647	0.928	0.832	0.997	0.990
p=3.0	0.050	0.049	0.064	0.065	0.158	0.149	0.412	0.371	0.691	0.610	0.901	0.843	0.997	0.991
p=3.5	0.049	0.049	0.061	0.061	0.158	0.151	0.405	0.371	0.648	0.595	0.887	0.842	0.994	0.988
p=5.0	0.050	0.049	0.065	0.066	0.161	0.155	0.364	0.349	0.662	0.635	0.879	0.857	0.994	0.991
p=10.0	0.049	0.049	0.062	0.062	0.156	0.153	0.368	0.367	0.632	0.628	0.838	0.832	0.985	0.985

Table 3. II - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Gamma) – AR (1) Model Without Intercept Term

ϕ_1=1.0	n=20			n=30			n=50			n=100			n=300		
	MML	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
k=3.0			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	1.006	0.662		1.003	0.781		1.002	0.873		1.001	0.938		1.000	0.980	
Stnd dev.	0.010	0.052	4	0.005	0.030	3	0.002	0.014	2	0.001	0.005	4	0.000	0.001	100
Skewness	0.344	0.384		0.224	0.258		0.180	0.136		0.129	0.071		0.057	0.025	
Kurtosis	3.487	3.845		3.265	3.432		3.156	3.213		3.169	3.164		3.094	3.048	
k=4.0															
Mean	1.003	0.662		1.002	0.782		1.001	0.873		1.000	0.938		1.000	0.980	
Stnd dev.	0.010	0.046	5	0.006	0.026	5	0.003	0.012	6	0.001	0.004	6	0.000	0.001	100
Skewness	0.177	0.321		0.122	0.133		0.120	0.078		0.052	0.046		0.042	0.028	
Kurtosis	3.250	3.757		3.132	3.213		3.303	3.167		3.096	3.097		2.901	3.082	
k=5.0															
Mean	1.002	0.663		1.001	0.782		1.001	0.873		1.000	0.938		1.000	0.980	
Stnd dev.	0.010	0.040	6	0.005	0.023	5	0.003	0.011	7	0.001	0.004	6	0.000	0.001	100
Skewness	0.010	0.198		0.100	0.123		0.046	0.077		0.026	0.045		0.031	0.019	
Kurtosis	3.184	3.393		3.208	3.165		3.156	3.194		3.044	3.054		3.016	2.996	
k=8.0															
Mean	1.001	0.664		1.000	0.782		1.000	0.873		1.000	0.938		1.000	0.980	
Stnd dev.	0.010	0.032	10	0.005	0.018	8	0.002	0.009	5	0.001	0.003	11	0.000	0.001	0
Skewness	0.056	0.152		0.040	0.079		0.012	0.046		0.004	0.044		0.000	0.024	
Kurtosis	3.085	3.188		3.088	3.148		3.036	2.988		2.983	3.180		2.995	2.965	

Table 3. II - B: Type I Error ($\varepsilon_t \sim$ Gamma) – AR (1) Model Without Intercept Term

	k=3.0		**k=4.0**		**k=5.0**		**k=8.0**	
n=20	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	4.472	6.414	4.530	6.337	4.553	6.304	4.565	6.225
Type I Error	0.046	0.047	0.046	0.051	0.049	0.049	0.051	0.051
n=30								
Critical Value	5.443	6.968	5.493	6.920	5.514	6.887	5.533	6.840
Type I Error	0.050	0.050	0.047	0.050	0.045	0.048	0.052	0.051
n=50								
Critical Value	7.014	8.143	7.057	8.118	7.077	8.102	7.097	8.072
Type I Error	0.051	0.050	0.048	0.051	0.051	0.050	0.049	0.050
n=100								
Critical Value	9.942	10.706	9.973	10.693	9.988	10.685	10.005	10.670
Type I Error	0.048	0.050	0.046	0.048	0.048	0.049	0.052	0.051
n=300								
Critical Value	17.277	17.699	17.296	17.695	17.305	17.693	17.317	17.688
Type I Error	0.050	0.050	0.052	0.050	0.051	0.049	0.050	0.046

Note: For all values, Three- Moment Chi- Square Approximation is valid.

Table 3. II - C: Values of the Power ($\varepsilon_t \sim$ Gamma) – AR (1) Model Without Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=20														
k=3.0	0.262	0.047	0.961	0.152	0.995	0.416	1.000	0.738	1.000	0.881	1.000	0.945	1.000	0.978
k=4.0	0.022	0.051	0.177	0.101	0.960	0.479	1.000	0.836	1.000	0.943	1.000	0.977	1.000	0.993
k=5.0	0.006	0.049	0.092	0.112	0.936	0.552	1.000	0.892	1.000	0.972	1.000	0.991	1.000	0.998
k=8.0	0.006	0.051	0.081	0.135	0.927	0.711	1.000	0.969	1.000	0.998	1.000	0.999	1.000	1.000
n=30														
k=3.0	0.654	0.050	0.958	0.155	1.000	0.770	1.000	0.969	1.000	0.993	1.000	0.998	1.000	0.999
k=4.0	0.104	0.050	0.670	0.173	1.000	0.848	1.000	0.990	1.000	0.999	1.000	1.000	1.000	1.000
k=5.0	0.028	0.048	0.448	0.196	1.000	0.905	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.010	0.051	0.285	0.262	1.000	0.979	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=50														
k=3.0	0.982	0.050	1.000	0.362	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	0.525	0.051	0.999	0.425	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	0.141	0.050	0.987	0.495	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.026	0.050	0.937	0.641	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=100														
k=3.0	1.000	0.050	1.000	0.933	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	0.981	0.048	1.000	0.974	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	0.734	0.049	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.117	0.051	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=300														
k=3.0	1.000	0.050	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	1.000	0.050	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	1.000	0.049	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.793	0.046	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. III - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Generalized Logistic) - AR (1) Model Without Intercept Term

ϕ_1=1.0	n=20			n=30			n=50			n=100			n=300		
	MML	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
b=0.5															
Mean	0.983	0.641	12	0.993	0.769	8	0.998	0.867	5	1.000	0.937	3	1.000	0.980	11
Stnd dev.	0.062	0.179		0.029	0.106		0.011	0.050		0.003	0.017		0.001	0.003	
Skewness	7.337	31.286		7.137	89.217		13.88	407.722		0.709	3456.39		0.117	102562	
Kurtosis	17.247	31.286		22.76	89.218		50.57	407.726		5.527	3456.39		3.307	102562	
b=3.0															
Mean	0.998	0.659	6	0.999	0.780	5	1.000	0.872	4	1.000	0.938	6	1.000	0.980	0
Stnd dev.	0.021	0.089		0.011	0.050		0.005	0.024		0.002	0.008		0.000	0.002	
Skewness	0.386	125.545		0.245	398.026		0.053	1807.37		0.027	14158.5		0.007	384192	
Kurtosis	5.018	125.533		4.859	398.026		3.372	1807.36		3.252	14158.7		3.018	384191.	
b=4.0															
Mean	1.000	0.659	6	1.000	0.780	5	1.000	0.872	4	1.000	0.938	2	1.000	0.980	0
Stnd dev.	0.017	0.070		0.009	0.040		0.004	0.019		0.001	0.007		0.000	0.001	
Skewness	0.034	201.663		0.040	640.810		0.023	2926.50		0.003	23165		0.000	590762	
Kurtosis	3.491	201.661		3.683	640.810		3.456	2926.49		3.058	23164.9		2.918	590762	
b=6.0															
Mean	1.001	0.662	6	1.000	0.781	5	1.000	0.873	4	1.000	0.938	4	1.000	0.980	0
Stnd dev.	0.013	0.055		0.007	0.030		0.003	0.015		0.001	0.005		0.000	0.001	
Skewness	0.002	334.635		0.003	1084.81		0.001	4690.31		0.002	37583.7		0.000	106009	
Kurtosis	3.307	334.635		3.239	1084.81		3.135	4690.32		3.034	37583.7		2.987	106009	

Table 3. III - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Generalized Logistic) - AR (1) Model Without Intercept Term (Continue)

ϕ_1=1.0	n=20			n=30			n=50			n=100			n=300		
	MML	LS	% Eff	MML	LS	% Eff	MML	LS	% Eff	MML	LS	% Eff	MML	LS	% Eff
b=6.0															
Mean	1.001	0.662	6	1.000	0.781	5	1.000	0.873	4	1.000	0.938	4	1.000	0.980	0
Stnd dev.	0.013	0.055		0.007	0.030		0.003	0.015		0.001	0.005		0.000	0.001	
Skewness	0.002	334.635		0.003	1084.81		0.001	4690.31		0.002	37583.7		0.000	106009	
Kurtosis	3.307	334.635		3.239	1084.81		3.135	4690.32		3.034	37583.7		2.987	106009	
b=8.0															
Mean	1.001	0.663	7	1.000	0.781	5	1.000	0.873	5	1.000	0.938	4	1.000	0.980	0
Stnd dev.	0.012	0.047		0.006	0.026		0.003	0.013		0.001	0.005		0.000	0.001	
Skewness	0.017	461.951		0.015	1430.65		0.003	6301.65		0.003	49139.3		0.000	133148	
Kurtosis	3.118	461.951		3.137	1430.65		2.996	6301.65		3.070	49139.3		2.999	133148	

Table 3. III - B: Type I Error ($\varepsilon_t \sim$ Generalized Logistic) - AR (1) Model Without Intercept Term

	b=0.5		**b=3.0**		**b=4.0**		**b=6.0**		**b=8.0**	
n=20	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	5.071*	7.602*	4.653	6.759	4.600	6.558	4.570	6.417	4.559	4.920
Type I Error	0.037	0.047	0.044	0.044	0.045	0.048	0.043	0.049	0.041	0.049
n=30										
Critical Value	5.826*	7.877*	5.588	7.111	5.556	7.062	5.539	6.971	5.533	6.928
Type I Error	0.037	0.039	0.045	0.049	0.049	0.050	0.047	0.048	0.043	0.047
n=50										
Critical Value	7.240	8.659	7.133	8.272	7.118	8.204	7.107	8.148	7.104	8.127
Type I Error	0.036	0.044	0.047	0.047	0.052	0.049	0.050	0.050	0.043	0.046
n=100										
Critical Value	10.065	10.939	10.029	10.767	10.022	10.733	10.018	10.708	10.016	10.696
Type I Error	0.047	0.043	0.049	0.048	0.048	0.049	0.050	0.049	0.048	0.053
n=300										
Critical Value	17.339	17.770	17.330	17.717	17.328	17.709	17.326	17.699	17.326	17.696
Type I Error	0.051	0.054	0.050	0.050	0.051	0.048	0.052	0.055	0.046	0.051

Note: For (*) values, Four- Moment F Approximation is valid; other cases satisfy Three- Moment Chi - Square Approximation

Table 3. III - C: Values of the Power (ε_t ~ Generalized Logistic) - AR (1) Model Without Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=20														
b=0.5	0.040	0.047	0.050	0.057	0.126	0.119	0.303	0.169	0.673	0.184	0.887	0.203	0.984	0.241
b=3.0	0.044	0.044	0.091	0.075	0.679	0.248	0.975	0.462	0.999	0.577	1.000	0.635	1.000	0.744
b=4.0	0.045	0.048	0.014	0.086	0.863	0.307	0.996	0.571	1.000	0.731	1.000	0.812	1.000	0.875
b=6.0	0.048	0.049	0.020	0.095	0.934	0.409	1.000	0.736	1.000	0.870	1.000	0.927	1.000	0.972
b=8.0	0.041	0.049	0.022	0.111	0.964	0.477	1.000	0.821	1.000	0.933	1.000	0.972	1.000	0.992
n=30														
b=0.5	0.037	0.039	0.064	0.069	0.405	0.205	0.922	0.294	0.991	0.332	0.999	0.395	1.000	0.526
b=3.0	0.047	0.047	0.598	0.220	1.000	0.829	1.000	0.959	1.000	0.985	1.000	0.990	1.000	0.999
b=4.0	0.049	0.050	0.296	0.127	0.997	0.602	1.000	0.862	1.000	0.941	1.000	0.963	1.000	0.985
b=6.0	0.047	0.048	0.418	0.156	1.000	0.731	1.000	0.960	1.000	0.988	1.000	0.995	1.000	0.997
b=8.0	0.043	0.046	0.470	0.169	1.000	0.828	1.000	0.985	1.000	0.998	1.000	0.999	1.000	1.000
n=50														
b=0.5	0.036	0.044	0.151	0.114	0.980	0.415	1.000	0.555	1.000	0.671	1.000	0.785	1.000	0.947
b=3.0	0.047	0.047	0.598	0.220	1.000	0.829	1.000	0.959	1.000	0.985	1.000	0.990	1.000	0.999
b=4.0	0.052	0.049	0.778	0.264	1.000	0.941	1.000	0.993	1.000	0.998	1.000	1.000	1.000	1.000
b=6.0	0.050	0.050	0.889	0.351	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=8.0	0.043	0.046	0.926	0.407	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. III - C: Values of the Power (ε_t ~ Generalized Logistic) - AR (1) Model Without Intercept Term (Continue)

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=100														
b=0.5	0.047	0.043	0.866	0.320	1.000	0.830	1.000	0.944	1.000	0.993	1.000	0.999	1.000	1.000
b=3.0	0.049	0.048	1.000	0.669	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=4.0	0.048	0.049	1.000	0.788	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=6.0	0.050	0.049	1.000	0.921	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=8.0	0.048	0.053	1.000	0.970	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=300														
b=0.5	0.051	0.054	1.000	0.967	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=3.0	0.050	0.050	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=4.0	0.051	0.048	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=6.0	0.052	0.055	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=8.0	0.046	0.051	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. IV - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student – t) – AR (1) Model With Intercept Term

ϕ_1=1.0	n=20			n=30			n=50			n=100			n=300		
	MML	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
p=2.0															
Mean	0.793	0.772		0.859	0.839		0.917	0.901		0.958	0.948		0.986	0.982	
Stnd dev.	0.178	0.183	95	0.124	0.130	91	0.077	0.083	86	0.038	0.043	78	0.013	0.015	75
Skewness	1.346	1.070		1.819	1.484		3.125	2.460		2.477	2.085		2.688	2.155	
Kurtosis	4.996	4.964		5.726	5.764		9.360	7.964		7.218	6.388		7.537	6.369	
p=2.5															
Mean	0.780	0.768		0.851	0.840		0.909	0.899		0.955	0.949		0.985	0.982	
Stnd dev.	0.181	0.184	97	0.124	0.128	94	0.081	0.085	91	0.039	0.043	82	0.013	0.015	75
Skewness	1.122	0.927		1.489	1.221		2.364	1.962		2.334	2.103		2.373	2.091	
Kurtosis	4.678	4.639		5.243	5.004		6.911	6.258		6.913	6.571		6.946	6.317	
p=3.5															
Mean	0.778	0.771		0.843	0.836		0.905	0.899		0.952	0.948		0.984	0.982	
Stnd dev.	0.179	0.181	98	0.128	0.130	97	0.080	0.082	95	0.041	0.043	91	0.014	0.015	87
Skewness	1.112	0.997		1.496	1.347		1.819	1.692		2.419	2.285		2.344	2.183	
Kurtosis	4.714	4.673		5.163	4.980		5.823	5.540		7.294	6.909		6.785	6.370	
p=5.0															
Mean	0.770	0.767		0.841	0.838		0.901	0.899		0.949	0.948		0.983	0.982	
Stnd dev.	0.183	0.183	100	0.130	0.131	98	0.080	0.081	98	0.042	0.044	91	0.014	0.015	87
Skewness	1.154	1.091		1.393	1.332		1.821	1.795		1.983	1.967		2.289	2.381	
Kurtosis	4.701	4.571		5.171	5.022		6.469	6.418		5.984	5.963		6.597	6.819	
p=10.0															
Mean	0.770	0.769		0.839	0.838		0.899	0.898		0.948	0.948		0.983	0.982	
Stnd dev.	0.182	0.182	100	0.130	0.130	100	0.081	0.082	98	0.043	0.043	100	0.015	0.015	100
Skewness	1.135	1.127		1.514	1.486		1.592	1.615		1.814	1.829		2.544	2.543	
Kurtosis	4.817	4.820		5.470	5.435		5.367	5.451		5.655	5.692		7.211	7.210	

Table 3. IV - B: Type I Error (ε_t ~ Student – t) – AR (1) Model With Intercept Term

	p=2.5		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
n=20	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	6.980*	7.044	6.972*	7.033*	6.963	7.006	7.037	7.053*	7.049*	7.053*
Type I Error	0.052	0.053	0.052	0.052	0.050	0.051	0.051	0.052	0.053	0.053
n=30										
Critical Value	7.619	7.714	7.642	7.708	7.658*	7.716*	7.657*	7.679*	7.731	7.736
Type I Error	0.051	0.050	0.052	0.052	0.053	0.054	0.052	0.052	0.052	0.053
n=50										
Critical Value	8.778	8.900	8.805	8.891	8.851	8.904	8.819	8.844	8.868*	8.875*
Type I Error	0.050	0.051	0.052	0.051	0.051	0.049	0.052	0.053	0.053	0.052
n=100										
Critical Value	11.213	11.339	11.243	11.334	11.281	11.352	11.308	11.342	11.320*	11.327*
Type I Error	0.050	0.050	0.052	0.0509	0.050	0.050	0.049	0.051	0.052	0.052
n=300										
Critical Value	18.020	18.113	18.089	18.163	18.068	18.122	18.096	18.123	18.112	18.117
Type I Error	0.048	0.047	0.050	0.050	0.052	0.052	0.049	0.049	0.051	0.052

Note: For (*) values, Three- Moment Chi- Square Approximation, is valid; other cases satisfy Four- Moment Approximation

Table 3. IV - C: Values of the Power (ε_t ~ Student – t) – AR (1) Model With Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=20														
p=2.5	0.052	0.053	0.056	0.057	0.075	0.079	0.096	0.099	0.121	0.124	0.156	0.166	0.258	0.265
p=3.0	0.052	0.052	0.058	0.057	0.075	0.073	0.099	0.101	0.125	0.133	0.163	0.167	0.265	0.274
p=3.5	0.050	0.051	0.055	0.058	0.073	0.074	0.100	0.103	0.130	0.130	0.168	0.172	0.261	0.268
p=5.0	0.051	0.052	0.056	0.057	0.075	0.077	0.103	0.104	0.136	0.136	0.167	0.170	0.267	0.270
p=10.0	0.053	0.053	0.057	0.056	0.076	0.076	0.097	0.098	0.138	0.138	0.171	0.171	0.266	0.264
n=30														
p=2.5	0.051	0.050	0.059	0.058	0.084	0.088	0.122	0.128	0.173	0.176	0.250	0.252	0.454	0.448
p=3.0	0.052	0.052	0.063	0.061	0.084	0.084	0.121	0.127	0.183	0.184	0.248	0.251	0.460	0.455
p=3.5	0.053	0.054	0.051	0.060	0.081	0.081	0.129	0.131	0.189	0.197	0.249	0.253	0.454	0.453
p=5.0	0.052	0.052	0.057	0.058	0.089	0.090	0.127	0.129	0.186	0.187	0.258	0.257	0.462	0.457
p=10.0	0.052	0.053	0.059	0.059	0.091	0.090	0.131	0.130	0.184	0.185	0.262	0.262	0.441	0.438
n=50														
p=2.5	0.050	0.051	0.065	0.064	0.109	0.113	0.187	0.190	0.330	0.321	0.531	0.496	0.889	0.837
p=3.0	0.052	0.051	0.064	0.065	0.109	0.107	0.206	0.208	0.325	0.321	0.527	0.507	0.872	0.842
p=3.5	0.051	0.049	0.064	0.063	0.103	0.104	0.193	0.195	0.343	0.333	0.530	0.509	0.847	0.821
p=5.0	0.052	0.053	0.060	0.059	0.109	0.110	0.199	0.199	0.333	0.332	0.514	0.504	0.830	0.818
p=10.0	0.053	0.052	0.063	0.062	0.111	0.112	0.200	0.200	0.342	0.341	0.504	0.501	0.823	0.821

Table 3. IV - C: Values of the Power (ε_t ~ Student – t) – AR (1) Model With Intercept Term (Continue)

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=100														
p=2.5	0.050	0.050	0.066	0.067	0.190	0.190	0.549	0.495	0.870	0.791	0.985	0.960	1.000	1.000
p=3.0	0.052	0.051	0.074	0.072	0.196	0.191	0.519	0.487	0.848	0.792	0.978	0.962	1.000	1.000
p=3.5	0.050	0.050	0.072	0.071	0.199	0.196	0.517	0.491	0.830	0.787	0.973	0.957	1.000	1.000
p=5.0	0.049	0.051	0.076	0.076	0.193	0.192	0.480	0.471	0.818	0.801	0.971	0.964	1.000	1.000
p=10.0	0.052	0.052	0.073	0.074	0.203	0.203	0.488	0.486	0.780	0.776	0.963	0.961	1.000	1.000
n=300														
p=2.5	0.048	0.047	0.125	0.130	0.874	0.782	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=3.0	0.050	0.050	0.124	0.125	0.862	0.786	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=3.5	0.052	0.052	0.121	0.126	0.836	0.777	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=5.0	0.049	0.049	0.127	0.128	0.810	0.782	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p=10.0	0.051	0.052	0.123	0.124	0.787	0.778	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. V - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Gamma) – AR (1) Model With Intercept Term

ϕ_1=1.0	n=20			n=30			n=50			n=100			n=300		
	MML	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
k=3.0			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	0.998	0.996		0.999	0.998		1.000	0.999		1.000	1.000		1.000	1.000	
Stnd dev.	0.017	0.023	55	0.009	0.012	56	0.004	0.006	44	0.001	0.002	25	0.000	0.000	100
Skew.	0.002	0.000		0.001	0.000		0.000	0.000		0.000	0.000		0.003	0.000	
Kurtosis	3.485	3.272		3.507	3.308		3.365	3.271		3.172	3.052		3.118	2.980	
k=4.0															
Mean	0.998	0.997		0.999	0.999		1.000	1.000		1.000	1.000		1.000	1.000	
Stnd dev.	0.016	0.020	64	0.009	0.011	67	0.004	0.005	64	0.001	0.002	25	0.000	0.000	100
Skew.	0.001	0.003		0.000	0.002		0.000	0.001		0.000	0.000		0.001	0.000	
Kurtosis	3.446	3.418		3.405	3.182		3.150	3.202		3.086	3.017		3.036	3.042	
k=5.0															
Mean	0.998	0.998		0.999	0.999		1.000	1.000		1.000	1.000		1.000	1.000	
Stnd dev.	0.015	0.018	69	0.008	0.010	64	0.004	0.004	100	0.001	0.002	25	0.000	0.000	100
Skew.	0.003	0.006		0.001	0.001		0.001	0.001		0.000	0.000		0.005	0.000	
Kurtosis	3.398	3.309		3.337	3.312		3.165	3.079		3.093	3.039		3.050	3.118	
k=8.0															
Mean	0.999	0.998		0.999	0.999		1.000	1.000		1.000	1.000		1.000	1.000	
Stnd dev.	0.013	0.014	86	0.007	0.008	77	0.003	0.003	100	0.001	0.001	100	0.000	0.000	100
Skew.	0.000	0.008		0.000	0.000		0.000	0.000		0.001	0.000		0.001	0.000	
Kurtosis	3.268	3.242		3.168	3.152		3.102	3.022		3.051	3.051		3.125	3.177	

Table 3. V - B: Type I Error (ε_t ~ Gamma) – AR (1) Model With Intercept Term

	k=3.0		**k=4.0**		**k=5.0**		**k=8.0**	
n=20	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	4.611	4.658	4.601	4.631	4.591	4.613	4.572	4.581
Type I Error	0.047	0.049	0.050	0.048	0.051	0.049	0.046	0.047
n=30								
Critical Value	5.563	5.598	5.560	5.583	5.554	5.570	5.541	5.550
Type I Error	0.049	0.049	0.048	0.049	0.050	0.049	0.051	0.049
n=50								
Critical Value	7.120	7.143	7.118	7.133	7.115	7.124	7.108	7.114
Type I Error	0.051	0.049	0.049	0.048	0.049	0.049	0.049	0.047
n=100								
Critical Value	10.023	10.035	10.022	10.030	10.021	10.027	10.018	10.021
Type I Error	0.050	0.051	0.051	0.050	0.049	0.046	0.047	0.050
n=300								
Critical Value	17.328	17.332	17.328	17.330	17.327	17.329	17.327	17.327
Type I Error	0.052	0.046	0.050	0.051	0.050	0.051	0.050	0.049

Note: For all values, Three- Moment Chi- Square Approximation is valid

Table 3. V - C: Values of the Power (ε_t ~ Gamma) – AR (1) Model With Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=20														
k=3.0	0.047	0.049	0.162	0.135	0.847	0.697	0.988	0.953	0.998	0.989	0.999	0.997	1.000	0.999
k=4.0	0.050	0.048	0.170	0.146	0.867	0.765	0.994	0.976	0.999	0.995	1.000	0.998	1.000	1.000
k=5.0	0.051	0.049	0.187	0.161	0.899	0.825	0.996	0.987	1.000	0.998	1.000	1.000	1.000	1.000
k=8.0	0.046	0.047	0.222	0.201	0.955	0.923	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
n=30														
k=3.0	0.049	0.049	0.324	0.235	0.991	0.944	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	0.048	0.049	0.351	0.281	0.993	0.968	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	0.050	0.049	0.377	0.309	0.997	0.987	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.051	0.049	0.452	0.398	0.999	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=50														
k=3.0	0.051	0.049	0.777	0.551	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	0.048	0.048	0.801	0.634	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	0.049	0.049	0.836	0.714	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.049	0.047	0.906	0.847	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=100														
k=3.0	0.050	0.051	1.000	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	0.051	0.050	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	0.049	0.046	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.047	0.050	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=300														
k=3.0	0.052	0.046	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=4.0	0.050	0.051	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=5.0	0.050	0.051	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
k=8.0	0.050	0.049	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. VI - A: Moment – Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Generalized Logistic) – AR (1) Model With Intercept Term

μ=0.0 σ=1.0 ϕ_1=1.0	**n=20**			**n=30**			**n=50**			**n=100**			**n=300**		
	MML	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
b=0.5															
Mean	0.945	0.939		0.976	0.973		0.992	0.991		0.998	0.998		1.000	1.000	
Stnd dev.	0.102	0.105	94	0.050	0.052	92	0.020	0.022	83	0.006	0.007	73	0.001	0.001	100
Skewness	5.467	4.164		5.493	4.277		1.383	1.046		0.027	0.028		0.000	0.001	
Kurtosis	14.040	12.820		17.219	15.157		9.339	8.493		3.916	3.840		3.420	3.272	
b=3.0															
Mean	0.990	0.988		0.995	0.994		0.998	0.998		1.000	1.000		1.000	1.000	
Stnd dev.	0.037	0.039	90	0.019	0.021	82	0.009	0.010	81	0.003	0.003	100	0.001	0.001	100
Skewness	0.021	0.016		0.000	0.000		0.000	0.000		0.000	0.001		0.003	0.004	
Kurtosis	4.301	4.269		3.696	3.730		3.296	3.342		3.171	3.176		3.023	2.965	
b=4.0															
Mean	0.994	0.993		0.997	0.996		0.999	0.999		1.000	1.000		1.000	1.000	
Stnd dev.	0.028	0.030	87	0.015	0.016	88	0.007	0.008	77	0.002	0.003	44	0.000	0.000	100
Skewness	0.002	0.003		0.001	0.000		0.001	0.002		0.000	0.000		0.000	0.000	
Kurtosis	3.628	3.622		3.386	3.674		3.107	3.182		3.216	3.118		2.913	3.971	

Table 3. VI - A: Moment – Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Generalized Logistic) – AR (1) Model With Intercept Term (Continue)

μ=0.0 σ=1.0 ϕ_1=1.0	**n=20**			**n=30**			**n=50**			**n=100**			**n=300**		
	MML	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
b=6.0															
Mean	0.997	0.996		0.998	0.998		0.999	0.999		1.000	1.000		1.000	1.000	
Stnd dev.	0.021	0.024	77	0.011	0.013	72	0.005	0.006	69	0.002	0.002	100	0.000	0.000	100
Skewness	0.001	0.001		0.002	0.002		0.000	0.000		0.001	0.000		0.000	0.001	
Kurtosis	3.460	3.393		3.135	3.338		3.194	3.194		3.045	3.127		2.981	3.011	
b=8.0															
Mean	0.997	0.997		0.999	0.999		1.000	0.999		1.000	1.000		1.000	1.000	
Stnd dev.	0.018	0.020	81	0.009	0.011	67	0.004	0.005	64	0.002	0.002	100	0.000	0.000	100
Skewness	0.000	0.000		0.003	0.002		0.000	0.000		0.001	0.000		0.002	0.000	
Kurtosis	3.284	3.475		3.350	3.323		3.111	3.164		3.041	3.001		2.909	3.012	

Table 3. VI - B: Type I Error (ε_t ~ Generalized Logistic) – AR (1) Model With Intercept Term

	b=0.5		**b=3.0**		**b=4.0**		**b=6.0**		**b=8.0**	
n=20	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	5.545*	5.649	4.794	4.818	4.707	4.733	4.643	4.666	4.614	4.635
Type I Error	0.045	0.041	0.045	0.045	0.047	0.049	0.046	0.046	0.047	0.048
n=30										
Critical Value	6.136	6.185	5.677	5.695	5.627	5.586	5.586	5.604	5.571	5.586
Type I Error	0.037	0.037	0.048	0.046	0.049	0.048	0.048	0.048	0.050	0.050
n=50										
Critical Value	7.371	7.415	7.185	7.197	7.157	7.168	7.134	7.144	7.125	7.135
Type I Error	0.037	0.038	0.049	0.049	0.042	0.045	0.050	0.053	0.050	0.050
n=100										
Critical Value	10.118	10.135	10.055	10.061	10.041	10.047	10.031	10.036	10.026	10.031
Type I Error	0.048	0.046	0.050	0.049	0.049	0.050	0.045	0.050	0.050	0.050
n=300										
Critical Value	17.355	17.361	17.338	17.340	17.334	17.336	17.330	17.332	17.329	17.331
Type I Error	0.053	0.051	0.052	0.048	0.052	0.047	0.048	0.050	0.046	0.045

Note: For (*) values, Four- Moment F Approximation is valid; other cases satisfy Three- Moment Chi - Square Approximation

Table 3. VI - C: Values of the Power (ε_t ~ Generalized Logistic) - AR (1) Model With Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=20														
b=0.5	0.045	0.041	0.065	0.058	0.167	0.157	0.325	0.354	0.511	0.503	0.666	0.653	0.868	0.849
b=3.0	0.045	0.045	0.093	0.094	0.491	0.467	0.817	0.793	0.940	0.931	0.981	0.974	0.997	0.995
b=4.0	0.047	0.049	0.119	0.115	0.596	0.563	0.907	0.881	0.977	0.965	0.992	0.988	0.999	0.998
b=6.0	0.046	0.046	0.145	0.128	0.756	0.687	0.973	0.947	0.994	0.988	0.999	0.997	1.000	0.999
b=8.0	0.047	0.048	0.162	0.145	0.827	0.751	0.987	0.975	0.998	0.996	1.000	0.999	1.000	0.999
n=30														
b=0.5	0.037	0.037	0.067	0.066	0.336	0.343	0.974	0.963	0.859	0.842	0.955	0.946	0.994	0.993
b=3.0	0.048	0.046	0.164	0.155	0.792	0.762	0.973	0.966	0.995	0.991	0.999	0.998	1.000	1.000
b=4.0	0.049	0.048	0.126	0.120	0.891	0.853	0.991	0.984	0.999	0.997	1.000	0.999	1.000	1.000
b=6.0	0.048	0.048	0.263	0.230	0.965	0.931	1.000	0.997	1.000	0.999	1.000	1.000	1.000	1.000
b=8.0	0.037	0.038	0.309	0.268	0.987	0.968	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
n=50														
b=0.5	0.037	0.038	0.152	0.146	0.779	0.757	0.976	0.972	0.998	0.997	1.000	0.999	1.000	1.000
b=3.0	0.049	0.049	0.371	0.343	0.982	0.975	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
b=4.0	0.042	0.045	0.467	0.418	0.997	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=6.0	0.050	0.053	0.630	0.548	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=8.0	0.050	0.050	0.727	0.622	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. VI - C: Values of the Power ($\varepsilon_t \sim$ Generalized Logistic) - AR (1) Model With Intercept Term (Continue)

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=100														
b=0.5	0.048	0.046	0.563	0.528	0.997	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=3.0	0.050	0.049	0.890	0.847	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=4.0	0.049	0.050	0.965	0.936	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=6.0	0.045	0.050	0.996	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=8.0	0.050	0.050	0.999	0.995	1.000	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000
n=300														
b=0.5	0.053	0.051	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=3.0	0.052	0.048	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=4.0	0.052	0.047	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=6.0	0.048	0.050	1.100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
b=8.0	0.046	0.045	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3. VII - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student–t) - AR (1) Outlier Model Without Intercept Term

ϕ_1=1.0	p=2.5			p=3.0			p=3.5			p=5.0			p=10.0		
n=50	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
e(i)=2 x e(i)			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	0.978	0.969		0.975	0.968		0.974	0.969		0.970	0.968		0.968	0.968	
Stnd dev.	0.048	0.056	73	0.052	0.059	78	0.053	0.058	84	0.056	0.059	90	0.057	0.058	97
Skewness	6.025	5.329		6.102	5.940		6.355	5.987		5.351	5.418		5.091	5.172	
Kurtosis	12.267	10.927		12.704	12.272		13.11	12.286		10.79	10.91		10.544	10.676	
e(i)=4 x e(i)															
Mean	0.979	0.967		0.976	0.965		0.973	0.964		0.970	0.964		0.966	0.963	
Stnd dev.	0.049	0.066	55	0.052	0.067	60	0.055	0.068	65	0.059	0.070	71	0.064	0.069	86
Skewness	10.610	9.568		9.607	8.935		8.282	8.204		9.393	8.749		8.163	8.031	
Kurtosis	20.095	17.492		17.922	16.848		14.87	15.00		16.77	15.42		14.601	14.233	

Table 3. VII - B: Type I Error ($\varepsilon_t \sim$ Student – t) - AR (1) Outlier Model Without Intercept Term

	p=2.5		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
n=50	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
e(i)=2 x e(i)										
Critical Value	7.917*	8.097*	7.958	8.098	7.985	8.108	8.070	8.121*	8.147	8.168*
Type I Error	0.050	0.051	0.050	0.050	0.052	0.050	0.052	0.052	0.051	0.050
e(i)=4 x e(i)										
Critical Value	7.877	8.204	7.972*	8.257*	8.054*	8.302*	8.076	8.269*	8.202*	8.299*
Type I Error	0.047	0.053	0.048	0.052	0.049	0.050	0.051	0.052	0.050	0.052

Note: For (*) values, Three- Moment Chi- Square Approximation, is valid; other cases satisfy Four – Moment F Approximation.

Table 3. VII - C: Values of the Power (ε_t ~ Student – t) - AR (1) Outlier Model Without Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
n=50	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
e(i)=2 x e(i)														
p=2.5	0.050	0.051	0.061	0.062	0.166	0.155	0.410	0.338	0.759	0.615	0.934	0.829	0.997	0.984
p=3.0	0.050	0.050	0.064	0.063	0.170	0.158	0.385	0.338	0.705	0.610	0.878	0.793	0.993	0.977
p=3.5	0.052	0.050	0.062	0.063	0.161	0.151	0.380	0.336	0.682	0.610	0.874	0.805	0.996	0.985
p=5.0	0.052	0.052	0.062	0.062	0.158	0.152	0.366	0.343	0.618	0.582	0.845	0.801	0.988	0.975
p=10.0	0.051	0.050	0.064	0.064	0.152	0.150	0.357	0.346	0.610	0.591	0.831	0.812	0.982	0.975
e(i)=4 x e(i)														
p=2.5	0.047	0.053	0.064	0.063	0.173	0.153	0.430	0.323	0.798	0.554	0.937	0.713	0.998	0.937
p=3.0	0.048	0.052	0.063	0.062	0.161	0.146	0.404	0.317	0.722	0.541	0.897	0.708	0.993	0.924
p=3.5	0.049	0.050	0.065	0.066	0.160	0.149	0.400	0.333	0.650	0.521	0.864	0.703	0.988	0.909
p=5.0	0.051	0.052	0.061	0.061	0.149	0.146	0.367	0.320	0.636	0.516	0.821	0.675	0.981	0.905
p=10.0	0.050	0.052	0.062	0.064	0.160	0.156	0.351	0.329	0.553	0.500	0.778	0.702	0.950	0.894

Table 3. VIII - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student–t) - AR(1) Outlier Model With Intercept Term

ϕ_1=1.0	p=2.5			p=3.0			p=3.5			p=5.0			p=10.0		
n=100	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
e(i)=2 x e(i)			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	0.941	0.934		0.938	0.933		0.939	0.934		0.935	0.932		0.933	0.932	
Stnd dev.	0.047	0.050	88	0.048	0.050	92	0.048	0.050	92	0.050	0.051	96	0.049	0.050	96
Skewness	1.965	1.828		1.948	1.854		1.960	1.933		1.777	1.745		1.465	1.464	
Kurtosis	6.258	6.111		6.186	6.154		6.413	6.400		5.608	5.553		5.034	.5.022	
e(i)=4 x e(i)															
Mean	0.910	0.899		0.907	0.897		0.905	0.897		0.904	0.896		0.899	0.895	
Stnd dev.	0.065	0.068	91	0.065	0.067	94	0.065	0.067	94	0.065	0.067	94	0.066	0.067	97
Skewness	2.237	2.042		1.925	1.878		2.050	1.927		1.808	1.754		1.859	1.854	
Kurtosis	7.097	6.803		6.144	5.983		6.769	6.446		6.063	5.881		5.907	5.891	

Table 3. VIII - B: Type I Error (ε_t ~ Student – t) - AR (1) Outlier Model With Intercept Term

	p=2.5		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
n=100	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
e(i)=2 x e(i)										
Critical Value	11.502	11.621	11.540	11.630	11.521	11.598	11.575	11.625	11.591	11.614
Type I Error	0.050	0.049	0.052	0.052	0.050	0.050	0.050	0.050	0.050	0.050
e(i)=4 x e(i)										
Critical Value	12.155	12.314	12.204	12.342	12.169	12.294	12.165	12.295	12.234	12.297
Type I Error	0.047	0.047	0.049	0.051	0.051	0.051	0.049	0.050	0.053	0.054

Note: All values satisfy Four- Moment F Approximation.

Table 3. VIII - C: Values of the Power ($\varepsilon_t \sim$ Student – t) - AR (1) Outlier Model With Intercept Term

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
n=100	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
e(i)=2 x e(i)														
p=2.5	0.050	0.049	0.064	0.067	0.100	0.115	0.303	0.312	0.645	0.595	0.926	0.866	1.000	0.997
p=3.0	0.052	0.052	0.058	0.061	0.111	0.128	0.297	0.304	0.673	0.634	0.903	0.854	0.998	0.991
p=3.5	0.050	0.050	0.065	0.065	0.106	0.116	0.308	0.316	0.647	0.623	0.894	0.854	0.999	0.995
p=5.0	0.050	0.050	0.068	0.069	0.116	0.123	0.323	0.328	0.665	0.641	0.899	0.873	0.998	0.994
p=10.0	0.050	0.050	0.067	0.068	0.130	0.135	0.329	0.329	0.639	0.632	0.855	0.843	0.995	0.992
e(i)=4 x e(i)														
p=2.5	0.047	0.047	0.040	0.051	0.031	0.065	0.084	0.160	0.241	0.319	0.545	0.545	0.930	0.845
p=3.0	0.049	0.051	0.045	0.050	0.041	0.068	0.093	0.152	0.274	0.330	0.521	0.520	0.910	0.840
p=3.5	0.051	0.051	0.048	0.054	0.047	0.073	0.119	0.168	0.277	0.329	0.547	0.541	0.899	0.829
p=5.0	0.049	0.050	0.050	0.053	0.051	0.076	0.123	0.174	0.320	0.355	0.565	0.554	0.894	0.840
p=10.0	0.053	0.054	0.054	0.057	0.062	0.078	0.158	0.186	0.334	0.353	0.557	0.551	0.856	0.820

Table 3. IX - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student – t) - AR (1) Model Without Intercept Term -Mixture Model:0.90LTS(p, σ^2)+0.10LTS(p, kσ^2); k=2,4

ϕ_1=1.0	p=2.5			p=3.0			p=3.5			p=5.0			p=10.0		
n=50	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
e(i)=2 x e(i)			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	0.977	0.967		0.974	0.966		0.971	0.967		0.969	0.967		0.968	0.967	
Stnd dev.	0.050	0.060	69	0.053	0.060	78	0.054	0.060	81	0.057	0.060	90	0.058	0.059	97
Skewness	4.638	3.781		4.500	4.048		5.266	4.496		4.813	4.460		4.316	4.309	
Kurtosis	10.122	8.562		10.173	9.277		12.749	10.718		10.769	10.178		10.280	10.350	
e(i)=4 x e(i)															
Mean	0.981	0.969		0.978	0.968		0.978	0.970		0.974	0.968		0.972	0.969	
Stnd dev.	0.044	0.060	54	0.047	0.060	61	0.046	0.058	63	0.050	0.060	69	0.055	0.061	81
Skewness	7.877	5.131		5.493	4.554		4.697	3.635		4.917	4.043		6.616	5.839	
Kurtosis	17.967	11.684		11.749	10.439		10.410	8.534		10.731	9.464		16.224	14.198	

Table 3. IX - B: Type I Error (ε_t ~ Student – t) - AR (1) Model Without Intercept Term Mixture Model:0.90LTS(p, σ^2)+0.10LTS(p,kσ^2); k=2,4

	p=2.5		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
n=50	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
e(i)=2 x e(i)										
Critical Value	7.925	8.120	8.005	8.148	7.995	8.120	8.078	8.142	8.132	8.158
Type I Error	0.052	0.051	0.048	0.048	0.051	0.050	0.048	0.050	0.051	0.050
e(i)=4 x e(i)										
Critical Value	7.793	8.100	7.823	8.084	7.868	8.097	7.947	8.123	8.001	8.103
Type I Error	0.049	0.051	0.049	0.049	0.051	0.049	0.049	0.053	0.050	0.050

Note: All cases satisfy Four- Moment F Approximation.

Table 3. IX - C: Values of the Power ($\varepsilon_t \sim$ Student – t) - AR (1) Model Without Intercept Term Mixture Model: 0.90LTS(p, σ^2)+0.10LTS(p, kσ^2); k=2,4

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
n=50	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
e(i)=2 x e(i)														
p=2.5	0.052	0.051	0.062	0.065	0.162	0.142	0.424	0.334	0.735	0.586	0.916	0.795	0.995	0.982
p=3.0	0.048	0.049	0.063	0.069	0.154	0.145	0.401	0.330	0.703	0.598	0.888	0.794	0.993	0.981
p=3.5	0.051	0.050	0.064	0.064	0.155	0.149	0.369	0.317	0.679	0.581	0.876	0.789	0.994	0.981
p=5.0	0.048	0.050	0.062	0.062	0.152	0.145	0.357	0.321	0.622	0.577	0.823	0.764	0.986	0.975
p=10.0	0.051	0.050	0.062	0.063	0.141	0.139	0.345	0.334	0.604	0.579	0.815	0.794	0.984	0.980
e(i)=4 x e(i)														
p=2.5	0.049	0.051	0.068	0.065	0.194	0.144	0.555	0.338	0.853	0.589	0.977	0.832	0.999	0.988
p=3.0	0.049	0.049	0.065	0.067	0.180	0.141	0.472	0.322	0.824	0.590	0.975	0.835	0.999	0.987
p=3.5	0.051	0.049	0.063	0.063	0.175	0.148	0.472	0.333	0.806	0.594	0.956	0.818	0.999	0.988
p=5.0	0.049	0.053	0.063	0.066	0.168	0.142	0.470	0.344	0.752	0.564	0.946	0.817	0.999	0.984
p=10.0	0.050	0.050	0.064	0.062	0.161	0.148	0.387	0.330	0.665	0.572	0.921	0.838	0.996	0.986

Table 3. X - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student – t) - AR (1) Model With Intercept Term - Mixture Model:0.90LTS(p, σ^2)+0.10LTS(p, kσ^2); k=2,4

ϕ_1=1.0	p=2.5			p=3.0			p=3.5			p=5.0			p=10.0		
n=100	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**	**MML**	**LS**	**%**
e(i)=2 x e(i)			**Eff**			**Eff**			**Eff**			**Eff**			**Eff**
Mean	0.956	0.948		0.954	0.948		0.952	0.948		0.951	0.948		0.949	0.947	
Stnd dev.	0.040	0.044	83	0.040	0.043	87	0.041	0.043	91	0.041	0.043	91	0.042	0.043	95
Skewness	2.937	2.578		2.100	2.077		2.246	2.085		2.165	2.065		1.991	1.961	
Kurtosis	8.449	7.590		6.357	6.442		6.577	6.316		6.490	6.241		6.244	6.200	
e(i)=4 x e(i)															
Mean	0.962	0.949		0.959	0.948		0.958	0.948		0.958	0.949		0.954	0.949	
Stnd dev.	0.036	0.042	73	0.037	0.043	74	0.038	0.043	78	0.037	0.042	78	0.40	0.043	87
Skewness	2.957	2.093		2.777	2.054		3.268	2.432		2.330	1.860		2.345	2.084	
Kurtosis	8.096	6.394		7.726	6.303		9.743	7.714		6.789	6.015		6.793	6.362	

Table 3. X - B: Type I Error (ε_t ~ Student – t) - AR (1) Model With Intercept Term –Mixture Model: 0.90LTS(p, σ^2)+0.10LTS(p,k σ^2); k=2,4

	p=2.5		**p=3.0**		**p=3.5**		**p=5.0**		**p=10.0**	
n=100	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
e(i)=2 x e(i)										
Critical Value	11.173	11.336	11.253	11.371	11.256	11.345	11.293	11.362	11.310	11.338
Type I Error	0.050	0.048	0.050	0.050	0.052	0.052	0.050	0.051	0.052	0.053
e(i)=4 x e(i)										
Critical Value	11.082	11.344	11.138	11.355	11.148	11.348	11.137	11.337	11.226	11.337
Type I Error	0.049	0.048	0.049	0.050	0.049	0.047	0.049	0.049	0.051	0.051

Note: All cases satisfy Four- Moment F Approximation.

Table 3. X - C: Values of the Power ($\varepsilon_t \sim$ Student – t) - AR (1) Model With Intercept Term – Mixture Model: 0.90LTS(p, σ^2)+0.10LTS(p, kσ^2); k=2,4

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
n=100	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
e(i)=2 x e(i)														
p=2.5	0.050	0.048	0.071	0.071	0.196	0.195	0.551	0.484	0.902	0.805	0.991	0.964	1.000	1.000
p=3.0	0.050	0.050	0.069	0.072	0.192	0.194	0.533	0.480	0.878	0.805	0.983	0.955	1.000	1.000
p=3.5	0.052	0.052	0.075	0.076	0.198	0.199	0.512	0.471	0.859	0.795	0.984	0.959	1.000	1.000
p=5.0	0.050	0.051	0.072	0.073	0.204	0.198	0.507	0.479	0.834	0.789	0.976	0.959	1.000	1.000
p=10.0	0.052	0.053	0.067	0.067	0.181	0.180	0.480	0.467	0.814	0.792	0.966	0.956	1.000	1.000
e(i)=4 x e(i)														
p=2.5	0.049	0.048	0.066	0.071	0.177	0.189	0.606	0.470	0.962	0.806	0.999	0.971	1.000	1.000
p=3.0	0.049	0.050	0.066	0.066	0.192	0.186	0.585	0.468	0.951	0.807	0.998	0.962	1.000	1.000
p=3.5	0.049	0.047	0.068	0.069	0.200	0.205	0.590	0.484	0.937	0.791	0.997	0.964	1.000	1.000
p=5.0	0.049	0.049	0.067	0.071	0.192	0.191	0.581	0.479	0.939	0.821	0.997	0.966	1.000	1.000
p=10.0	0.051	0.051	0.065	0.068	0.188	0.187	0.558	0.502	0.897	0.816	0.990	0.965	1.000	1.000

Table 3. XI - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 ($\varepsilon_t \sim$ Student – t) - AR (1) Model Without Intercept Term -Contamination Model:0.90LTS(p, σ^2)+0.10U(-0.5,0.5)

ϕ_1=1.0	p=2.5			p=3.0			p=3.5			p=5.0			p=10.0		
n=50	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
Mean	0.976	0.966		0.974	0.967		0.971	0.967		0.968	0.966		0.967	0.966	
Stnd dev.	0.052	0.062	70	0.055	0.062	79	0.054	0.060	81	0.059	0.061	94	0.060	0.061	97
Skewness	5.295	4.512		5.295	4.724		3.951	3.816		5.045	4.432		4.564	4.500	
Kurtosis	11.213	9.725		11.331	10.146		8.730	8.417		11.536	10.150		10.428	10.263	

Table 3. XI - B: Type I Error ($\varepsilon_t \sim$ Student – t) - AR (1) Model Without Intercept Term Contamination Model: 0.90LTS(p, σ^2)+0.10U(-0.5,0.5)

	p=2.5		p=3.0		p=3.5		p=5.0		p=10.0	
n=50	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	7.946	8.139	8.021*	8.165*	8.048	8.140	8.125	8.174*	8.145*	8.162*
Type I Error	0.051	0.050	0.049	0.049	0.049	0.050	0.049	0.049	0.049	0.050

Note: For (*) values, Three- Moment Chi- Square Approximation, is valid; other cases satisfy Four- Moment F Approximation

Table 3. XI - C: Values of the Power (ε_t ~ Student – t) - AR (1) Model Without Intercept Term -Contamination Model: 0.90LTS(p, σ^2)+0.10U(-0.5,0.5)

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=50														
p=2.5	0.051	0.050	0.063	0.064	0.161	0.144	0.401	0.322	0.698	0.567	0.903	0.791	0.994	0.981
p=3.0	0.049	0.049	0.066	0.065	0.164	0.149	0.377	0.319	0.656	0.559	0.887	0.798	0.994	0.982
p=3.5	0.049	0.050	0.065	0.065	0.155	0.144	0.377	0.326	0.619	0.550	0.859	0.800	0.994	0.983
p=5.0	0.049	0.049	0.063	0.062	0.157	0.153	0.336	0.320	0.582	0.539	0.834	0.790	0.988	0.980
p=10.0	0.049	0.050	0.062	0.061	0.144	0.143	0.313	0.311	0.593	0.582	0.808	0.791	0.981	0.978

Table 3. XII - A: Moment - Based Measures for the Sampling Distribution of ϕ_1 (ε_t ~ Student – t) - AR (1) Model With Intercept Term -Contamination Model: 0.90LTS(p, σ^2)+0.10U(-0.5,0.5)

ϕ_1=1.0	p=2.5			p=3.0			P=3.5			p=5.0			p=10.0		
n=100	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**	**MML**	**LS**	**% Eff**
Mean	0.955	0.948	82	0.954	0.949	86	0.953	0.948	86	0.950	0.947	91	0.949	0.948	95
Stnd dev.	0.039	0.043		0.039	0.042		0.039	0.042		0.042	0.044		0.042	0.043	
Skewness	2.309	2.051		2.181	2.047		1.805	1.723		2.143	2.114		2.138	2.123	
Kurtosis	6.998	6.522		6.899	6.622		5.876	5.721		6.683	6.655		6.727	6.729	

Table 3. XII - B: Type I Error (ε_t ~ Student – t) - AR (1) Model With Intercept Term Contamination Model: 0.90LTS(p, σ^2)+0.10U(-0.5,0.5)

	p=2.5		p=3.0		p=3.5		p=5.0		p=10.0	
n=100	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0	d_1	d_0
Critical Value	11.194	11.342	11.249	11.363*	11.270	11.359	11.298	11.350	11.346	11.361
Type I Error	0.050	0.051	0.049	0.051	0.049	0.049	0.049	0.051	0.051	0.051

Note: For (*) value, Three- Moment Chi- Square Approximation is valid; other cases satisfy Four- Moment F Approximation.

Table 3. XII - C: Values of the Power (ε_t ~ Student – t) - AR (1) Model With Intercept Term Contamination Model: 0.90LTS(p, σ^2)+0.10U(-0.5,0.5)

	ϕ_1 **=1.0**		ϕ_1 **=0.99**		ϕ_1 **=0.95**		ϕ_1 **=0.90**		ϕ_1 **=0.85**		ϕ_1 **=0.80**		ϕ_1 **=0.70**	
	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0	R_1	R_0
n=100														
p=2.5	0.050	0.051	0.071	0.070	0.206	0.200	0.534	0.465	0.895	0.807	0.991	0.964	1.000	1.000
p=3.0	0.049	0.051	0.070	0.073	0.193	0.191	0.536	0.488	0.871	0.801	0.985	0.961	1.000	1.000
p=3.5	0.049	0.049	0.072	0.071	0.196	0.192	0.519	0.488	0.831	0.783	0.979	0.956	1.000	1.000
p=5.0	0.049	0.051	0.073	0.075	0.197	0.199	0.491	0.473	0.834	0.801	0.973	0.960	1.000	1.000
p=10.0	0.051	0.051	0.073	0.071	0.193	0.194	0.488	0.479	0.803	0.790	0.962	0.957	1.000	0.999