University of Central Florida

STARS

Electronic Theses and Dissertations

2012

Autoregressive Models

Kelly Wade
University of Central Florida

Part of the Mathematics Commons

Find similar works at: https://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Wade, Kelly, "Autoregressive Models" (2012). *Electronic Theses and Dissertations*. 2322.
https://stars.library.ucf.edu/etd/2322

AUTOREGRESSIVE MODELS

by

KELLY WADE
B.A.S. Daytona State College, 2010

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Mathematics
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2012

© 2012 Kelly Wade

ABSTRACT

Consider a sequence of random variables which obeys a first order autoregressive model with unknown parameter alpha. Under suitable assumptions on the error structure of the model, the limiting distribution of the normalized least squares estimator of alpha is discussed. The choice of the normalizing constant depends on whether alpha is less than one, equals one, or is greater than one in absolute value. In particular, the limiting distribution is normal provided that the absolute value of alpha is less than one, but is a function of Brownian motion whenever the absolute value of alpha equals one. Some general remarks are made whenever the sequence of random variables is a first order moving average process.

This is dedicated to my parents, Floyd and Lois Wade, who have supported me in all my life's endeavors with love and encouragement. Also, to my love, Lauren Axelrod, who has been the best roommate and study buddy a college student could hope to have. Finally, to my son, Chase Wade; I hope that this will be an encouragement to you. You can do anything you set your mind to.

ACKNOWLEDGMENTS

Sincere thanks to Gary Richardson. Your instruction and persistence has been of great benefit. Very warm thanks to Norma Robles. She has been a constant resource for any and all things U.C.F. and C.O.S. related. Special thanks to Ram Mohapatra and Marianna Pensky. Finally, a great cheer of appreciation for Xin Li who gave me a wonderful opportunity.

TABLE OF CONTENTS

INTRODUCTION

Consider a sequence $\{Y_t\}$, $t \geq 0$ of random variables defined on a probability space $(\Omega, \Im, P)$ that obey the first order autoregressive model $Y_t = \alpha Y_{t-1} + \varepsilon_t$, where α is an unknown parameter. The error structure $\{\varepsilon_t\}$ is assumed to be IID, independent individually distributed, with $E(\varepsilon_1) = 0$ and $0 < var\varepsilon_1 = \sigma^2 < \infty$. No specific distribution of the errors is assumed.

This leads to using the least squares procedure to estimate α. Let $\hat{\alpha}_n$ denote the least squares estimator (LSE) of α, which is a function of $Y_1, Y_2, \ldots, Y_n$. In addition to estimating α, the limiting distribution of $\hat{\alpha}_n$, whenever properly normalized and centered, can lead to a test of hypothesis for α. If the limiting distribution is unknown, then a simulation study can be used to estimate tail probabilities for a test of hypothesis.

The LSE $\hat{\alpha}_n$ is the value of α which makes $Q_n(\alpha) = \sum_{t=1}^{n}(Y_t - \alpha Y_{t-1})^2$ a minimum. In particular, $\hat{\alpha}_n = \frac{\sum_{t=1}^{n} Y_t Y_{t-1}}{\sum_{t=1}^{n}(Y_{t-1})^2}$. substituting the model assumption $Y_t = \alpha Y_{t-1} + \varepsilon_t$ gives:

$$\hat{\alpha}_n = \frac{\sum_{t=1}^{n}(\alpha Y_{t-1} + \varepsilon_t) Y_{t-1}}{\sum_{t=1}^{n}(Y_{t-1})^2}, \text{ and}$$

$$\hat{\alpha}_n - \alpha = \frac{\sum_{t=1}^{n}(Y_{t-1}\varepsilon_t)}{\sum_{t=1}^{n}(Y_{t-1})^2}. \qquad (1.1)$$

The purpose of this work is to show that $\{c_n(\hat{\alpha}_n - \alpha)\}$ converges in distribution whenever the normalizing $c_n{}'s$ are properly chosen. The choice of c_n varies according to the cases: $|\alpha| < 1$, $\alpha = 1$, and $|\alpha| > 1$. The case $|\alpha| < 1$, ($\alpha = 1$, $|\alpha| < 1$) is studied in chapter 2 (chapter 3, chapter 4), respectively.

The Martingale Central Limit Theorem is the primary tool used to show that whenever $|\alpha| < 1$, the limiting distribution is normal. Whenever $\alpha = 1$ the limiting distribution is shown to be a function of Brownian motion. Donsker's Functional Central Limit Theorem is essential here. Order in probability techniques are utilized in the $|\alpha| > 1$ case.

The results presented here are known. However, this work collects them into one resource with detailed proofs. It is hoped that a convenient comparison of the three cases of the AR (1) model is of interest to the reader. Each case requires uniquely different methods for proof. Excellent references on the theory of Time Series are the books written by Brockwell and Davis [2] and Fuller [4]. Hasza [3] investigated the case whenever $|\alpha| > 1$ in his Ph.D. dissertation.

Finally, an outline of the primary ideas used in proving that the Gauss-Newton estimator is asymptotically normal for the order one moving average model is given in the last section.

AR (1)-CASE I: $|\boldsymbol{\alpha}| < 1$

It is assumed throughout this section that $\{Y_t\}, t \geq 0$ is a sequence of random variables defined on the probability space $(\Omega, \Im, P)$ that obeys the model:

$$Y_t = \alpha Y_{t-1} + \varepsilon_t\,, \;\; |\alpha| < 1, \tag{2.1}$$

where $\varepsilon_t\,, t = 0, \pm1, \pm2, \ldots$ are IID random variables with $E(\varepsilon_1) = 0$ and $0 < var\varepsilon_1 = \sigma^2 < \infty$. Recall that the least squares estimator (LSE) of α is given by:

$$\hat{\alpha}_n = \frac{\sum_{t=1}^{n}(Y_t Y_{t-1})}{\sum_{t=1}^{n}(Y_{t-1})^2}. \tag{2.2}$$

Using the model assumption (2.1), $\hat{\alpha}_n = \frac{\sum_{t=1}^{n}(\alpha Y_{t-1}+\varepsilon_t)Y_{t-1}}{\sum_{t=1}^{n}(Y_{t-1})^2}$, and thus,

$$\hat{\alpha}_n - \alpha = \frac{\sum_{t=1}^{n}(Y_{t-1}\varepsilon_t\,)}{\sum_{t=1}^{n}(Y_{t-1})^2}. \tag{2.3}$$

Recall that a sequence $\{V_n\}$ of random variables defined $(\Omega, \Im, P)$ converges in distribution to V provided that $F_n \to F$ pointwise except possibly at values where F is discontinuous. Here F_n is the distribution function of V_n and F is the distribution function of V. The above is denoted by $V_n \overset{D}{\to} V$. Moreover, $V_n \to V$ in probability if for each $\delta > 0$,

$P\{|V_n - V| > \delta\} \to 0$, as $n \to \infty$, denoted by $V_n \overset{P}{\to} V$. Also, $\{V_n\}$ converges to V in $L^p, p > 0$, , provided $E|V_n - V|^p \to 0$ as $n \to \infty$, denoted by $V_n \overset{L^p}{\to} V$. Listed below are some basic properties of these convergence notions.

Theorem 2.1 Assume that $\{V_n\}$ and $\{W_n\}$ are sequences defined on $(\Omega, \Im, P)$. Suppose that $f: R \to R$ is continuous except for a set of P^V measure zero. Then:

(i) $V_n \overset{L^p}{\to} V \Rightarrow V_n \overset{P}{\to} V \Rightarrow V_n \overset{D}{\to} V, p > 0$

(ii) $V_n \overset{D}{\to} V \; iff \; V_n \overset{P}{\to} V$ whenever V is a constant rv, a.s.

(iii) $V_n - W_n \overset{P}{\to} 0$ and $W_n \overset{P}{\to} W \Rightarrow V_n \overset{P}{\to} W$

(iv) $V_n - W_n \overset{P}{\to} 0$ and $W_n \overset{D}{\to} W \Rightarrow V_n \overset{D}{\to} W$

(v) $V_n \overset{P}{\to} V \Rightarrow f(V_n) \overset{P}{\to} f(V)$

(vi) $V_n \overset{D}{\to} V \Rightarrow f(V_n) \overset{D}{\to} f(V)$

(vii) $V_n \overset{D}{\to} V \; iff \; \varphi_{V_n} \to \varphi_V$ pointwise, where φ_{V_n} denotes the characteristic function of V_n.

The two results stated below are used to show that $\left\{\frac{1}{n}\sum_{t=1}^{n} Y_t^2\right\}$ converges in probability. These results can be found in Proposition 6.3.9 and Proposition 6.3.10 of Brockwell and Davis [2].

Theorem 2.2 Suppose that $\{V_n\}\ n \geq 1, \{W_{nk}\}\ 1 \leq k \leq n, \{W_k\}, k \geq 1$, and W denotes random variables defined on $(\Omega, \Im, P)$. If:

(i) $W_{nk} \xrightarrow{D} W_k$ as $n \to \infty$, for each fixed $k \geq 1$.

(ii) $W_k \xrightarrow{D} W$ as $k \to \infty$.

(iii) $\lim_{k\to\infty} \text{limsup}_{n\to\infty} P\{|V_n - W_{nk}| > \delta\} = 0$, for each fixed $\delta > 0$. Then

$V_n \xrightarrow{D} W$ as $n \to \infty$.

The next result is called the "Weak Law of Large Numbers for Moving Averages."

Theorem 2.3 Let $\{Z_t :\ t = 0, \pm 1, \pm 2, \ldots\}$ be a sequence of IID random variables defined on $(\Omega, \Im, P)$, where $E|Z_1| < \infty$. If $\{\varphi_j\}, j \geq 0$ is a sequence of real numbers with $\sum_{j=0}^{\infty} |\varphi_j| < \infty$, define $V_t = \sum_{j=0}^{\infty} \varphi_j Z_{t-j}$. Then $\frac{1}{n} \sum_{t=1}^{n} V_t \xrightarrow{P} (\sum_{j=0}^{\infty} \varphi_j) E(Z_1)$.

Assume that $\{Y_t\}\ t \geq 0$ obeys model (2.1), then:

$Y_1 = \alpha Y_0 + \epsilon_1, Y_2 = \alpha Y_1 + \epsilon_2 = \alpha(\alpha Y_0 + \epsilon_1) + \epsilon_2 = \alpha^2 Y_0 + \alpha\epsilon_1 + \epsilon_2 , Y_3 = \alpha Y_2 + \epsilon_3 = \alpha(\alpha^2 Y_0 + \alpha\epsilon_1 + \epsilon_2) + \epsilon_3 = \alpha^3 Y_0 + \alpha^2\epsilon_1 + \alpha\epsilon_2 + \epsilon_3.$

In general, an induction argument shows that:

$$Y_t = \alpha^t Y_0 + \sum_{i=1}^{t} \alpha^{t-i} \epsilon_i, \text{ for all } t \geq 1. \tag{2.4}$$

Theorem 2.4 Suppose that $\{Y_t\}\ t \geq 0$ obeys model (2.1), then:

$$\frac{1}{n}\sum_{t=1}^{n} Y_t^2 \xrightarrow{P} \frac{\sigma^2}{1-\alpha^2}.$$

Proof: Since only asymptotic results are needed, one can assume that $Y_0 \equiv 0$ and thus, $Y_t = \sum_{i=1}^{t} \alpha^{t-i}\, \epsilon_{i.}$. Note that $Y_t = \sum_{j=0}^{t-1} \alpha^j\, \epsilon_{t-j}$, and hence $Y_t^2 = \sum_{i,j=0}^{t-1} \alpha^i \alpha^j\, \epsilon_{t-i}\epsilon_{t-j}$. Define $U_t = \sum_{i,j=0}^{\infty} \alpha^i \alpha^j\, \epsilon_{t-i}\epsilon_{t-j}$, $V_t = \sum_{i=0}^{\infty} \alpha^{2i}\, \epsilon_{t-i}^2$, $W_t = \sum_{i,j=0, i\neq j}^{\infty} \alpha^i \alpha^j\, \epsilon_{t-i}\epsilon_{t-j}$; then

$U_t = V_t + W_t$. Observe that $E(V_t) = \sum_{i=0}^{\infty} \alpha^{2i}\sigma^2 = \frac{\sigma^2}{1-\alpha^2}$ and $E|W_t| \leq \sum_{i,j=0\ ,i\neq j}^{\infty} |\alpha|^i |\alpha|^j\, (E(\epsilon_1))^2 < \infty$. Hence, V_t and W_t are finite a.s.

Observe that:

$$E|U_t - Y_t^2| = E\left|\sum_{i,j=0,}^{\infty} \alpha^i \alpha^j\, \epsilon_{t-i}\epsilon_{t-j} - \sum_{i,j=0}^{t-1} \alpha^i \alpha^j\, \epsilon_{t-i}\epsilon_{t-j}\right|$$

$$\leq \sum_{i=0,j=t}^{\infty} |\alpha^i||\alpha^j|\sigma^2 + \sum_{i=t,j=0}^{\infty} |\alpha^i||\alpha^j|\sigma^2 \leq M|\alpha|^t \to 0 \text{ as } t \to \infty.$$

It follows that $U_t - Y_t^2 \xrightarrow{L_1} 0$, and thus, $E\left|\frac{1}{n}\sum_{t=1}^{n}(U_t - Y_t^2)\right| \leq \frac{1}{n}\sum_{t=1}^{T} E|U_t - Y_t^2| + \frac{1}{n}\sum_{t=T+1}^{n} E|U_t - Y_t^2|$. Given, $\delta > 0$, choose T such that $E|U_t - Y_t^2| < \frac{\delta}{2}$ for all $t \geq T$.

Moreover, choose $N > T$ for which $\frac{1}{n}\sum_{t=1}^{T} E|U_t - Y_t^2| < \frac{\delta}{2}$ for all $n \geq N$. Hence,

$E\left|\frac{1}{n}\sum_{t=1}^{n}(U_t - Y_t^2)\right| < \delta$ for all $n \geq N$, and thus $\frac{1}{n}\sum_{t=1}^{n}(U_t - Y_t^2) \xrightarrow{L_1} 0$.

Next, it is shown that $\frac{1}{n}\sum_{t=1}^{n} U_t \xrightarrow{P} \frac{\sigma^2}{1-\alpha^2}$. It follows from Theorem 2.3 that $\frac{1}{n}\sum_{t=1}^{n} V_t \xrightarrow{P} \sum_{i=0}^{\infty} \alpha^{2i}\sigma^2 = \frac{\sigma^2}{1-\alpha^2}$. It remains to use Theorem 2.2 to verify that $\frac{1}{n}\sum_{t=1}^{n} W_t \xrightarrow{P} 0$. Note that $\frac{1}{n}\sum_{t=1}^{n} W_t = \sum_{i,j=0,i\neq j}^{\infty} \alpha^i\alpha^j \cdot \frac{1}{n}\sum_{t=1}^{n} \epsilon_{t-i}\epsilon_{t-j}$. Observe that $cov(\epsilon_{s-i}\epsilon_{s-j}, \epsilon_{t-i}\epsilon_{t-j}) = 0$ whenever $s \neq t$. Indeed, if $s - i = t - j$ and $s - j = t - i, i = j$, which is contrary to the definition of W_t. It follows that for $i \neq j$, $var\left(\frac{1}{n}\sum_{t=1}^{n} \epsilon_{t-i}\epsilon_{t-j}\right) = \frac{1}{n^2}\sum_{t=1}^{n} \sigma^4 = \frac{\sigma^4}{n}$, and thus by Chebyshev's inequality, $\frac{1}{n}\sum_{t=1}^{n} \epsilon_{t-i}\epsilon_{t-j} \xrightarrow{P} 0 \; as \; n \to \infty$. Denote $T_n = \sum_{i,j=0,i\neq j}^{\infty} \alpha^i\alpha^j \cdot \frac{1}{n}\sum_{t=1}^{n} \epsilon_{t-i}\epsilon_{t-j}$, $\quad T_{nk} = \sum_{i,j=0,i\neq j}^{k} \alpha^i\alpha^j \cdot \frac{1}{n}\sum_{t=1}^{n} \epsilon_{t-i}\epsilon_{t-j}$, and note that $T_{nk} \xrightarrow{P} T_k \equiv 0 \; as \; n \to \infty$.

Observe that: $|T_n - T_{nk}| \leq \sum_{i=o,j=k+1}^{k,\infty}|\alpha^i||\alpha^j|\frac{1}{n}\sum_{t=1}^{n}|\epsilon_{t-i}||\epsilon_{t-j}| +$

$\sum_{i=k+1,j=0}^{\infty,k}|\alpha^i||\alpha^j|\frac{1}{n}\sum_{t=1}^{n}|\epsilon_{t-i}||\epsilon_{t-j}| + \sum_{i,j=k+1,i\neq j}^{\infty}|\alpha^i||\alpha^j||\epsilon_{t-i}||\epsilon_{t-j}|$, and thus,

$$E|T_n - T_{nk}| \leq \frac{|\alpha^{k+1}|}{(1-|\alpha|)^2}E|\epsilon_1|E|\epsilon_2| + \frac{|\alpha^{k+1}|}{(1-|\alpha|)^2}E|\epsilon_1|E|\epsilon_2| + \frac{|\alpha^{k+1}|^2}{(1-|\alpha|)^2}\; E|\epsilon_1|E|\epsilon_2|\;.$$

Then $lim_{k\to\infty}\, limsup_{n\to\infty}\; E|T_n - T_{nk}| = 0$, and thus by Theorem 2.2, $T_n \xrightarrow{P} 0$. However, $T_n = \frac{1}{n}\sum_{t=1}^{n} W_t$, and $\frac{1}{n}\sum_{t=1}^{n} U_t = \frac{1}{n}\sum_{t=1}^{n} V_t + \frac{1}{n}\sum_{t=1}^{n} W_t \xrightarrow{P} \frac{\sigma^2}{1-\alpha^2}$. It was shown above that $\frac{1}{n}\sum_{t=1}^{n}(U_t - Y_t^2) \xrightarrow{L_1} 0$, and thus by Theorem 2.1 (i), $\frac{1}{n}\sum_{t=1}^{n}(U_t - Y_t^2) \xrightarrow{P} 0$. Since $\frac{1}{n}\sum_{t=1}^{n} U_t$

$\xrightarrow{P} \frac{\sigma^2}{1-\alpha^2}$ it follows from Theorem 2.1 (iii) that $\frac{1}{n}\sum_{t=1}^{n} Y_t^2 \xrightarrow{P} \frac{\sigma^2}{1-\alpha^2}$ ■

Recall from (1.1) that $\sqrt{n}(\hat{\alpha}_n - \alpha) = \frac{\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(Y_{t-1}\epsilon_t)}{\frac{1}{n}\sum_{t=1}^{n}(Y_{t-1})^2}$. It remains to show that the limiting distribution of $\left\{\frac{1}{n}\sum_{t=1}^{n} Y_{t-1}\epsilon_t\right\}$ is normal. The Martingale Central Limit Theorem is used to verify this. The notion of uniform integrability of a sequence of random variables is used in the proof. In particular, a sequence $\{V_t, t \geq 1\}$ of random variables defined on $(\Omega, \Im, P)$ is said to be uniformly integrable provided that for each $\delta > 0$ there exists a real number c such that for all $t \geq 1, E\left[|V_t| \cdot 1_{\{|V_t|>c\}}\right] < \delta.$ Observe that uniform integrability implies uniform boundedness, that is, $sup_{t\geq 1} E|V_t| < \infty$.

Lemma 2.1 Let $\{V_t, t \geq 1\}$ be a sequence of random variables defined on $(\Omega, \Im, P)$ satisfying $sup_{t\geq 1} E|V_t|^p = K < \infty$, where $p > 1$. Then, $\{V_t, t \geq 1\}$ is uniformly integrable.

Proof: Given $\delta > 0$, note that $\lim_{x\to\infty} \frac{|x|^p}{|x|} = +\infty$, and thus, there exists $c > 0$ such that for all $|x| > c$, $\frac{|x|^p}{|x|} > \frac{2K}{\delta}$. Then for all $|x| > c$, $|x| < \frac{|x|^p}{2K} \cdot \delta$, and thus $E\left|V_t \cdot 1_{\{|V_t|>c\}}\right| \leq \frac{E|V_t|^p}{2K} \cdot \delta$ $< \delta$, for all $t \geq 1$. Hence $\{V_t, t \geq 1\}$ is uniformly integrable. ■

A basic ingredient of the Martingale Central Limit Theorem is the concept of conditional expectation. More precisely, let X denote an integrable random variable defined on $(\Omega, \Im, P)$. Let $\mathfrak{Y} \subseteq \mathcal{F}$ be a sub σ - field of $\mathcal{F}$. The condition expectation of X is defined by $E(X|\mathfrak{Y}) \coloneqq Y$ such that Y is $\mathfrak{Y}$ -measurable and $\int_A XdP = \int_A YdP$ for each $A \in \mathfrak{Y}$. The existence and uniqueness (almost surely) of Y is based on Radon-Nikodym Theorem. A list of some basic properties of conditional expectation is given below. Let $\sigma(X)$ denote the smallest σ -field such that X is measurable, that is, $\sigma(X) = \{X^{-1}(B): B \in \mathcal{B}(\mathcal{R})\}$, where $\mathcal{B}(\mathcal{R})$ denotes the Borel σ-field on $\mathcal{R}$.

Theorem 2.5 Assume that X, Y are integrable random variables defined on $(\Omega, \Im, P)$, and let $\mathfrak{Y} \subseteq \mathcal{H} \subseteq \mathcal{F}$ sub- σ-fields of $\mathcal{F}$. The following results are valid a.s:

(i) $E(X|\mathfrak{Y}) = X$ whenever X is $\mathfrak{Y}$ -measurable

(ii) $E[E(X|\mathfrak{Y})] = E(X)$

(iii) $E((aX + bY)|\mathfrak{Y}) = aE(X|\mathfrak{Y}) + bE(Y|\mathfrak{Y})$, where a, b are real numbers

(iv) $X \leq Y$ implies $E(X|\mathfrak{Y}) \leq E(Y|\mathfrak{Y})$

(v) $E(XY|\mathfrak{Y}) = YE(X|\mathfrak{Y})$, whenever Y is $\mathfrak{Y}$ -measurable and XY is integrable

(vi) $E(E(X|\mathcal{H})|\mathfrak{Y}) = E(X|\mathfrak{Y})$

(vii) $E(X|\mathfrak{Y}) = E(X)$ provided that $\sigma(X)$ and $\mathfrak{Y}$ are independent σ-fields.

Note: The Martingale Central Limit Theorem suited for our context is stated below.

Theorem 2.6 Let $\{S_n, n \geq 0\}$ be a sequence of mean zero square integrable random variables defined on $(\Omega, \Im, P)$. Assume that $\{S_n, \Im_n, n \geq 1\}$ is a martingale, where $\{\Im_n, n \geq 0\}$ is an increasing sequence of sub- σ-fields of $\Im$. Suppose that $S_0 \equiv 0$ and $X_n \coloneqq S_n - S_{n-1}$. Denote $\sigma_n^2 = E(X_n^2|\Im_{n-1})$, $V_n^2 = \sum_{j=1}^{n} \sigma_j^2$, and $s_n^2 = E(V_n^2) = E(S_n^2)$. Assume that:

(i) $\frac{V_n^2}{s_n^2} \xrightarrow{P} 1$

(ii) For each $\delta > 0, \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{j=1}^{n} E(X_j^2 \cdot 1_{\{|X_j| > \delta s_n\}}) = 0,$

Then $\frac{S_n}{s_n} \xrightarrow{D} \mathcal{Z} \sim \mathcal{N}(0,1)$.

Theorem 2.7 Suppose that $\{Y_t, t \geq 0\}$ obeys model (2.1); moreover, assume that

$E|\epsilon_1|^{2+r} < \infty$ for some $r > 0$. Then $\frac{\sqrt{1-\alpha^2}}{\sigma^2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} Y_{t-1}\epsilon_t \xrightarrow{D} \mathcal{Z} \sim \mathcal{N}(0,1)$.

Proof: The Martingale Central Limit Theorem is used to verify this. Denote $\sigma^2 = var\epsilon_1$, $S_n = \sum_{t=1}^{n} Y_{t-1}\epsilon_t$, $\sigma_t^2 = E(Y_{t-1}^2\epsilon_t^2|\mathcal{F}_{t-1})$, $s_n^2 = E(V_n^2) = E(S_n^2)$, and fix $\delta > 0$. Here $\mathcal{F}_t = \sigma\{\epsilon_1, \epsilon_2, \dots \epsilon_t\}$, that is, the smallest σ -field such that each σ_i, $1 \leq i \leq t$ is measurable.

Note that, $E(S_{n+1}|\mathcal{F}_n) = E\big((S_n + Y_n\epsilon_{n+1})\big|\mathcal{F}_n\big) = S_n + E(Y_n\epsilon_{n+1}|\mathcal{F}_n) = S_n + Y_nE(\epsilon_{n+1}|\mathcal{F}_n) = S_n + Y_nE(\epsilon_n) = S_n$. Hence $\{S_n, \mathcal{F}_n, n \geq 1\}$ is a martingale. Denote $\sigma_t^2 = E(Y_{t-1}^2\epsilon_t^2|\mathcal{F}_{t-1}) =$

$Y_{t-1}^2 E(\epsilon_t^2|\mathcal{F}_{t-1}) = Y_{t-1}^2 E(\epsilon_1^2) = Y_{t-1}^2\sigma^2$, where $\mathcal{F}_0 = \{\phi, \Omega\}$ and $Y_0 \equiv 0$ is $\mathcal{F}_0$-measurable. Define $V_n^2 = \sum_{t=1}^n \sigma_t^2 = \sigma^2 \sum_{t=1}^n Y_{t-1}^2$, and $s_n^2 = E(V_n^2) = E(S_n^2) = \sigma^2 \sum_{t=1}^n varY_{t-1}$. Note that since $Y_t = \alpha Y_{t-1} + \epsilon_t$, $t \geq 0$, $Y_t = \alpha^t Y_0 + \sum_{i=1}^t \alpha^{t-i}\epsilon_i = \alpha^t Y_0 + \sum_{i=0}^{t-1} \alpha^i \epsilon_{t-i}$. Since $|\alpha| < 1$, one can select $Y_0 \equiv 0$ since we are interested in asymptotic results. Let $Y_t = \sum_{i=0}^{t-1}\alpha^i\epsilon_{t-i}$, then $varY_t = \sum_{i=0}^{t-1}\alpha^{2i}\sigma^2 = \frac{(1-\alpha^{2t})\sigma^2}{1-\alpha^2}$. Then $s_n^2 = \sigma^4 \sum_{t=1}^n \frac{1-\alpha^{2t}}{1-\alpha^2} = \frac{\sigma^4}{1-\alpha^2}[n - \frac{(1-\alpha^{2n})\alpha^2}{1-\alpha^2}]$. It follows from Theorem 2.4 that $\frac{V_n^2}{s^2} = \frac{\sigma^2 \frac{1}{n}\sum_{t=1}^n Y_{t-1}^2}{\frac{\sigma^4}{1-\alpha^2}[1 - \frac{\alpha^2(1-\alpha^{2n})}{n}]} \xrightarrow{P} \frac{\frac{\sigma^4}{(1-\alpha^2)}}{\frac{\sigma^4}{(1-\alpha^2)}} = 1$ as $n \to \infty$, i.e. $\frac{V_n^2}{s_n^2} \xrightarrow{P} 1$ as $n \to \infty$.

Next, it is shown that $\frac{1}{s_n^2}\sum_{t=1}^n E[Y_{t-1}^2\,\epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > \delta s_n\}}] \to 0$ for $\delta > 0$ fixed. First it is shown that $\{Y_{t-1}^2\epsilon_t^2\}_{t\geq 1}$ is uniformly integrable. According to Lemma 2.1, it is sufficient to show that for $r > 0$, given in the hypothesis, $sup_{t\geq 1} E|Y_{t-1}^2\epsilon_t^2|^{1+\frac{r}{2}} = sup_{t\geq 1}E|Y_{t-1}^2\epsilon_t^2|^{2+r} < \infty$. Let $p = \frac{2+r}{1+r}, q = 2 + r$; then $p > 1$, $q > 1$, and $\frac{1}{p} + \frac{1}{q} = 1$. Then using Holder's inequality,

$$|Y_{t-1}|^{2+r} \leq \left|\sum_{i=0}^{t-1}|\alpha|^i|\epsilon_{t-i}|\right|^{2+r} \leq \left|\sum_{i=0}^{t-1}\left(|\alpha|^i\right)^{\frac{1}{p}}\left(|\alpha|^i\right)^{\frac{1}{q}}|\epsilon_{t-i}|\right|^{2+r} \leq$$

$$\left|\left(\sum_{i=0}^{t-1}|\alpha|^i\right)^{\frac{1+r}{2+r}}\left(\sum_{i=0}^{t-1}|\alpha|^i|\epsilon_{t-i}|^{2+r}\right)^{\frac{1}{2+r}}\right|^{2+r} = \left(\sum_{i=0}^{t-1}|\alpha|^i\right)^{1+r}\sum_{i=0}^{t-1}|\alpha|^i|\epsilon_{t-i}|^{2+r}.$$

Then, $E|Y_{t-1}\epsilon_t|^{2+r} = E|\epsilon_1|^{2+r}E|\epsilon_1|^{2+r}(\sum_{i=0}^{t-1}|\alpha|^i)^{2+r} = O(1)$, and thus $\{Y_{t-1}^2\epsilon_t^2\}_{t\geq 1}$ is uniformly integrable. Hence, given $\delta > 0$ there exists $M_\delta > 0$ such that

$E[Y_{t-1}^2\epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > M_\delta\}}] < \delta$ for all $t \geq 1$.

Since $s_n^2 = O(n)$, choose n_0 such that for all $\geq n_0$, $\delta s_n \geq M_\delta$. Then for all $n \geq n_0$, $E[Y_{t-1}^2 \epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > \delta s_n\}}] < \delta$ and thus, $\frac{1}{n}\sum_{t=n_0}^{n} E[Y_{t-1}^2\, \epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > \delta s_n\}}] < \delta$ for all $n \geq n_0$. Choose n_1 such that for all $n \geq n_1$, $\frac{1}{n}\sum_{t=1}^{n} E[Y_{t-1}^2\, \epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > \delta s_n\}}] < \delta$. However, $s_n^2 \sim \frac{n\alpha^4}{1-\alpha^2}$ as $n \to \infty$ and thus, $\frac{1}{s_n^2}\sum_{t=1}^{n} E[Y_{t-1}^2\, \epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > \delta \epsilon_n\}}] < \delta$, for n sufficiently large. Hence:

$lim_{n\to\infty} \frac{1}{s_n^2}\sum_{t=1}^{n} E[Y_{t-1}^2\, \epsilon_t^2 \cdot 1_{\{|Y_{t-1}\epsilon_t| > \delta \epsilon_n\}}] = 0.$ It follows from Martingale Convergence Theorem that $\frac{1}{s_n}\sum_{t=1}^{n} Y_{t-1}\epsilon_t \xrightarrow{D} Z \sim N(0,1)$. Again $s_n^2 \sim \frac{n\sigma^4}{1-\alpha^2}$, implies that $\frac{\sqrt{1-\alpha^2}}{\sigma^2}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Y_{t-1}\epsilon_t \xrightarrow{D} Z$. ■

Theorem 2.8 Suppose that the hypotheses of Theorem 2.7 are satisfied and let $\hat{\alpha}_n$ denote the LSE of α. Then $\sqrt{n}(\hat{\alpha}_n - \alpha) \xrightarrow{D} N(0,1-\alpha^2)$.

Proof: It follows that $\sqrt{n}(\hat{\alpha}_n - \alpha) = \frac{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\frac{1}{n}\sum_{t=1}^{n} Y_{t-1}^2}$, and by Theorem 2.4 and Theorem 2.7 above, $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} Y_{t-1}\epsilon_t \xrightarrow{D} \frac{\sigma^2}{\sqrt{1-\alpha^2}} Z$ and $\frac{1}{n}\sum_{t=1}^{n} Y_{t-1}^2 \xrightarrow{P} \frac{\sigma^2}{1-\alpha^2}$. Hence,

$\sqrt{n}(\hat{\alpha}_n - \alpha) \xrightarrow{D} \sqrt{1-\alpha^2}\, Z = N(0,1-\alpha^2)$. ■

AR (1)-CASE II: $\boldsymbol{\alpha = 1}$

Suppose that $\{Y_t\ , t \geq 0\}$ is a sequence of random variables defined on $(\Omega, \Im, P)$ which obeys the model:

$$Y_t = \alpha Y_{t-1} + \epsilon_t \tag{3.1}$$

where $\alpha = 1$, and $\epsilon_t, t \geq 1$ are IID mean zero with $0 < \text{var}\epsilon_1 = \sigma^2 < \infty$. Since $\alpha = 1$, it follows from (1.1) that the LSE of α obeys

$$\widehat{\alpha}_n - 1 = \frac{\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2} \tag{3.2}$$

If $\alpha = 1$ then $Y_t = Y_{t-1} + \epsilon_t$, and the time series$\{Y_t\}$ is called a random walk. Many papers have been written about when the process is a random walk. This is of interest to economists. Testing the hypothesis$H_O: \alpha = 1$ vs. $H_A: |\alpha| < 1$ is discussed below.

Lemma 3.1 Assume that $\{Y_t\}_{t\geq 0}$ is a time series obeying the model $Y_t = Y_{t-1} + \epsilon_t\ , t \geq 1$, and that $\{\epsilon_t\}$ is an IID sequence with $E(\epsilon_1) = 0$ and $var\epsilon_1 = \sigma^2 < \infty$. Then $\frac{1}{n\sigma^2}\sum_{t=1}^{n} Y_{t-1}\epsilon_t \xrightarrow{D} \frac{Z^2}{2} - \frac{1}{2}$ as $n \to \infty$, where $Z \sim N(0,1)$.

Proof: Iterating on $Y_t = Y_{t-1} + \epsilon_t\ , t \geq 1\ , Y_1 = Y_0 + \epsilon_1, Y_2 = Y_1 + \epsilon_2 = \ Y_0 + \epsilon_1 + \epsilon_2, Y_3 =$

$Y_2 + \epsilon_3 = Y_0 + \epsilon_1 + \epsilon_2 + \epsilon_3$, etc. Then $Y_{t-1} = Y_0 + \sum_{i=1}^{t-1} \epsilon_i$, and thus $\sum_{t=1}^{n} Y_{t-1}\epsilon_t =$

$\sum_{t=1}^{n}(Y_0 + \sum_{i=1}^{t-1} \epsilon_i)\epsilon_t = Y_0 \sum_{t=1}^{n} \epsilon_t + \sum_{t=1}^{n} \sum_{i=1}^{t-1} \epsilon_i \, \epsilon_t = Y_0 \sum_{t=1}^{n} \epsilon_t + \frac{1}{2}(\sum_{t=1}^{n} \epsilon_t)^2 - \frac{1}{2}\sum_{t=1}^{n} {\epsilon_t}^2.$

Using Strong Law of Large Numbers and Central Limit Theorem, $\frac{1}{n\sigma^2}\sum_{t=1}^{n} Y_{t-1}\epsilon_t =$

$\frac{Y_0}{n\sigma^2}\sum_{t=1}^{n} \epsilon_t + \frac{1}{2}(\sum_{t=1}^{n} \frac{\epsilon_t}{\sqrt{n}\sigma})^2 - \frac{1}{2n\sigma^2}\sum_{t=1}^{n} {\epsilon_t}^2 \xrightarrow{D} \frac{Y_0 E(\epsilon_1)}{\sigma^2} + \frac{1}{2}Z^2 - \frac{1}{2\sigma^2}E({\epsilon_1}^2) = 0 + \frac{1}{2}Z^2 - \frac{1}{2} =$

$\frac{1}{2}Z^2 - \frac{1}{2}$ as $n \to \infty$. ■

A stochastic process $\{X_t : t \in T\}$ defined on a probability space $(\Omega, \mathfrak{I}, P)$ is called a Gaussian process provided that for each $t_1 \leq t_2 \leq \cdots \leq t_k$ in T, $(X_{t_1}, X_{t_2}, \dots , X_{t_k})$ is a k -dimensional normal random vector, that is, $\sum_{i=1}^{k} \lambda_i X_{t_i}$ is a one –dimensional normal random variable, for each $\lambda' = (\lambda_1, \lambda_2, \cdots , \lambda_k) \in \mathbb{R}^k$. This includes the degenerate case, where $\sum_{1}^{k} \lambda_i X_{t_i}$ is constant almost surely Moreover, a stochastic process $\{W_t : t \in T\}$ is called a Wiener process or Brownian motion provided:

(i) $\{W_t\}$ is a mean zero, Guassian process.

(ii) $cov(W_s, W_t) = (s \wedge t)\sigma^2$, $s, t \in T$.

Here σ^2 is a parameter. The index set T is chosen to be $I = [0,1]$ in our context.

Lemma 3.2 A stochastic process $\{W_t : t \in I\}$ is a Wiener process if and only if:

(i) $W_0 = 0$ almost surely

(ii) $W_t - W_s \sim N(0, (t-s)\sigma^2)$ whenever $s \leq t$

(iii) $W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, \cdots, W_{t_k} - W_{t_{k-1}}$ are independent random variables for $0 \leq t_1 < t_2 < t_3 \cdots < t_k \leq 1$.

Proof: Assume that $\{W_t : t \in I\}$ obeys (i), (ii), and (iii). It is shown that the process is Gaussian. Let us check this for $k = 3$ since the idea extends. Suppose that:

$0 < t_1 < t_2 < t_3 < 1$ and $\lambda' = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$. Then, $\lambda_1 W_{t_1} + \lambda_2 W_{t_2} + \lambda_3 W_{t_3} =$

$\lambda_1 W_{t_1} + \lambda_2 [W_{t_1} + (W_{t_2} - W_{t_1})] + \lambda_3 [W_{t_1} + (W_{t_2} - W_{t_1}) + (W_{t_3} - W_{t_2})] =$

$\lambda_3 [W_{t_1} + (W_{t_2} - W_{t_1}) + (W_{t_3} - W_{t_2})] =$

$(\lambda_1 + \lambda_2 + \lambda_3) W_{t_1} + (\lambda_2 + \lambda_3)(W_{t_2} - W_{t_1}) + \lambda_3 (W_{t_3} - W_{t_2})$ is normal since $W_{t_1} - W_0, W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}$ are independent random variables by (iii) and $W_0 = 0$ by (i). This idea extends to any k, and thus $\{W_t : t \in I\}$ is a Gaussian process. Also, if $0 \leq s \leq t \leq 1$, then $cov(W_s, W_t) = cov(W_s, W_s + (W_t - W_s)) = cov(W_s, W_s) + cov(W_s, W_t - W_s) = varW_s + 0.$

Since $W_s, W_t - W_s$ are independent random variables. Hence, $cov(W_s, W_t) = s\sigma^2$ by (i) and (ii) and thus $\{W_t : t \in I\}$ is a Wiener process.

Conversely, assume that $\{W_t : t \in I\}$ is a Wiener process. Since it is Gaussian and

$cov(W_0, W_0) = 0 \cdot \sigma^2 = 0$, that is,. $E(W_0^2) = 0$, it follows that $W_0 = 0$ a.s. and (i) is satisfied. Moreover, since $\{W_t: t \in I\}$ is a Gaussian process it follows that $W_t - W_s$ is normally distributed with mean zero. If $0 \leq s \leq t \leq 1$, then $var(W_t - W_s) = varW_t + varW_s - 2cov(W_s, W_t) =$

$t\sigma^2 + s\sigma^2 - 2(s \wedge t)\sigma^2 = \sigma^2(s + t - 2s) = (t - s)\sigma^2$. Hence $W_t - W_s$, and thus (ii) is fulfilled.

Next, if $0 \leq t_1 < t_2 < t_3 \cdots < t_k \leq 1$, $W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, \cdots, W_{t_k} - W_{t_{k-1}}$, are independent if and only if they are uncorrelated since each random variable is normal. Note that if $i < j$, $cov(W_{t_i} - W_{t_{i-1}}, W_{t_j} - W_{t_{j-1}}) = cov(W_{t_i} - W_{t_{i-1}}, W_{t_j} - W_{t_{j-1}}) = cov(W_{t_i}, W_{t_j}) -$ $cov(W_{t_{i-1}}, W_{t_j}) - cov(W_{t_i}, W_{t_{j-1}}) + cov(W_{t_{i-1}}, W_{t_{j-1}}) = t_i\sigma^2 - t_{i-1}\sigma^2 - t_i\sigma^2 + t_{i-1}\sigma^2 = 0$ Hence (iii) is satisfied. ■

Remark 3.1 Condition (ii) in Lemma 3.2 is called stationary. Condition (iii) is referred to as independent increments.

A function space Central Limit Theorem is used to show that the normalized numerator of $\hat{\alpha}_n - 1$ converges in distribution. The function space is the set of all right continuous maps defined on *I* = [0, 1] whose left – hand limits exist. This set is denoted *D(I).* A suitable topology for *D(I)* has been defined by A.V. Skorohod (for example, see [1], p.112). This space is separable. Moreover, there exists a complete metric ρ that induces the Skorohod topology (see [1], p.112). Further, if $f_n, f, n \geq 1$, are continuous elements in *D(I)*, then $f_n \overset{\rho}{\to} f$ if and only if $f_n \to f$ uniformly on I.

Let $\mathcal{B}$ denote the Borel σ-field on *D(I)*, and assume $V_n, V : (\Omega, \Im, P) \to (D(I), \mathcal{B})$, $n \geq 1$, are random elements (measurable functions). Define $V_n \overset{D}{\to} V$ if and only if for each bounded, continuous map $h : (D(I), \rho) \to \mathcal{R}$, $E(h(V_n)) \to E(h(V))$. The following fundamental result, known as the Continuous Mapping Theorem, is needed ([1], Theorem 5.1).

Theorem 3.1 Assume that $V_n \overset{D}{\to} V$ in *D(I)*, $h: (D(I), \mathcal{B}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable and $P^V\{f \in D(I): h$ is discontinuous at f}=0. Then $h(V_n) \overset{D}{\to} h(V)$.

M. Donsker proved the following Central Limit Theorem for a sequence of random elements in $(D(I), \mathcal{B})$. (See [1], Theorem 16.1).

Theorem 3.2 Let $\{\epsilon_n\}$ be a sequence IID, mean zero, with variance $0 < \sigma^2 < \infty$. Denote $X_n(t) \coloneqq \frac{1}{\sqrt{n}\sigma} \sum_{i=1}^{[\![nt]\!]} \epsilon_i$ where $t \in I$ and $[\![nt]\!]$ is the greatest integer which is less than or equal to

nt. Then $X_n \xrightarrow{D} W$ in $(D(I), \mathcal{B})$, where $\{W(t), t \in I\}$ is a Wiener process on I.

Quite often one wants to prove that a sequence $\{Z_n\}$ of random variables converges in distribution. Assume there exists a sequence $V_n \xrightarrow{D} V$ in *D(I)*, and a function h : (*D(I)*,*P*)→ ℝ that is continuous almost surely $[P^V]$. If $Z_n = h(V_n)$, then by Theorem 3.1, $Z_n \xrightarrow{D} h(V)$. In our use below, $V_n = X_n$ and $V = W$, where W denotes a Wiener process on I.

Lemma 3.3 Assume that the time series $\{Y_t\}_{t \geq 0}$ obeys $Y_t = Y_{t-1} + \epsilon_t, t \geq 1,$ and $\{\epsilon_t\}$ are IID, $E(\epsilon_1) = 0$, and $var\epsilon_1 = \sigma^2 < \infty$. Then:

$\frac{1}{n^2\sigma^2}\sum_{t=1}^{n} Y_{t-1}^2 \xrightarrow{D} \int_I W^2(t)dt$ as $n \to \infty$.

Proof: Define $X_n(t) \coloneqq \frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{[\![nt]\!]} \epsilon_i, 0 \leq t \leq 1$. According to Theorem 3.2, $X_n \xrightarrow{D} W$ in *D*([0,1]), where *D*([0,1]) is the set of all right-continuous maps on I with left-hand limits and W is a Wiener process on I. Define *h* : *D*([0,1]) → ℝ by $h(f) = \int_I f dx$.Then *h* is continuous almost surely $[P^W]$ on *D*([0,1]), and thus by Theorem 3.1, $h(X_n^2) \xrightarrow{D} h(W^2)$. Moreover,

$$h(X_n^2) = \int_I X_n^2(t)dt = \frac{\epsilon_1^2}{n^2\sigma^2} + \frac{(\epsilon_1+\epsilon_2)^2}{n^2\sigma^2} + \cdots + \frac{(\epsilon_1+\cdots+\epsilon_{n-1})^2}{n^2\sigma^2} =$$

$$\frac{1}{n^2\sigma^2}\sum_{t=1}^{n}\left(\sum_{i=1}^{t-1} \epsilon_i\right)^2 \xrightarrow{D} \int_I W^2(t)dt.$$

As shown in Lemma 3.1, $Y_{t-1} = Y_0 + \sum_{i=1}^{t-1} \epsilon_i$, and thus,

$\frac{1}{n^2\sigma^2}\sum_{t=1}^{n} Y_{t-1}^2 = \frac{n}{n^2\sigma^2} Y_0^2 + \frac{2Y_0}{n^2\sigma^2}\sum_{t=1}^{n}\sum_{i=1}^{t-1}\epsilon_i + \frac{1}{n^2\sigma^2}\sum_{t=1}^{n}(\sum_{i=1}^{t-1}\epsilon_i)^2$. Observe that

$\int_I X_n(t)dt = \frac{\epsilon_1}{n\sqrt{n}\sigma} + \frac{(\epsilon_1+\epsilon_2)}{n\sqrt{n}\sigma} + \cdots + \frac{(\epsilon_1+\cdots+\epsilon_{n-1})}{n\sqrt{n}\sigma} \overset{D}{\to} \int_I W(t)dt$ since $X_n \overset{D}{\to} W$ and $h(f) = \int_I f dt$

is continuous almost surely[P^W]. That is, $\frac{1}{n\sqrt{n}\sigma}\sum_{t=1}^{n}\sum_{i=1}^{t-1}\epsilon_i \overset{D}{\to} \int_I W(t)dt$ as $n \to \infty$. Hence,

$\frac{1}{n^2\sigma^2}\sum_{t=1}^{n}\sum_{i=1}^{t-1}\epsilon_i \overset{D}{\to} 0$, and thus it follows that $\frac{1}{n^2\sigma^2}\sum_{t=1}^{n} Y_{t-1}^2 \overset{D}{\to} \int_I W^2(t)dt$ as $n \to \infty$. ■

Theorem 3.3 Assume that $\{Y_t : t \geq 0\}$ is a time series that obeys model (3.1). Then

$n(\hat{\alpha}_n - 1) \overset{D}{\to} \frac{Z^2-1}{2\int_I W^2(t)dt}$, where $\hat{\alpha}_n$ is the least squares estimator of α as defined in (3.2).

Since $\alpha = 1$, according to (3.1), $\hat{\alpha}_n - 1 = \frac{\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2}$, and thus

$n(\hat{\alpha}_n - 1) = \frac{\frac{1}{n\sigma^2}\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\frac{1}{n^2\sigma^2}\sum_{t=1}^{n} Y_{t-1}^2} \overset{D}{\to} \frac{\frac{1}{2}Z^2-\frac{1}{2}}{\int_I W^2(t)dt} = \frac{Z^2-1}{2\int_I W^2(t)dt}$ by Lemma 3.1 and 3.3. ■

Remark 3.2 The above shows that $\{n(\hat{\alpha}_n - 1)\}$ converges in distribution. However, the distribution of $\frac{Z^2-1}{2\int_I W^2(t)dt}$ is not known (like normal or chi-square), and thus for the needed sample size must be approximated using simulation of $n(\hat{\alpha}_n - 1)$, before testing $H_0: \alpha = 1$.

AR (1)-CASE III: $|\alpha| > 1$

Consider the time series $\{Y_t\}_{t\geq 0}$ defined on $(\Omega, \mathfrak{J}, P)$ which obeys the model:

$$Y_t = \alpha Y_{t-1} + \epsilon_t \quad , \ t \geq 0, \tag{4.1}$$

where $\{\epsilon_t\}$ are IID random variables with $E(\epsilon_1) = 0$, $0 < var\epsilon_1 < \infty$, and $|\alpha| > 1$. Recall from (1.1) that the LSE $\hat{\alpha}_n$ of α satisfies $\hat{\alpha}_n - \alpha = \frac{\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2}$. Utilizing order in probability techniques, it is shown in Theorem 4.1 that $\hat{\alpha}_n - \alpha = O_P(|\alpha|^{-n})$.

Recall that if $\{a_n\}$ is a sequence of real numbers, then $a_n = O(1)$ means that there exists M such that $|a_n| \leq M \ \forall \ n \geq 1$, that is, the sequence is bounded. Further $a_n = o(1)$ means that $a_n \to 0$ as $n \to \infty$. Note that $a_n = o(1)$ implies $a_n = O(1)$. More generally, if $b_n > 0, n \geq 1$, then $a_n = O(b_n)$ means that $\left|\frac{a_n}{b_n}\right| = O(1)$ and $a_n = o(b_n)$ means that $\frac{a_n}{b_n} \to 0$ as $\to \infty$.

Example 4.1 Let $a_n = 1 - \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right), n \geq 1$. Then $a_n = 1 - \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) = \frac{3}{n} - \frac{2}{n^2}$.
Hence $a_n = \frac{3}{n} + O\left(\frac{1}{n^2}\right)$ and $a_n = \frac{3}{n} + o\left(\frac{1}{n^r}\right)$, $0 < r < 2$. ■

Let $\{X_n\}$ be a sequence of random variables defined on $(\Omega, \mathfrak{J}, P)$ and $\{a_n\}$ a sequence of positive real numbers. Then $X_n = O_P(a_n)$ means that for each $\epsilon > 0$ there exists M_ϵ such that

$P\{\frac{|X_n|}{a_n} \geq M_\epsilon\} < \epsilon \ \forall\ n \geq 1$. Further, $X_n = o_P(a_n)$ means that $\frac{X_n}{a_n} \xrightarrow{P} 0$ as $n \to \infty$, and

$X_n = X + o_P(a_n)$ is defined by $X_n - X = o_P(a_n)$.

Lemma 4.1 Assume that $\{X_n\}, \{Y_n\}$ are sequences of random variables defined on $(\Omega, \Im, P)$ and let $\{a_n\}, \{b_n\}$ be sequences of positive real numbers. Then:

(i) $X_n = o_P(a_n)$ and $Y_n = O_P(b_n) \Rightarrow X_nY_n = o_P(a_nb_n)$

(ii) $X_n = o_P(a_n)$ and $Y_n = o_P(b_n) \Rightarrow X_n + Y_n = o_P(a_n \vee b_n)$

(iii) $X_n = o_P(a_n)$ and $r > 0 \Rightarrow |X_n|^r = o_P(a_n^r)$.

(iv) (i),(ii),and (iii) hold with o_P replaced by O_P.

(v) $X_n \xrightarrow{P} X \Rightarrow X_n = O_P(1)$.

Proof:

(i) Fix $\epsilon > 0$, note that for $M > 0$, $\left\{\frac{|X_nY_n|}{a_nb_n} > \epsilon\right\}$

$\subseteq \left\{\frac{|Y_n|}{b_n} \leq M, \frac{|X_n|}{a_n} > \frac{\epsilon}{M}\right\} \cup \left\{\frac{|Y_n|}{b_n} > M, \frac{|X_nY_n|}{a_nb_n} > \epsilon\right\} \subseteq \left\{\frac{|X_n|}{a_n} > \frac{\epsilon}{M}\right\} \cup \left\{\frac{|Y_n|}{b_n} > M\right\}$. Given $\delta > 0$ choose M such that $P\left\{\frac{|Y_n|}{b_n} > M\right\} < \frac{\delta}{2}$ for each $n \geq 1$. Since $\frac{|X_n|}{a_n} \xrightarrow{P} 0$, $P\left\{\frac{|X_n|}{a_n} > \frac{\epsilon}{M}\right\} < \frac{\delta}{2}$ for n sufficiently large. Hence, $P\left\{\frac{|X_nY_n|}{a_nb_n} > \epsilon\right\} \leq P\left\{\frac{|X_n|}{a_n} > \frac{\epsilon}{M}\right\} + P\left\{\frac{|Y_n|}{b_n} > M\right\} < \delta$ for n sufficiently large, and thus $X_nY_n = o_P(a_nb_n)$.

(ii) Given $\epsilon > 0$, $\left\{\frac{|X_n+Y_n|}{a_n \vee b_n} > \epsilon\right\} \subseteq \left\{\frac{|X_n|}{a_n \vee b_n} > \frac{\epsilon}{2}\right\} \cup \left\{\frac{|Y_n|}{a_n \vee b_n} > \frac{\epsilon}{2}\right\} \subseteq \left\{\frac{|X_n|}{a_n} > \frac{\epsilon}{2}\right\} \cup \left\{\frac{|Y_n|}{b_n} > \frac{\epsilon}{2}\right\}$. and thus $P\left\{\frac{|X_n+Y_n|}{a_n \vee b_n} > \epsilon\right\} \leq P\left\{\frac{|X_n|}{a_n} > \frac{\epsilon}{2}\right\} + P\left\{\frac{|Y_n|}{b_n} > \frac{\epsilon}{2}\right\} \to 0$ as $n \to \infty$. Hence $X_n + Y_n = o_P(a_n \vee b_n)$.

(iii) Given $\epsilon > 0$, $\left\{\frac{|X_n|^r}{{a_n}^r} > \epsilon\right\} = \left\{\frac{|X_n|}{a_n} > \epsilon^{\frac{1}{r}}\right\}$, $P\left\{\frac{|X_n|^r}{{a_n}^r} > \epsilon\right\} = P\left\{\frac{|X_n|}{a_n} > \epsilon^{\frac{1}{r}}\right\} \to 0$, and thus $|X_n|^r = o_P(a_n^r)$.

(iv) Assume that $X_n = O_P(a_n)$. Given $\epsilon > 0$ choose M such that $P\left\{\frac{|X_n|}{a_n} > M^{\frac{1}{r}}\right\} < \epsilon$ for all $n \geq 1$. Then $P\left\{\frac{|X_n|^r}{{a_n}^r} > M\right\} < \epsilon$ for all $n \geq 1$, hence, $|X_n|^r = O_P(a_n^r)$.

(v) Fix $\epsilon > 0$. Since$\{|X| \geq n\} \downarrow \varphi$, there exists M_ϵ such that, $\left\{|X| \geq \frac{M_\epsilon}{2}\right\} < \frac{\epsilon}{2}$. Moreover, $X_n \xrightarrow{P} X$ implies that there exists n_0 such that $P\left\{|X_n - X| \geq \frac{M_\epsilon}{2}\right\} < \frac{\epsilon}{2}$ for $n \geq n_0$. Note that, $\{|X_n| > M_\epsilon\} \subseteq \left\{|X_n - X| > \frac{M_\epsilon}{2}\right\} \cup \left\{|X| > \frac{M_\epsilon}{2}\right\}$, and thus $P\{|X_n| > M_\epsilon\} \leq P\left\{|X_n - X| > \frac{M_\epsilon}{2}\right\} + P\left\{|X| > \frac{M_\epsilon}{2}\right\} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ for $n \geq n_0$. One can choose $N_\epsilon > M_\epsilon$ such that for all $n \geq 1$, $P\{|X_n| > N_\epsilon\} < \epsilon$, and thus $X_n = O_P(1)$. ■

Lemma4.2 Let $\{X_n\}$, $\{Y_n\}$ be sequences of random variables defined on $(\Omega, \Im, P)$ and let $\{a_n\}$

be a sequence of positive real numbers. Then:

(i) $E(X_n^2) = O(a_n^2)$ implies $X_n = O_P(a_n)$

(ii) $E(|X_n|) = O(a_n)$ implies $X_n = O_P(a_n)$. Also (i) implies (ii)

(iii) $X_n = O_P(a_n)$ and $a_n \to 0$ implies $X_n = o_P(1)$.

Proof:

(i) Choose K such that $\frac{E(X_n^2)}{a_n^2} \leq K$. Given $\epsilon > 0$, using Markov's inequality,

$P\left\{\frac{|X_n|}{a_n} \geq M\right\} \leq \frac{E(X_n)^2}{a_n^2 M^2} \leq \frac{K}{M^2}$. Let $M = \sqrt{\frac{K}{\epsilon}}$; then $P\left\{\frac{|X_n|}{a_n} \geq M\right\} \leq \frac{K}{\frac{K}{\epsilon}} = \epsilon$ for all $n \geq 1$.

Hence $X_n = O_P(a_n)$.

(ii) Choose K such that $\frac{E|X_n|}{a_n} \leq K$ for all $n \geq 1$. Then

$P\left\{\frac{|X_n|}{a_n} \geq M\right\} \leq \frac{E|X_n|}{a_n M} \leq \frac{K}{M} \leq \epsilon$ when $M = \frac{K}{\epsilon}$, $n \geq 1$. Hence $X_n = O_P(a_n)$ according to Liapounov's inequality, $E|X_n| \leq (E|X_n|^2)^{\frac{1}{2}}$. Then $\frac{E|X_n|}{a_n} \leq \left[\frac{E(X_n^2)}{a_n^2}\right]^{\frac{1}{2}}$, and thus $E(X_n^2) = O(a_n^2)$ implies that $E|X_n| = O(a_n)$.

(iii) Given $\epsilon > 0$, it must be shown that $P\{|X_n| \geq \epsilon\} \to 0$ as $n \to \infty$. Fix $\delta > 0$. Since $X_n = O_P(a_n)$, there exists $M_\delta > 0$ such that $P\left\{\frac{|X_n|}{a_n} \geq M_\delta\right\} < \delta$ for all $n \geq 1$.

However, $a_n \to 0$ and thus $M_\delta a_n \leq \epsilon$ for all $n \geq n_0$. Hence for all $n \geq n_0$,

$P\{|X_n| \geq \epsilon\} \leq P\{|X_n| \geq M_\delta a_n\} = P\left\{\frac{|X_n|}{a_n} \geq M_\delta\right\} < \delta$, and thus $X_n = o_P(1)$. ■

Lemma 4.3 Assume that $\{X_n\}$,$\{Y_n\}$ are sequences of random variables defined on $(\Omega, \mathfrak{I}, P)$. Then:

(i) $X_n \overset{P}{\to} X$, $Y_n \overset{P}{\to} Y$ implies $X_n \pm Y_n \overset{P}{\to} X \pm Y$

(ii) $X_n \overset{P}{\to} X$, $Y_n \overset{P}{\to} Y$ implies $X_n Y_n \overset{P}{\to} XY$

(iii) $X_n \overset{P}{\to} X$, $Y_n \overset{P}{\to} Y$ implies $\frac{X_n}{Y_n} \overset{P}{\to} \frac{X}{Y}$ provided $P\{Y = 0\} = 0$

(iv) $X_n - Y_n \overset{P}{\to} 0$ implies $X_n \overset{D}{\to} X$ if and only if $Y_n \overset{D}{\to} X$.

(v) $(X_n, Y_n) \overset{D}{\to} (X, Y) \Rightarrow X_n \pm Y_n \overset{D}{\to} X \pm Y, X_n Y_n \overset{D}{\to} XY$, and $\frac{X_n}{Y_n} \overset{D}{\to} \frac{X}{Y}$ provided $P\{Y = 0\} = 0$.

(vi) $X_n \overset{D}{\to} X$, $Y_n \overset{D}{\to} a \in \mathbb{R}$ implies $(X_n, Y_n) \overset{D}{\to} (X, a)$.

(vii) $Y_n = X_n + O_P(a_n) \overset{D}{\to} Y$ if and only if $X_n \overset{D}{\to} Y$, provided $a_n \to 0$.

Given the model $Y_t = \alpha Y_{t-1} + \epsilon_t$, where $|\alpha| > 1$ and $\{\epsilon_t\}$ are IID, $E(\epsilon_1) = 0$, and $0 < var\epsilon_1 = \sigma^2 < \infty$, $t \geq 1$. Let Y_0 be the initial random variable. Recall that $\hat{\alpha}_n - \alpha = \frac{\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\sum_{t=1}^{n} Y_{t-1}^2}$, and thus:

$$\alpha^n(\hat{\alpha}_n - \alpha) = \frac{\alpha^{-n}\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2}. \quad (4.2)$$

Iterating the model $Y_t = \alpha Y_{t-1} + \epsilon_t$ gives $Y_t = \alpha^t Y_0 + \sum_{j=1}^{t} \alpha^{t-j}\epsilon_j$. Let $X_t := Y_0 + \sum_{j=1}^{t} \alpha^{-j}\epsilon_j$.

Since $var\left(\sum_{j=1}^{\infty}\alpha^{-j}\epsilon_j\right) = \sum_{j=1}^{\infty}\alpha^{-2j}\sigma^2 < \infty$, it follows from one of Kolmogorov's theorems that $X_t \xrightarrow{a.s.} Y_0 + \sum_{j=1}^{t}\alpha^{-j}\epsilon_j =: X$. Note that $Y_t = \alpha^t X_t$ and thus $\sum_{t=1}^{n} Y_{t-1}^2 = \sum_{t=1}^{n}\alpha^{2(t-1)}X_{t-1}^2$.

Then $\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 = \sum_{t=1}^{n}\alpha^{2(t-n)-2}X_{t-1}^2 = \sum_{t=1}^{n}\alpha^{2(t-n)-2}\left(X + (X_{t-1} - X)\right)^2 =$ $\sum_{t=1}^{n}\alpha^{2(t-n)-2}X^2 + 2\sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_{t-1} - X) + \sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_{t-1} - X)^2$. Let $j = n - t$, then $\sum_{t=1}^{n}\alpha^{2(t-n)-2}X^2 = X^2\sum_{j=0}^{n-1}\alpha^{-2(j+1)}$ Hence:

$$\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 = X^2\sum_{j=0}^{n-1}\alpha^{-2(j+1)} + 2\sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_{t-1} - X) + \sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_{t-1} - X)^2 . \tag{4.3}$$

Lemma 4.4 Suppose that the sequence $\{Y_n\}$ obeys the model and assumptions listed in (4.1). Then:

(i) $X^2\sum_{j=0}^{n}\alpha^{-2(j+1)} = (\alpha^2 - 1)^{-1}X^2 + O_p(|\alpha|^{-2n})$

(ii) $E(X_{t-1} - X)^2 = O(|\alpha|^{-2t})$

(iii) $\sum_{t=1}^{n}\alpha^{2(t-n)-2}(X_{t-1} - X)^2 = O_p(n|\alpha|^{-2n})$

(iv) $\sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_{t-1} - X) = O_p(|\alpha|^{-n})$

(v) $\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 = (\alpha^2-1)^{-1}X^2 + O_p(|\alpha|^{-n})$

(vi) $\alpha^{-2n}Y_n \to (\alpha^2-1)^{-1}X^2$ almost surely.

Proof:

(i) Note that $\sum_{j=0}^{n}\alpha^{-2(j+1)} = \alpha^{-2}\sum_{j=0}^{n}\alpha^{-2j} = \alpha^{-2}\frac{1-\alpha^{-2(n+1)}}{1-\alpha^{-2}} = \frac{1-\alpha^{-2(n+1)}}{\alpha^2-1} = \frac{1}{\alpha^2-1} + O(|\alpha|^{-2n})$ Since $X^2 = O_P(1)$, it follows from Lemma 4.1 (iv) that $X^2\sum_{j=0}^{n}\alpha^{-2(j+1)} = (\alpha^2-1)^{-1}X^2 + X^2 O(|\alpha|^{-2n}) = (\alpha^2-1)^{-1}X^2 + O_p(|\alpha|^{-2n})$.

(ii) Since $X_{t-1} - X = \sum_{j=t}^{\infty}\alpha^{-j}\epsilon_j$, it follows that $E(X_{t-1}-X)^2 = var(X_{t-1}-X) = \sum_{j=t}^{\infty}\alpha^{-2j}\sigma^2 = \alpha^{-2t}\sigma^2\sum_{j=0}^{\infty}\alpha^{-2j} = \alpha^{-2t}\frac{\sigma^2}{1-\alpha^{-2}} = O(|\alpha|^{-2t})$.

(iii) According to (ii), $E[\sum_{t=1}^{n}\alpha^{2(t-n)-2}(X_{t-1}-X)^2] = \sum_{t=1}^{n}\alpha^{2(t-n)-2}E(X_{t-1}-X)^2 \le M\sum_{t=1}^{n}\alpha^{2(t-n)-2}\alpha^{-2t} = \text{M}\alpha^{-2n-2}\sum_{t=1}^{n}1 = O(n|\alpha|^{-2n})$. It follows from Lemma 4.2 (ii) that $\sum_{t=1}^{n}\alpha^{2(t-n)-2}(X_{t-1}-X)^2] = O_P\,(n|\alpha|^{-2n})$.

(iv) Applying Cauchy's inequality and (ii) above,

$E\left|\sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_t-X)\right| \le \sum_{t=1}^{n}\alpha^{2(t-n)-2}(E|X|^2E|X_t-X|^2)^{\frac{1}{2}} \le (E|X|^2)^{\frac{1}{2}}\sum_{t=1}^{n}\alpha^{2(t-n)-2}(E|X_t-X|^2)^{\frac{1}{2}} \le M\sum_{t=1}^{n}\alpha^{2(t-n)-2}|\alpha|^{-t} = \alpha^{-2}M|\alpha|^{-n}\sum_{t=1}^{n}|\alpha|^{-(n-t)} = |\alpha|^{-n-2}M\left(\frac{1-|\alpha|^{-n}}{1-|\alpha|^{-1}}\right) = O(|\alpha|^{-n})$.

It follows from Lemma 4.2 (ii) that $\sum_{t=1}^{n}\alpha^{2(t-n)-2}X(X_t-X) = O_P(|\alpha|^{-n})$.

(v) It follows from (4.3), Lemma 4.4 (i), and parts (i),(iii), and (iv) above that

$\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 = (\alpha^2-1)^{-1}X^2 + O_P(|\alpha|^{-n})$.

(vi) Given $\delta > 0$, employing (4.3), (i), (v), and the expectations calculated in (iii) and (iv), $\sum_{n=1}^{\infty} P\{|\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 - (\alpha^2-1)^{-1}X^2| > \delta\} \leq M\sum_{n=1}^{\infty}|\alpha|^{-n} < \infty$, for some $M > 0$. It follows from the Borel-Cantelli Lemma that $\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 \to (\alpha^2-1)^{-1}X^2$ almost surely. ■

Next, consider the numerator of $\alpha^{-n}\sum_{t=1}^{n} Y_{t-1}\epsilon_t$ from (4.2). Note that:

$$\alpha^{-n}\sum_{t=1}^{n} Y_{t-1}\epsilon_t = \alpha^{-n}\sum_{t=1}^{n}\alpha^{t-1}X_{t-1}\epsilon_t = \alpha^{-n}\sum_{t=1}^{n}\alpha^{t-1}\big(X+(X_{t-1}-X)\big)\epsilon_t =$$

$$X\sum_{t=1}^{n}\alpha^{-(n-t+1)}\epsilon_t + \sum_{t=1}^{n}\alpha^{-(n-t+1)}(X_{t-1}-X)\epsilon_t. \qquad (4.4)$$

Lemma 4.5 Given that the sequence $\{Y_n\}$ satisfies the model and assumption listed in (4.1) then:

(i) $X_{t-1} - X = O_P(|\alpha|^{-t})$

(ii) $\sum_{t=1}^{n}\alpha^{-(n-t+1)}(X_{t-1}-X)\epsilon_t = O_P(n|\alpha|^{-n})$

(iii) $\sum_{t=1}^{n}\alpha^{-(n-t+1)}(X_{t-1}-X)\epsilon_t \to 0$ almost surely.

Proof:

(i) Recall $X_{t-1} - X = \sum_{j=t}^{\infty}\alpha^{-j}\epsilon_j$. Then by Lemma 4.2 (i) and Lemma 4.4 (ii) $X_{t-1} - X = O_P(|\alpha|^{-t})$.

(ii) According to Lemma 4.4 (ii), $E(X_{t-1} - X)^2 = O_P(|\alpha|^{-2t})$. Hence

$E\left|\sum_{t=1}^{n} \alpha^{-(n-t+1)}(X_{t-1} - X)\epsilon_t\right| \leq |\alpha|^{-1} \sum_{t=1}^{n} |\alpha|^{-(n-t)} \left(E(X_{t-1} - X)^2 E(\epsilon_t^2)\right)^{\frac{1}{2}} \leq$

$|\alpha|^{-1}\sigma M \sum_{t=1}^{n} |\alpha|^{-(n-t)}|\alpha|^{-t} = O(n|\alpha|^{-n})$. Then by Lemma 4.2 (ii),

$\sum_{t=1}^{n} \alpha^{-(n-t+1)}(X_{t-1} - X)\epsilon_t = O_P(n|\alpha|^{-n})$.

(iii) As shown in the proof of (ii) above, $\sum_{n=1}^{\infty} E\left|\sum_{t=1}^{n} \alpha^{-(n-t+1)}(X_{t-1} - X)\epsilon_t\right| \leq$ $\sum_{n=1}^{\infty} Mn|\alpha|^{-n}$, and the latter series converges by the ratio test. Then by the Borel-Cantelli Lemma, $\sum_{t=1}^{n} \alpha^{-(n-t+1)}(X_{t-1} - X)\epsilon_t \to 0$ almost surely. ■

Theorem 4.1 Let $\{Y_n\}$ be a sequence which satisfies the model and assumption listed in 4.1.

Then $\hat{\alpha}_n - \alpha = O_P(|\alpha|^{-n})$.

Proof: According to (4.2), $\alpha^n(\hat{\alpha}_n - \alpha) = \frac{\alpha^{-n}\sum_{t=1}^{n} Y_{t-1}\epsilon_t}{\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2}$, and by Lemma 4.4 (vi), $\alpha^{-2n}\sum_{t=1}^{n} Y_{t-1}^2 \to (\alpha^2 - 1)^{-1}X^2$ almost surely. By (4.4) and Lemma 4.5 (ii), $\alpha^{-n}\sum_{t=1}^{n} Y_{t-1}\,\epsilon_t = X\sum_{t=1}^{n} \alpha^{-(n-t+1)}\epsilon_t + O_P(1)$. Moreover, $X\sum_{t=1}^{n} \alpha^{-(n-t+1)}\epsilon_t =$ $O_P(1)\,O_P(1) = O_P(1)$, and thus $\alpha^{-n}\sum_{t=1}^{n} Y_{t-1}\,\epsilon_t = O_P(1)$. Hence $\alpha^n(\hat{\alpha}_n - \alpha) = O_P(1)$, and thus $\hat{\alpha}_n - \alpha = O_P(|\alpha|^{-n})$. ■

MA (1)

Consider the first order moving average model: $Y_t = \epsilon_t + \theta_0 \epsilon_{t-1}, t \geq 1$, where ϵ_t are IID, $t \geq 1, E(\epsilon_t) = 0$ and $var\epsilon_t = \sigma^2 < \infty$. The $Y_t{}'$s, $1 \leq t \leq n$ are observable, and the goal is to define an estimator $\hat{\theta}_n$ in terms of $Y_1, Y_2, \cdots, Y_n$ such that $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N(0, \Delta^2)$, for some $\Delta > 0$. In particular, this would imply that $\hat{\theta}_n \xrightarrow{P} \theta_0$; that is, $\{\hat{\theta}_n\}$ is a consistent estimator of the unknown parameter θ_0. It is assumed that $\theta_0 \in \Theta = [-r, r] \ni 0 < \mathrm{r} < 1$. This is needed for convergence properties.

One might proceed as follows. Note that $Y_t = \epsilon_t + \theta_0 \epsilon_{t-1}$, define $f(x, \theta) = \theta x$, where $|\theta| \leq r < 1$, and consider the least squares sum $Q_n(\theta) = \sum_{t=1}^{n} (Y_t - \theta\epsilon_{t-1})^2$. Since Q_n is continuous in θ, there exists $\hat{\theta}_n \in [-r, r]$ such that $Q_n(\hat{\theta}_n) = \min_{|\theta| \leq r} Q_n(\theta)$. It might be tempting to use $\hat{\theta}_n$ as an estimator of θ_0. The problem is that one needs to know $\epsilon_0, \epsilon_1, \cdots, \epsilon_{n-1}$in order to find $\hat{\theta}_n$. Since $\epsilon_0, \epsilon_1, \cdots, \epsilon_{n-1}$ are not observable, it follows that $\hat{\theta}_n$ is not a valid estimator.

Note that $\epsilon_t = Y_t - \theta_0 \epsilon_{t-1}, t \geq 1$. Let $\delta_0 = \epsilon_0$ be an imposed initial random variable. Define $\epsilon_1(\theta) = Y_1 - \theta\delta_0$ and $\epsilon_t(\theta) = Y_t - \theta\epsilon_{t-1}(\theta), t \geq 2$. Iterating, $\epsilon_2(\theta) = Y_2 - \theta\epsilon_1(\theta) =$

$Y_2 - \theta(Y_1 - \theta\delta_0) = Y_2 - \theta Y_1 + \theta^2\delta_0, \epsilon_3(\theta) = Y_3 - \theta\epsilon_2(\theta) = Y_3 - \theta(Y_2 - \theta Y_1 + \theta^2\delta_0) =$

$Y_3 - \theta Y_2 + \theta^2 Y_1 - \theta^3\delta_0$, etc. Hence $\epsilon_t(\theta) = \sum_{k=0}^{t-1} (-\theta)^k Y_{t-k} + (-\theta)^t \delta_0,\ t \geq 1$, is computable for various values of θ. By Taylor's expansion, $\epsilon_t = \epsilon_t(\theta_0) = \quad \epsilon_t(\theta) + \frac{\partial \epsilon_t(\theta)}{\partial \theta}(\theta_0 - \theta) + R_t,\ t \geq 1$. Denote $W_t(\theta) = -\frac{\partial \epsilon_t(\theta)}{\partial \theta}$ and thus $\epsilon_t(\theta) = W_t(\theta)(\theta_0 - \theta) + \epsilon_t - R_t$.

Note that $W_t(\theta) = \sum_{k=0}^{t-1} k(-\theta)^{k-1}\, Y_{t-k} + t(-\theta)^{t-1}\delta_0$. Let $\bar{\theta}_n$ denote an initial estimator of θ_0, where $\left|\bar{\theta}_n(w)\right| \leq r$ for each $w \in \Omega$. Then $\epsilon_t(\bar{\theta}_n) = W_t(\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + \epsilon_t - R_t$. An estimator $\hat{\delta}_n$ of $\theta_0 - \bar{\theta}_n$ is obtained by ignoring R_t and using the least squares estimator, $\hat{\delta}_n = \sum_{t=1}^{n} \frac{W_t(\bar{\theta}_n)\epsilon_t(\bar{\theta}_n)}{W_t^2(\bar{\theta}_n)}$. The improved Gauss-Newton estimator $\hat{\theta}_n$ of θ_0 is $\hat{\theta}_n = \hat{\delta}_n + \bar{\theta}_n$.

Since $\epsilon_t(\bar{\theta}_n) = W_t(\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + \epsilon_t - R_t$, $\hat{\delta}_n = \frac{\sum_{t=1}^{n} W_t(\bar{\theta}_n)[W_t(\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + \epsilon_t - R_t]}{\sum_{t=1}^{n} W_t^2(\bar{\theta}_n)} = (\theta_0 - \bar{\theta}_n) + \frac{\sum_{t=1}^{n} W_t(\bar{\theta}_n)(\epsilon_t - R_t)}{\sum_{t=1}^{n} W_t^2(\bar{\theta}_n)}$ and thus $\hat{\theta}_n - \theta_0 = \hat{\delta}_n - (\theta_0 - \bar{\theta}_n) = \frac{\sum_{t=1}^{n} W_t(\bar{\theta}_n)(\epsilon_t - R_t)}{\sum_{t=1}^{n} W_t^2(\bar{\theta}_n)}$. Then $\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) = \frac{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} W_t(\bar{\theta}_n)(\epsilon_t - R_t)}{\frac{1}{n}\sum_{t=1}^{n} W_t^2(\bar{\theta}_n)}$. It remains to show that $\frac{1}{n}\sum_{t=1}^{n} W_t^2(\bar{\theta}_n) \xrightarrow{P} b$, a nonzero constant random variable, $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} W_t(\bar{\theta}_n)R_t \xrightarrow{P} 0$, and $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} W_t(\bar{\theta}_n)\epsilon_t \xrightarrow{D} N(0, p^2)$, for some $p > 0$. Then it follows that $\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{D} \frac{1}{b} N(0, p^2) = N\left(0, \frac{p^2}{b^2}\right) = N(0, \Delta^2)$, where $\Delta = \frac{p}{b}$. Showing the above results are not easy; let us first look at the general parameter estimation problem in nonlinear regression. This provides the general technique needed.

Consider the general regression model $Y_t = f(X_t, \theta_0) + \epsilon_t, t \geq 1$, where θ_0 is the unknown parameter, x_t and Y_t are observable random variables, $1 \leq t \leq n$. In general, $\theta_0 \in \mathbb{R}^p$;

However, in our case $\theta_0 \in \mathbb{R}$. Suppose that the parameter space Θ is a compact subset of $\mathbb{R}$ and $f(\cdot,\theta)\colon \Theta \to \mathbb{R}$ is continuous in θ. Let us assume that $\{\epsilon_t, t \geq 1\}$ are IID, mean zero and variance$\sigma^2 < \infty$. As before, suppose that $\{\bar{\theta}_n\}$is a sequence of initial estimators of θ_0. Assume that $f(\cdot,\theta)$has two continuous derivatives in θ. By Taylor's expansion, $f(X_t,\theta_0) = f(X_t,\bar{\theta}_n) + \frac{\partial f}{\partial \theta}(X_t,\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + \frac{1}{2}\frac{\partial^2}{\partial \theta^2}\left(X_t,\tilde{\theta}_n\right)(\theta_0 - \bar{\theta}_n)^2$, for some $\tilde{\theta}_n$ on line between $\theta_0$and$\bar{\theta}_n$. Here $\tilde{\theta}_n$ is a random variable. Then $Y_t = f(X_t,\theta_0) + \epsilon_t = f(X_t,\bar{\theta}_n) + \frac{\partial f}{\partial \theta}(X_t,\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + \frac{1}{2}\frac{\partial^2 f}{\partial \theta^2}\left(X_t,\tilde{\theta}_n\right)(\theta_0 - \bar{\theta}_n)^2 + e_t$. Define $F_n'(\theta) = \left(\frac{\partial f}{\partial \theta}(X_1,\theta), \frac{\partial f}{\partial \theta}(X_2,\theta), \cdots, \frac{\partial f}{\partial \theta}(X_t,\theta)\right)$, $Y_n' = (Y_1, Y_2, \cdots, Y_n)$, $\lambda_n' = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)$, and $f_n'(\theta) = \left(f(X_1,\theta), f(X_2,\theta), \cdots, f(X_t,\theta)\right)$. Then $Y_n - f_n(\bar{\theta}_n) \approx F_n(\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + \lambda_n$, and thus the least squares estimator of $\theta_0 - \bar{\theta}_n$ is $\hat{\delta}_n = \frac{F_n'(\bar{\theta}_n)\left(Y_n - f_n(\bar{\theta}_n)\right)}{F_n'(\bar{\theta}_n)F_n(\bar{\theta}_n)}$. The Gauss-Newton estimator of θ_0 is defined by $\hat{\theta}_n = \hat{\delta}_n + \bar{\theta}_n$.

Again, observe that $\hat{\delta}_n = \frac{F_n'(\bar{\theta}_n)\left(F_n(\bar{\theta}_n)(\theta_0 - \bar{\theta}_n) + R_n + \lambda_n\right)}{F_n'(\bar{\theta}_n)F_n(\bar{\theta}_n)} = (\theta_0 - \bar{\theta}_n) + \frac{F_n'(\bar{\theta}_n)(R_n + \lambda_n)}{F_n'(\bar{\theta}_n)F_n(\bar{\theta}_n)}$, where $R_n' = \frac{1}{2}\left(\frac{\partial^2 f}{\partial \theta^2}\left(X_1,\tilde{\theta}_n\right), \frac{\partial^2 f}{\partial \theta^2}\left(X_2,\tilde{\theta}_n\right), \cdots, \frac{\partial^2 f}{\partial \theta^2}\left(X_n,\tilde{\theta}_n\right)\right)(\theta_0 - \bar{\theta}_n)^2$. Hence $\hat{\theta}_n - \theta_0 = \hat{\delta}_n - (\theta_0 - \bar{\theta}_n) = \frac{F_n'(\bar{\theta}_n)(R_n + \lambda_n)}{F_n'(\bar{\theta}_n)F_n(\bar{\theta}_n)}$, and thus $\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) = \frac{\frac{1}{\sqrt{n}}F_n'(\bar{\theta}_n)(R_n + \lambda_n)}{\frac{1}{n}F_n'(\bar{\theta}_n)F_n(\bar{\theta}_n)}$. As before, it must be shown that $\frac{1}{n}F_n'(\bar{\theta}_n)F_n(\bar{\theta}_n) \xrightarrow{P} b$, where b is a nonzero real number, $\frac{1}{\sqrt{n}}F_n'(\bar{\theta}_n)R_n \xrightarrow{P} 0$, and

$\frac{1}{\sqrt{n}} F_n'(\bar{\theta}_n)\lambda_n \xrightarrow{D} N(0, p^2)$. Then it follows that $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N\left(0, \frac{p^2}{b^2}\right) = N(0, \Delta^2)$.

REFERENCES

[1] Billingsley, P. (1968), Convergence of Probability Measures. Wiley, New York.

[2] Brockwell, P.J. and Davis, R.A. (1991), Time Series: Theory and Methods, Springer Series in Statistics, 2nd edition, New York.

[3] Hasza, D.P. (1977), Estimation in Nonstationary Time Series, PhD. Dissertation, Statistics Department, Iowa State University.

[4] Fuller, W.A. (1996), An Introduction to Time Series, Wiley, 2nd edition, New York.