Convergence of averages of scaled functions of I(1) linear processes

Robert M. de Jong
Department of Economics
Michigan State University

May 30, 2000

Abstract

Econometricians typically make use of functional central limit theorems to prove results for I(1) processes. For example, to establish the limit distributions of unit root tests such as the Phillips-Perron and Dickey-Fuller tests, the functional central limit theorem plays a crucial role. In this paper, it is pointed out that for linear processes, minimal conditions that ensure that only a central limit theorem holds are sufficient for establishing limit distributions of such tests. This eliminates the need to impose the stronger functional central limit theorem conditions and implies convergence of Dickey-Fuller type unit root tests under minimal conditions.

1 Introduction

Unit root tests and the analysis of econometric procedures in the presence of unit roots is by now standard in econometrics. Hamilton (1994), Chapter 17, for example, provides an overview of procedures that are in use to deal with unit root processes. Typically, asymptotic distribution proofs for unit root tests rely on convergence results for statistics of the form

$$(n^{-1/2}x_n, n^{-1}\sum_{t=1}^{n}\Psi(n^{-1/2}x_t)'f(t/n))', \tag{1}$$

where x_t is a unit root process that can be written as

$$x_t = x_0 + \sum_{j=1}^{t}\varepsilon_j, \tag{2}$$

where x_0 is random or nonrandom, $\Psi(.)$ is a $(k \times 1)$ vector, and $f(.)$ is a $(k \times k)$ matrix that is assumed to be continuous for $\xi \in [0,1]$ with respect to the norm $|.|$, which, for a matrix A, is defined as $|A| = \max_{i,j}|A_{ij}|$. The ε_j are assumed to satisfy a weak dependence condition. Typically, we prove convergence in distribution of the statistic of Equation (1) by defining

$$W_n(\xi) = n^{-1/2}x_{[\xi n]}/\sigma \tag{3}$$

where $\sigma^2 = \lim_{n\to\infty} E(n^{-1/2}x_n)^2$, and by noting that from the functional central limit theorem (FCLT) it follows that $W_n(\xi) \Rightarrow W(\xi)$, where $W(.)$ denotes Wiener measure and "$\Rightarrow$" denotes weak convergence. An appeal to the continuous mapping theorem then completes the proof of convergence of the statistics of Equation (1). FCLTs typically require a set of assumptions that is at least as strict as the set of assumptions needed for a central limit theorem (CLT) to hold. This is because, in addition to the conditions needed for the CLT to hold, a "tightness" or "stochastic equicontinuity" condition has to be proven in order to strengthen the CLT to an FCLT. In this note, we point out that for *linear processes* ε_j, the convergence in distribution proof for the statistic given in Equation (1) can be completed under the conditions for a CLT for linear processes that Phillips and Solo (1992) call "a minimal result", rather than the stronger FCLT conditions. This is because our approach avoids the conditions of the FCLT that are implicitly imposed to verify the "tightness" or "stochastic equicontinuity" condition.

2 Main result

Throughout this paper, we assume that ε_j and x_t are scalars, while $\Psi(.)$ is possibly vector-valued, and $X_n \xrightarrow{d} X$ denotes convergence in distribution. The following assumption will be needed for the $\Psi(.)$ function:

Assumption 1 *$\Psi(.)$ is Lipschitz-continuous, i.e.*

$$|\Psi(x) - \Psi(y)| \leq L(x,y)|x-y|^\alpha \tag{4}$$

for some $\alpha \in (0,1]$ and all $x, y \in \mathbb{R} \times \mathbb{R}$, and $L(x,y)$ is continuous on $\mathbb{R} \times \mathbb{R}$.

To derive our first result, we need the following assumption.

Assumption 2

1. *For all $1 \leq a < b$ and some constant $C > 0$, $E(\sum_{j=a}^{b} \varepsilon_j)^2 \leq C(b-a)$.*
2. *For each $0 \leq \xi_1 \leq \xi_2 \leq \ldots \leq \xi_K$, $(X_n(\xi_1), X_n(\xi_2), \ldots, X_n(\xi_K))$ is asymptotically normally distributed, where $X_n(\xi_j) = n^{-1/2} \sum_{j=1}^{[\xi_j n]} \varepsilon_j$.*
3. *$C(\xi, \xi') = \lim_{n\to\infty} E(n^{-1/2} \sum_{j=1}^{[\xi n]} \varepsilon_j)(n^{-1/2} \sum_{j=1}^{[\xi' n]} \varepsilon_j)$ is well-defined and continuous for all $\xi, \xi' \in (0,1]^2$, and there exists a Gaussian process $X(.)$ with continuous sample paths that has $C(\xi, \xi')$ as its covariance kernel.*

Assumption 3 *For constants $C_1 > 0$, $C_2 > 0$, and $C_3 > 0$, $|L(x,y)| \leq C_1 + C_2|x| + C_3|y|$.*

The above assumptions are sufficient to prove the following result, that econometricians usually derive by an appeal to a FCLT:

Theorem 1 *Under Assumptions 1 and 2, and 3, we have*

$$(n^{-1/2}x_n, n^{-1}\sum_{t=1}^{n} \Psi(n^{-1/2}x_t)' f(t/n))' \xrightarrow{d} (X(1), \int_0^1 \Psi(X(\xi))' f(\xi) d\xi)'. \tag{5}$$

Interestingly, if Ψ is continuous but not Lipschitz-continuous as assumed in Assumption 3, the argument in this paper apparently is invalid. Assumption 3 severely limits the class of functions $\Psi(.)$ that can be considered. For example, among the functions $\Psi(x) = |x|^\beta$, values of β that exceed 2 cannot be considered. However, the above theorem is general enough to cover all statistics in Chapter 17 of Hamilton (1994) that Hamilton needs to analyze unit root tests; i.e., Dickey-Fuller type coefficient tests, t-tests and F-tests based on regressions that exclude both a constant and a trend, include an intercept yet exclude a trend, or include both an intercept and a trend. See Proposition 17.1 and 17.3 of Hamilton (1994), p. 486 and 505/506.

3 Linear processes

A linear processes ε_j is defined as

$$\varepsilon_j = \sum_{k=0}^{\infty} \phi_k \eta_{j-k}, \tag{6}$$

for some sequence of constants ϕ_k. The η_j are typically assumed to satisfy a weak dependence condition, such as independence or some form of mixing. Here, we will assume the following for the η_j:

Assumption 4 *η_j is a sequence of i.i.d. random variables with mean zero and variance $\gamma^2 \in (0, \infty)$.*

Finiteness of $E\varepsilon_j^2$ now implies $\sum_{k=0}^{\infty} \phi_k^2 < \infty$. For linear processes, an FCLT for $n^{-1/2} \sum_{j=1}^{[\xi n]} \varepsilon_j$ was established in Theorem 3.4 of Phillips and Solo (1992) under Assumption 4 and the condition $\sum_{k=0}^{\infty} k^{1/2} |\phi_k| < \infty$. However, Theorem 3.11 of Phillips and Solo (1992) establishes a CLT for $n^{-1/2} \sum_{j=1}^{n} \varepsilon_j$ under the condition $\sum_{k=0}^{\infty} \phi_k < \infty$. Clearly, if $|\phi_k| = O(k^{-1-\delta})$ for some $0 < \delta < 1/2$, the second assumption will hold, while the first will not necessarily hold. (Note that both results also require $\sum_{k=0}^{\infty} \phi_k^2 < \infty$). Also, the condition $\sum_{k=0}^{\infty} \phi_k < \infty$ is minimal. This is because under Assumption 4 and if $0 < \sum_{k=0}^{\infty} \phi_k < \infty$ and $\sum_{k=0}^{\infty} \phi_k^2 < \infty$,

$$n^{-1/2} \sum_{j=1}^{n} \varepsilon_j \xrightarrow{d} N(0, \gamma^2 (\sum_{k=0}^{\infty} \phi_k)^2). \tag{7}$$

See Hall and Heyde (1980) and Phillips and Solo (1992, Theorem 3.11) for this result. Davidson (2000) analyzes the issues involved in establishing the FCLT for linear processes, but fails to establish an FCLT for linear process under the conditions $0 < |\sum_{k=0}^{\infty} \phi_k| < \infty$ and $\sum_{k=0}^{\infty} \phi_k^2 < \infty$. The theorem below shows that under this set of minimal conditions, we can prove convergence in distribution of the statistic of Equation (1) as well:

Theorem 2 *Under Assumption 1, 2 and 3, if $0 < \sum_{k=0}^{\infty} \phi_k < \infty$ and $\sum_{k=0}^{\infty} \phi_k^2 < \infty$, then*

$$(n^{-1/2} x_n, n^{-1} \sum_{t=1}^{n} \Psi(n^{-1/2} x_t)' f(t/n))' \xrightarrow{d} (X(1), \int_0^1 \Psi(\sigma W(\xi))' f(\xi) d\xi)', \tag{8}$$

where $\sigma^2 = \gamma^2 \sum_{k=0}^{\infty} \phi_k$.

Note that under the conditions of Theorem 2, it will follow that a weak law of large numbers holds for $n^{-1} \sum_{t=1}^{n} \varepsilon_t^2$. Therefore, limit distributions of Dickey-Fuller type tests can be obtained under the conditions of Theorem 2. Clearly the results of this paper by themselves are not sufficient for analyzing Phillips-Perron type tests, because the derivation of the limit distribution of such tests also requires a consistency proof for heteroscedasticity and autocorrelation consistent covariance matrices.

Mathematical proofs

Proof of Theorem 1:

Below, without loss of generality, we will act as if $x_0 = 0$. Note that T_n converges in distribution to T if

$$\lim_{K\to\infty} \limsup_{n\to\infty} E|T_n - \tilde{T}_{nK}| = 0, \tag{9}$$

$\tilde{T}_{nK} \xrightarrow{d} \tilde{T}_K$ (as $n \to \infty$) for any finite K, and $\tilde{T}_K \xrightarrow{d} T$ (as $K \to \infty$). This can be seen as through inspection of characteristic functions:

$$\begin{aligned}
&\lim_{n\to\infty} E\exp(i\xi T_n) \\
&= \lim_{K\to\infty}\lim_{n\to\infty} E\exp(i\xi\tilde{T}_{nK}) + \lim_{K\to\infty}\lim_{n\to\infty}(E\exp(i\xi T_n) - E\exp(i\xi\tilde{T}_{nK})) \\
&= \lim_{K\to\infty} E\exp(i\xi T_K) + 0 = E\exp(i\xi T),
\end{aligned} \tag{10}$$

where the second equality follows from the convergence in distribution of $\tilde{T}_{nK}$ for any finite K, and the second term is zero because

$$\begin{aligned}
&|\lim_{K\to\infty}\lim_{n\to\infty}(E\exp(i\xi T_n) - E\exp(i\xi\tilde{T}_{nK}))| \\
&\leq \lim_{K\to\infty}\lim_{n\to\infty} |\xi| E|T_n - \tilde{T}_{nK}| = 0
\end{aligned} \tag{11}$$

by assumption and because $|\exp(i\xi x) - \exp(i\xi y)| \leq |\xi||x-y|$. Next, note that the result of the theorem follows if for all λ_1 and λ_2,

$$T_n = n^{-1/2}x_n\lambda_1 + n^{-1}\sum_{t=1}^{n}\Psi(n^{-1/2}\sum_{j=1}^{t}\varepsilon_j)'f(t/n)\lambda_2 \tag{12}$$

converges in distribution to $X(1)\lambda_1 + \int_0^1 \Psi(X(\xi))'f(\xi)d\xi\lambda_2$. Define

$$\tilde{T}_{nK} = n^{-1/2}x_n\lambda_1 + K^{-1}\sum_{i=1}^{K}\Psi(n^{-1/2}\sum_{j=1}^{[in/K]}\varepsilon_j)'f(i/K)\lambda_2. \tag{13}$$

Next, note that as n grows large, for any finite K, by Assumption 2.2, $\tilde{T}_{nK}$ converges in distribution to

$$\tilde{T}_K = X(1)\lambda_1 + K^{-1}\sum_{i=1}^{K}\Psi(X(i/K))'f(i/K)\lambda_2, \tag{14}$$

where $X(.)$ is the Gaussian process described earlier. As $K \to \infty$, by continuity of $\Psi(.)$ and $f(.)$,

$$\tilde{T}_K = X(1)\lambda_1 + K^{-1}\sum_{i=1}^{K}\Psi(X(i/K))'f(i/K)\lambda_2 \xrightarrow{d} X(1)\lambda_1 + \int_0^1 \Psi(X(\xi))'f(\xi)d\xi\lambda_2. \tag{15}$$

Furthermore,

$$|T_n - \tilde{T}_{nK}|$$

$$\leq |K^{-1}\sum_{i=1}^{K}(\sum_{t=[(i-1)(n/K)]+1}^{[in/K]} Kn^{-1}\Psi(n^{-1/2}\sum_{j=1}^{t}\varepsilon_j)'f(t/n) - \Psi(n^{-1/2}\sum_{j=1}^{[in/K]}\varepsilon_j)f(i/K))||\lambda_2|$$

$$\leq |K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)(\Psi(n^{-1/2}\sum_{j=1}^{t}\varepsilon_j) - \Psi(n^{-1/2}\sum_{j=1}^{[ni/K]}\varepsilon_j)'f(t/n))||\lambda_2|$$

$$+|K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)(f(t/n) - f(i/K))\Psi(n^{-1/2}\sum_{j=1}^{[in/K]}\varepsilon_j)||\lambda_2| = (P_1 + P_2)|\lambda_2|, \tag{16}$$

say, and $E|P_2|$ is bounded by

$$K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)|f(t/n) - f(i/K)|E|\Psi(n^{-1/2}\sum_{j=1}^{[in/K]}\varepsilon_j)|$$

$$\leq K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)|f(t/n) - f(i/K)|(\Psi(0) + C_1 + C_2E|n^{-1/2}\sum_{j=1}^{[in/K]}\varepsilon_j)|$$

$$= O(\sup_{x,y:|x-y|\leq K^{-1}}|f(x) - f(y)|), \tag{17}$$

and the last term approaches zero as $K \to \infty$ by the assumed continuity of $f(.)$ on [0,1]. By Assumption 3, we can bound P_1 by

$$\sup_{x\in[0,1]} |f(x)|K^{-1} \sum_{i=1}^{K} \sum_{t=[(i-1)(n/K)]+1}^{[in/K]} (K/n)|n^{-1/2} \sum_{j=1}^{t} \varepsilon_j - n^{-1/2} \sum_{j=1}^{[in/K]} \varepsilon_j|^{\alpha}$$

$$\times (C_1 + C_2|n^{-1/2} \sum_{j=1}^{t} \varepsilon_j| + C_3|n^{-1/2} \sum_{j=1}^{[in/K]} \varepsilon_j|)$$

$$= \sup_{x\in[0,1]} |f(x)|(S_1/C_1 + S_2/C_2 + S_3/C_3), \tag{18}$$

say, where the last result follows from Assumption 3. We can bound the expectation of S_1 as follows:

$$EK^{-1} \sum_{i=1}^{K} \sum_{t=[(i-1)(n/K)]+1}^{[in/K]} (K/n)|n^{-1/2} \sum_{j=1}^{t} \varepsilon_j - n^{-1/2} \sum_{j=1}^{[in/K]} \varepsilon_j|^{\alpha}$$

$$\leq K^{-1} \sum_{i=1}^{K} \sum_{t=[(i-1)(n/K)]+1}^{[in/K]} (K/n)(E|n^{-1/2} \sum_{j=1}^{t} \varepsilon_j - n^{-1/2} \sum_{j=1}^{[in/K]} \varepsilon_j|^2)^{\alpha/2}$$

$$= K^{-1} \sum_{i=1}^{K} \sum_{t=[(i-1)(n/K)]+1}^{[in/K]} (K/n)(E|n^{-1/2} \sum_{j=t+1}^{[in/K]} \varepsilon_j|^2)^{\alpha/2}$$

$$\leq K^{-1} \sum_{i=1}^{K} (Cn^{-1}(n/K))^{\alpha/2} \leq C^{\alpha/2} K^{-\alpha/2}. \tag{19}$$

The first inequality is Jensen's, the second follows by Assumption 1, and the third inequality is Jensen's again. By the Cauchy-Schwartz inequality,

$$E|S_2| = EK^{-1} \sum_{i=1}^{K} \sum_{t=(i-1)(n/K)+1}^{[in/K]} (K/n)|n^{-1/2} \sum_{j=t+1}^{[in/K]} \varepsilon_j|^{\alpha}|n^{-1/2} \sum_{j=1}^{t} \varepsilon_j|$$

$$\leq (K^{-1} \sum_{i=1}^{K} \sum_{t=[(i-1)(n/K)]+1}^{[in/K]} (K/n)E|n^{-1/2} \sum_{j=t+1}^{[in/K]} \varepsilon_j|^{2\alpha})^{1/2}$$

$$\times(K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)E(n^{-1/2}\sum_{j=1}^{t}\varepsilon_j)^2)^{1/2}$$

$$\leq (K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)E(n^{-1/2}\sum_{j=t+1}^{[in/K]}\varepsilon_j)^2)^{\alpha/2}(K^{-1}\sum_{i=1}^{K}\sum_{t=[(i-1)(n/K)]+1}^{[in/K]}(K/n)C)^{1/2}$$

$$\leq K^{-\alpha/2}C^{\alpha/2}C^{1/2} = O(K^{-\alpha/2}). \tag{20}$$

For $E|S_3|$ a similar result is easily obtained, which completes the proof.

Proof of Theorem 2:

Assumption 2.1 holds because Phillips and Solo (1992) in their proof of Theorem 3.11 show that $\lim_{n\to\infty} n^{-1}Ew_n^2 = 0$, where $w_n = \sum_{j=1}^n(\varepsilon_j - (\sum_{k=0}^\infty \phi_k)\eta_j)$. Also, under our assumptions, $\sup_{n\geq 1} n^{-1}E(\sum_{j=1}^n \eta_j)^2 < \infty$. Therefore, by stationarity of ε_j, $E(\sum_{j=a}^b \varepsilon_j)^2 = E(\sum_{j=1}^{b-a+1}\varepsilon_j)^2 \leq (b-a+1)\sup_{n\geq 1} n^{-1}E(\sum_{j=1}^n \varepsilon_j)^2 \leq C(b-a)$ for some $C > 0$. Also, Assumption 2.2 holds by Theorem 3.11 of Phillips and Solo (1992). Finally, for $\xi < \xi'$,

$$\lim_{n\to\infty} E(n^{-1/2}\sum_{j=1}^{[\xi n]}\varepsilon_j)(n^{-1/2}\sum_{j=1}^{[\xi' n]}\varepsilon_j)$$

$$= \lim_{n\to\infty} E(n^{-1/2}\sum_{j=1}^{[\xi n]}\varepsilon_j)^2 + \lim_{n\to\infty} E(n^{-1/2}\sum_{j=1}^{[\xi n]}\varepsilon_j)(n^{-1/2}\sum_{j=[\xi n]+1}^{[\xi' n]}\varepsilon_j), \tag{21}$$

and the first term converges to $\sigma^2\xi$ because it equals

$$\lim_{n\to\infty} E((\sum_{k=0}^{\infty}\phi_k)(n^{-1/2}\sum_{j=1}^{[\xi n]}\eta_j) + w_{[\xi n]})^2$$

$$= \lim_{n\to\infty}(\sum_{k=0}^{\infty}\phi_k)^2 E(n^{-1/2}\sum_{j=1}^{[\xi n]}\eta_j)^2 = \lim_{n\to\infty}\gamma^2(\sum_{k=0}^{\infty}\phi_k)^2 n^{-1}[n\xi] = \sigma^2\xi \tag{22}$$

which follows because $\lim_{n\to\infty} n^{-1}Ew_n^2 = 0$, as shown in the proof of Theorem 3.11 of Phillips and Solo (1992), and by applying the Cauchy-Schwartz inequality . The second term converges to zero because

$$\lim_{n\to\infty} E(n^{-1/2}\sum_{j=1}^{[\xi n]}\varepsilon_j)(n^{-1/2}\sum_{j=[\xi n]+1}^{[\xi' n]}\varepsilon_j)$$

$$= \lim_{n\to\infty} E(\sum_{k=0}^{\infty} \phi_k n^{-1/2} \sum_{j=1}^{[\xi n]} \eta_j + w_{[\xi n]})(\sum_{k=0}^{\infty} \phi_k n^{-1/2} \sum_{j=[\xi n]+1}^{[\xi' n]} \eta_j + w_{[\xi' n]} - w_{[\xi n]}) = 0, \quad (23)$$

again because $\lim_{n\to\infty} n^{-1} E w_n^2 = 0$, by applying the Cauchy-Schwartz inequality, and because

$$E(n^{-1/2} \sum_{j=1}^{[\xi n]} \eta_j)(n^{-1/2} \sum_{j=[\xi n]+1}^{[\xi' n]} \eta_j) = 0. \quad (24)$$

These results identify $X(.)$ as $\sigma W(.)$, thereby completing the proof.

References

Davidson, J. (2000), When is a time series I(0)? Evaluating the memory properties of dynamic models, mimeo, Cardiff University.

Hamilton, J.D. (1994), Time series analysis, Princeton: Princeton University Press.

Hall, P. and C.C. Heyde (1980), Martingale Limit Theory and Its Applications, Academic Press, New York.

Phillips, P. and V. Solo (1992), Asymptotics for linear processes, Annals of Statistics, Vol. 20, No. 2, p. 971-1001.