Very Introductory Elements of Unit Root Econometrics

Consider the AR(1) recursion

$$y_t = \beta_{1_0} y_{t-1} + \varepsilon_t$$ where

$(\varepsilon_t)_{t\in\mathbb{Z}}$ is a white noise process, and $\beta_{1_0}=1$.

Hence the UDC condition is not satisfied and it is possible to prove that there doesn't exist a linear process causal solution with absolutely summable coefficients on $\mathbb{Z}$. We thus study solutions on $\mathbb{N}$. Given the initial condition $y_0=0$, we consider causal solutions on $\mathbb{N}$ of

$$(*)\quad \left.\begin{array}{l} y_0 = 0, \\ y_t = \beta_{1_0} y_{t-1} + \varepsilon_t,\ t\geq 1,\ \beta_{1_0}=1 \end{array}\right\}$$

for which it is easy to see that we obtain the unique causal linear solution over $\mathbb{N}$ which is $(y_t)_{t\in\mathbb{N}}$, with $y_t = \begin{cases} 0, & t=0 \\ \sum_{i=1}^{t} \varepsilon_i, & t>0 \end{cases}$

for $(\varepsilon_t)_{t\in\mathbb{N}}$ the restriction of the aforementioned white noise process on $\mathbb{N}$. Obviously $E(y_t)=0$, $Var(y_t) = \sigma^2 t$, $\forall t>0$, where $0<\sigma^2 := E(\varepsilon_0^2) < +\infty$. Hence $(y_t)_{t\in\mathbb{Z}}$ cannot be stationary (neither strictly, nor weakly - why?), while if $(\varepsilon_t)_{t\in\mathbb{N}}$ satisfies conditions for the validity of some CLT, then $y_T/\sqrt{T}$ converges in distribution to some normal random variable. The framework of $(*)$ is called a unit root framework, while the $(y_t)_{t\in\mathbb{N}}$ solution is called a unit root process over $\mathbb{N}$, w.r.t. $(\varepsilon_t)_{t\in\mathbb{N}}$, and under the initial condition $y_0=0$.[△]

Given the importance of the random walk processes in several areas of empirical hypothesis (e.g. the random walk model for the log-prices of financial assets), it is important to examine the limit theory of the OLSE under $(*)$, and subsequently use it for the examination of testing procedures for hypotheses structures

$$\mathcal{H}_0: \beta_{1_0}=1$$

$$\mathcal{H}_1: \beta_{1_0}\in(-1,1),$$ i.e. unit root tests.

△ When $(\varepsilon_t)_{t\in\mathbb{N}}$ is iid then $(y_t)_{t\in\mathbb{N}}$ is called random walk process.

Standard Wiener Process or Brownian Motion

We need some preparatory work. We first (without mathematical rigour) define a continuous time [on [0,1]] Gaussian process, the distribution of which, which is a Gaussian measure on a function space, will be the normal distribution counterpart, w.r.t. the limiting behavior of appropriate partial sum processes.

Definition. Given a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a standard Wiener process (or Brownian Motion) $W := (W(t))_{t\in[0,1]}$ such that the following properties hold:

1. $W(0) = 0$, $\mathbb{P}$ a.s.,
2. $\forall t, s \in [0,1]$, $h \in \mathbb{R}$: $t+h, s+h \in [0,1]$, $W(t+h) - W(s+h)$ is independent of $W(t) - W(s)$, and $W(t+h) - W(s+h)$ has the same distribution with $W(t) - W(s)$, i.e. the process has independent and stationary increments,
3. $\forall\, 0 < t < s < 1$, $W(t) - W(s) \sim N(0, |t-s|)$, i.e. the process has normally distributed increments.

Using among others the Danniell-Kolmogorov theorem, we can (in a variety of ways) prove that W exists. E.g. we can use the Karhunen-Loève Theorem to establish the existence of the process by representing W as a stochastic series w.r.t. sequences of normal random variables and appropriately orthogonal polynomials. It is possible to prove the following results.

Proposition. In the framework of the definition:

1. W exists,
2. $\forall t > 0$, $W(t) \sim N(0, t)$,

* 3. $W(\cdot, \omega): [0,1] \to \mathbb{R}$ is continuous for $\mathbb{P}$ almost all $\omega \in \Omega$,

** 4. $W(\cdot, \omega): [0,1] \to \mathbb{R}$ is Lebesgue almost nowhere differentiable $\mathbb{P}$ a.s. □

*. 3 essentially refers to versions of W something that lies outside our scope.

** Remember that $W(\cdot,\omega)$ is a function $[0,1]\to\mathbb{R}$ for any $\omega\in\Omega$, while $W(t,\cdot)$ is a random variable $\forall t\in[0,1]$. The dependence on Ω is supressed for notational simplicity.

Given 3, the integral $\int_0^1 f(W_r)dr$, for $f:\mathbb{R}\to\mathbb{R}$ continuous is a well defined random variable when f is continuous, since for P almost all $\omega\in\Omega$, $\int_0^1 f(W(r,\omega))dr$ is a well defined Riemann integral.

E.g. when $f(x)=x^2$, $\int^1 W_r^2 dr$ is a well defined P a.s. positive random variable [A].

An FCLT and the CMT for Appropriately Continuous Functionals

Given the $(\varepsilon_t)_{t\in\mathbb{Z}}$ process as above, and for any $T\in\mathbb{N}^*$, consider the partial sum process (defined on $[0,1]$)

$$W_T(r) := \frac{1}{\sqrt{T}}\sum_{i=1}^{\lfloor rT\rfloor}\varepsilon_i \quad \text{where}$$

$\lfloor x\rfloor$ is the integer part of $x\in\mathbb{R}$. Notice that $W_T(0)=0$ by convention, while $W_T(1)=\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\varepsilon_i$. Furthermore, when $t\leq T$, $\frac{t-1}{T}\leq r<t/T$

then $W_T(r)=\frac{1}{\sqrt{T}}\sum_{i=1}^{\lfloor rT\rfloor}\varepsilon_i = \frac{1}{\sqrt{T}}\sum_{i=1}^{t-1}\varepsilon_i = W_T\left(\frac{t-1}{T}\right)$ but when $r=t/T$

then $W_T(r)=\frac{1}{\sqrt{T}}\sum_{i=1}^{\lfloor rT\rfloor}\varepsilon_i = \frac{1}{\sqrt{T}}\sum_{i=1}^{t}\varepsilon_i = W_T(t/T)$. Hence, $\forall\omega\in\Omega$

$W_T(\cdot,\omega)$ is right continuous with finite left limit. E.g. from the integral properties,

$$\int_0^1 W_T^2(r)dr = \sum_{t=1}^{T}\int_{\frac{t-1}{T}}^{t/T} W_T^2(r)dr = \sum_{t=1}^{T} W_T^2\left(\frac{t-1}{T}\right)\int_{t-1/T}^{t/T}dr$$

$$= \frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^2$$ (B). ((B) works under the convention $\sum_{i=1}^{0} := 0$)

We are interested in the asymptotic behavior of W_T as $T\to+\infty$. The following functional CLT (FCLT) specifies that the latter process (properly standardized) converges weakly to W, as processes with values in an appropriate function space. The proof goes by verifying fidi convergence via a multivariate extension of the CLT for stationary and ergodic s.m.d. processes, and by the verification of a property for the sequence $(W_T)_{T\geq 1}$ called asymptotic tightness.

Theorem [FCLT-Stationary and ergodic s.m.d.]. Suppose that $(\varepsilon_t)_{t\in\mathbb{N}}$ is a strictly stationary and ergodic s.m.d. with respect to some filtration. Then if $0<\sigma^2:=E(\varepsilon_0^2)$, as $T\to+\infty$

$$\frac{1}{\sigma}W_T \xrightarrow{d} W. \square$$

Comments. The aforementioned fidi convergence implies for example that $\forall r\in[0,1]$, $\frac{1}{\sigma}W_T(r)\xrightarrow{d} W(r)\sim N(0,r)$. Hence due to the CMT, $\frac{1}{\sigma^2}W_T^2(1)\xrightarrow{d} W^2(1)\sim\chi^2_1$. (e.g. $\frac{1}{\sigma}W_T(1)\xrightarrow{d} W(1)\sim N(0,1)$, which is essentially—why?—the relevant—which?—CLT).

Theorem (Functional) CMT. If G is a continuous functional defined on the aforementioned function space then as $T\to+\infty$

$$G\left(\frac{1}{\sqrt{\sigma}}W_T\right)\xrightarrow{d} G(W). \square$$

Example. $G(f):=\sigma f$. Then $W_T\xrightarrow{d}\sigma W$. [C]

Example. $G(f):=\sigma^2 f^2$. Then $W_T^2\xrightarrow{d}\sigma^2 W^2$. [D]

Example. $G(f):=\int_0^1\sigma^2 f(r)dr$. Then $\int_0^1 W_T^2(r)\xrightarrow{d}\sigma^2\int_0^1 W^2(r)dr$. [E]

Proposition [JC]

$$\begin{pmatrix} \frac{1}{T}\sum_{i=1}^{T}\varepsilon_i^2 \\ W_T^2(1) \\ \int_0^1 W_T^2(r)dr \end{pmatrix} \xrightarrow{d} \sigma^2 \begin{pmatrix} 1 \\ W^2(1) \\ \int_0^1 W^2(r)dr \end{pmatrix}, \text{ as } T\to+\infty. \square$$

Proof. Birkhoff's LLN implies (why?) $1/T\sum_{i=1}^{T}\varepsilon_i^2 \to \sigma^2$ $\mathbb{P}$ a.s. as $T\to+\infty$. The latter limit is non stochastic. This along with the FCLT, [D], [E], the fact that fidi convergence is implied by the FCLT, imply the result. $\square$

The OLSE and a Dickey-Fuller-Type Unit Root Test.

In the context of the relevant linear statistical model consider the OLSE, $\beta_T = \sum_{t=1}^{T} y_t y_{t-1} \Big/ \sum_{t=1}^{T} y_{t-1}^2 = 1 + \sum_{t=1}^{T}\varepsilon_t y_{t-1} \Big/ \sum_{t=1}^{T} y_{t-1}^2$.

Proposition [OLSE]. Suppose that the assumptions of the FCLT are satisfied and $(y_t)_{t=1,\dots,T}$ is part of the causal linear solution of (*). Then

$$T(\beta_T-1) \xrightarrow{d} \frac{1}{2}\frac{W^2(1)-1}{\int_0^1 W^2(r)dr}. \square$$

Proof. $T(\beta_T-1) = \left(1/T\sum_{t=1}^{T}\varepsilon_t y_{t-1}\right) \Big/ \left(1/T^2\sum_{t=1}^{T} y_{t-1}^2\right)$.

i. For the denominator we have that $\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 =$

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 = \frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{i=1}^{t-1}\varepsilon_i\right)^2 \overset{(B)}{=} \int_0^1 W_T^2(r)dr.$$

ii. For the numerator, $\left(\sum_{t=1}^{T} y_t^2\right) = \sum_{t=1}^{T}(y_{t-1}+\varepsilon_t)^2 =$
$\sum_{t=1}^{T} y_{t-1}^2 + 2\sum_{t=1}^{T}\varepsilon_t y_{t-1} + \sum_{t=1}^{T}\varepsilon_t^2$, hence $\sum_{t=1}^{T}\varepsilon_t y_{t-1} = \frac{1}{2}\left(\sum_{t=1}^{T} y_t^2\right.$
$\left. - \sum_{t=1}^{T} y_{t-1}^2 - \sum_{t=1}^{T}\varepsilon_t^2\right) = \frac{1}{2}\left(y_T^2 - y_0^2 - \sum_{t=1}^{T}\varepsilon_t^2\right) = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t^2$. Hence
$\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t y_{t-1} = \frac{1}{2}y_T^2/T - \frac{1}{2}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2 = \frac{1}{2}W_T^2(1) - \frac{1}{2}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2$.

Considering the expressions in i, ii, the result follows from Proposition [JC], [A] and the CLT (explain!). □

Comments. 1. $\beta_T \to \beta_0$ in probability, with rate T under the current assumption framework. Thus it is called superconsistent.

2. The distribution of $\frac{1}{2}(W_{(1)}^2 - 1)/\int_0^1 W_{(r)}^2 dr$ is well defined, independent of σ^2, and has non-zero mean. Hence $T(\beta_T - 1)$ is asymptotically biased.

Consider the hypothesis structure $H_0: \beta_{1_0} = 1$.
$H_1: \beta_{1_0} \in (-1,1)$

For the test statistic $df_T: T^2(\beta_T - 1)^2$ have that due to the OLSE proposition, that under the H_0, $df_T \rightsquigarrow \frac{1}{4}(W_{(1)}^2-1)^2/\left(\int_0^1 W_{(r)}^2 dr\right)^2$.

The latter has a well defined distribution. For $q_{df}(1-\alpha) :=$
$= \inf\left(x \in \mathbb{R},\ P\left(\frac{1}{4}(W_1^2-1)/\left[\int_0^1 W_r^2 dr\right]^2 \leq x\right) \geq 1-\alpha\right)$

(we can prove that $q_{df}(1-\alpha) > 0$, $\forall \alpha \in (0,1)$ and it is approximable by Monte Carlo Simulations). Consider the following unit root testing procedure for significance level, $\alpha \in (0,1)$:

Reject at α, H_0, iff $df_T > q_{df}(1-\alpha)$.

Given the previous it is ease to show that the procedure is asymptotically exact (Show it!). Consider the following additional result.

Proposition. Suppose that $(\varepsilon_t)_{t\in\mathbb{N}}$ is the relevant part of $(\varepsilon_t)_{t\in\mathbb{Z}}$ which is an ARCH(1) process, that satisfies $E(z_0^3)=0$, $E(z_0^4)=u<+\infty$, $\alpha<\frac{1}{\sqrt{u}}$. Then the testing procedure defined above

is consistent.

Proof. Under the assumption framework above, and if $\mathcal{H}_1$ is true then (explain) $\sqrt{T}(b_T-b_0)\xrightarrow{d} N(0,v(\omega,\alpha,u,b_0))$.
Hence, $T(b_T-1) = T(b_T-b_0)+T(b_0-1)\xrightarrow{d}-\infty$ (why?).
Then $T^2(b_T-1)^2\xrightarrow{d}+\infty$, and $\lim_{T\to+\infty}P(T^2(b_T-1)^2>q_{df}(1-\alpha))$

$$=P(q_{df}(1-\alpha)<+\infty)=1. \square$$

A plethora of unit root tests with differing local power properties have been examined in the relevant literature.

[The notes are in a state of perpetual correction. They do not substitute the lectures. Please r eport any typos to stelios@aueb.gr or the course's e-class.]

For an excellent introductory treatment of such-like notions see inter-alia:

White, H. (2014). Asymptotic theory for econometricians. Academic press.