607
Unit root asymptotics

Jon Faust

http://e105.org/e607

November 30, 2015

▶ **Claims**

- We have already proven a lot of results under the CAN framework
- We have asserted that the asymptotic theory is discontinuous when AR roots reach the unit circle.

  stats. behave differently asymptotically if roots are on the unit circle as oppose to just outside.

- Here we review some of the theory allowing us to prove this.

▶ **Background: standard Brownian motion, also called Weiner process**

▶ **Standard Brownian motion**

- A standard brownian motion is a continuous time process on $r = [0, 1]$, $W(r)$
- Any increment $W(r) - W(l)$, $r > l$ is $N(0, (r - l))$
- Nonoverlapping increments are jointly normal and uncorrelated.
- $W(0) = 0$.

▶ **Continuous time vs. discrete time random walk**

- Discrete time, standard Gaussian r.w.:

$$x_t = \sum_{s=1}^{t} \varepsilon_s, \;\; \varepsilon \sim iidN(0, 1)$$

- $y_{t_2} - y_{t_1}$—the increment between t_1 and t_2—is just the sum of the εs occuring in the interval.
- Thus, the variance is just the length of the interval, $(t_1 - t_2)$

- And nonoverlapping increments are jointly normal and uncorrelated.
- Thus, the Brownian motion looks like a continuous time Gaussian random walk.

▶ **Where we are headed**

- A certain class of persistent processes in the large T limit behave like Brownian motions.
- I'll make this a bit precise, but details, as always, left aside

▶ **A general randow walk-ish process**

- Define the process

$$x_t = \sum_{j=1}^{t} \varepsilon_j$$

- This looks like a random walk, but we'll allow ε to be messy.
- ε can he heteroskedastic and dependent, so long as it satisfies a CLT.
- Let just assume $E\varepsilon_j = 0$ and ε satisfies a CLT:

$$\sqrt{S}\frac{\sum_{j=t_1}^{t_2} \varepsilon_j}{(t_2 - t_1} \to_d N(0, \sigma^2)$$

for some σ^2

That is, over any sufficiently long span, from t_1 to t_2, normalized $\bar{\varepsilon}$ is approximately normal.

▶ **Where we are headed**

- Functional central limit theorems show that

$$\tilde{X}(r) \to_d W(r), r \in [0, 1]$$

where $\tilde{X}(r)$ is a straightforward transform of the x_ts.

▶ **Defining $\tilde{X}$**

- Define

$$\tilde{X}(r) = frac1\sigma T^{1/2} x_{[rT]}$$

where [.] means round down to nearest integer.

- Think of r as a share of the way through the sample, so that $\tilde{X}(r)$ is just the x at position $[rT]$
- Then we also normalize by the deterministic factor $(\sigma T^{1/2})^{-1}$.

▶ **Distribution of increments**

- Take the incrment from ℓ to r

$$\delta(\ell, r) = frac1\sigma T^{1/2} \left(\tilde{X}(r) - \tilde{X}(\ell)\right)$$

- For simplicity, assume that ℓT and rT are integers.

$$\delta(\ell, r) = frac1\sigma T^{1/2} \sum_{s=\ell T+1}^{rT} \varepsilon_s$$

- Ths sum has $T(r-\ell)$ terms so multiply and divide by this value to get:

$$\delta(\ell, r) = frac1\sigma T^{1/2} T(r-\ell)\bar{\varepsilon}_{\ell,r}$$

where the notation for the sample mean should be clear.

- Or:

$$\begin{aligned} \delta(\ell, r) &= \frac{(T(r-\ell))^{1/2}}{\sigma T^{1/2}} \left\{(T(r-\ell))^{1/2}\bar{\varepsilon}_{\ell,r}\right\} \\ \delta(\ell, r) &= \frac{((r-\ell))^{1/2}}{\sigma} \left\{(T(r-\ell))^{1/2}\bar{\varepsilon}_{\ell,r}\right\} \end{aligned}$$

- The item in curly brackets converges to $N(0, \sigma^2)$

  remember our εs satisfy a CLT.

- Thus,

$$\delta(\ell, r) \rightarrow_d N(0, (r-\ell))$$

▶ **Two Nonoverlapping increments**

- We'll need that nonoverlapping increments to be uncorrelated in the limit.
- Since our εs can have very general dependence, nonoverlapping increments need not be uncorrelated, but....
- Fix a right end of the first increment, r_1 and a left end of the second l_2.
- As T gets large, an arbitrarily large number of εs arrive between r_1T and ℓ_2T.
- That is, in terms of the original discrete time process the two intervals are far apart in time.
- Since a CLT applies to the εs, dependence must die out

  so nonoverlapping increments will be uncorrelated in the limit

▶ **Thus,**

- So long as we stick with rs and ℓ such that rT and ℓ are integers, it looks like we have the result we want for large T
- Finishing up: we need to allow for all rs and ℓs in the unit interval.
- And then make sure we are have a coherent definition of what it means for one random function to converge to another.
- These are deep, but don't add much of practical relevance beyond what we've said here.

▶ **What's it worth**

- Our statistics that are most naturally written in terms of the x_ts can also be written in terms of $\tilde{X}(r)$
- And we have a continuous mapping theorem saying that functions of $\tilde{X}$ converge in distribution to the same function of $W(r)$:
$$\phi(\tilde{X}) \to_d \phi(W)$$

▶ **Famous example: Spurious regress**

- In time series, Granger named the curious results we get when we regress one random walk on another the spurious regression problem.
- P.C.B Phillips later showed the $\hat{\beta}$ in this case converges in distribution rather than having a plim.
$$\hat{\beta} \to_d FBW$$
where FBW is 'function of Brownian motions.'
- This means, for example, that $\sqrt{T}(\hat{\beta} - 0)$ will have variance diverging with T.

▶ **Let's sketch the proof**

- Two independent random walks:
$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_{y,t} \\ x_t &= x_{t-1} + \varepsilon_{x,t} \end{aligned}$$
εs mean zero, iid, and variances σ_x^2 and σ_y^2.
- The regression of y on x gives,
$$\hat{\beta} = \frac{\sum x_t y_t}{\sum x_t^2}$$
- Easy to show that this is
$$\hat{\beta} = \frac{\sigma_y}{\sigma_x} \frac{\int_0^1 \tilde{X}(r)\tilde{Y}(r)dr}{\int_0^1 \tilde{X}(r)^2 dr}$$

- Given the above, by our functional CLT/continuous mapping theorem:

$$\hat{\beta} \rightarrow_d \frac{\sigma_y}{\sigma_x} \sigma \frac{\int_0^1 W_x(r) W_y(r) dr}{\int_0^1 W_x(r)^2 dr}$$

  where W_x and W_y are two independent standard Brownian motions.

▶ **OK, that was trivial**

- We basicaly replaced sums by integrals
- We don't have closed form or more meaningful expressions for any of these FBMs, so this doesn't help us understand the shape of the distribution.
- What did we learn?

▶ **What this kind of theory teaches us**

- We learned what power of T we need to normalize by to get $\hat{\beta} - \beta$ to have a well-defined asymptotic distribution.

  T^0 in the spurious regression case, rather than the usual $T^{1/2}$.

- We learned what parameters of the problem, if any, the limit distribution depends on

  (σ_y, σ_x).

- These are very useful facts in helping us understand problems, know what is asymptotically pivotal, and run meaningful Monte Carlos and Bootstraps etc.

▶ **How we'll use this stuff**

- I won't refer to these details much stuff explicitly very much
- I'll just pull out the keys:
  - What power of T, if any, leads to convergence in distribution.
  - What parameters (if any) does the distribution depend on.