Lecture 2C: Unit Root Processes

Vitor Possebom

EESP-FGV

Econometrics 2

Administrative

- Recommended Reading: Hamilton - Chapter 17.1-17.4.
- Problem Set 2 - Deadline: May 23rd at 9:00 am

Outline

1. Motivation
2. Useful Definitions and Results
3. Asymptotic Properties of an AR(1) when $\rho = 1$: Dickey-Fuller Tests
4. Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

Motivation

Motivation

In this lecture, we will pay close attention to the following model:

$$Y_t = \alpha + \rho \cdot Y_{t-1} + \epsilon_t,$$

where $\rho = 1$ and $\{\epsilon_t\}$ is a Gaussian White Noise Process.

This type of process is very special because it generates many problems.

Motivation: Problems generated by Unit Root Processes

1. **Downward bias of autoregressive coefficients:** ρ can be consistently estimated by OLS but the estimator is biased toward zero. This bias is roughly $\mathbb{E}[\hat{\rho} - 1] \approx -{}^{5.3}/_{T}$. This estimation bias causes forecasts of Y_t to perform worse than a pure random walk model.

2. **Non-normally distributed t-statistics:** The nonnormal distribution of $\hat{\rho}$ translates to a nonnormal distribution of its t-statistic so that normal critical values are invalid, implying that usual confidence intervals and hypothesis tests are invalid.

Motivation: Problems generated by Unit Root Processes

3. **Spurious Regression:** When two unit root processes are regressed onto each other, the estimated relationship may appear highly significant using conventional normal critical values although the series are unrelated. (Hamilton's Chapter 18.3)

To have a better understanding of all these problems, we will use a Monte Carlo simulation (`code01_unit-root-problems.R`).

To have some fun, you can find many examples of spurious regressions here: `https://www.tylervigen.com/spurious-correlations`.

Useful Definitions and Results

Outline

1. Motivation
2. Useful Definitions and Results
3. Asymptotic Properties of an AR(1) when $\rho = 1$: Dickey-Fuller Tests
4. Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

Standard Brownian Motion

Definition: Standard Brownian Motion (`https://youtu.be/7mmeksMiXp4`)

Standard Brownian Motion $W(\cdot)$ is a continuous-time stochastic process, associating each date $t \in [0, 1]$ with the scalar $W(t)$ such that:

1. $W(0) = 0$
2. For any dates $0 \leq t_1 < t_2 < \ldots < t_k \leq 1$, the changes $[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \ldots, [W(t_k) - W(t_{k-1})]$ are independent multivariate Gaussian with $[W(s) - W(t)] \sim N(0, s - t)$.
3. For any given realization $W(t)$ is continuous in t with probability 1.

Convergence in Distribution for Random Functions

Convergence in Distribution for Random Functions

Definition: Convergence in Distribution for Random Functions

Let $S(\cdot)$ represent a continuous-time stochastic process with $S(r)$ representing its value at some date r for $r \in [0, 1]$. Moreover, suppose that, for any given realization, $S(\cdot)$ is a continuous function of r with probability 1. For $\{S_T(\cdot)\}_{T=1}^{+\infty}$ a sequence of continuous functions, we say that $S_T(\cdot) \xrightarrow{d} S(\cdot)$ if all of the following hold:

1. For any finite collection of k particular dates, $0 \leq r_1 < r_2 < \ldots < r_k \leq 1$, the sequence of $k-$dimensional random vectors $\{Y_T\}_{T=1}^{+\infty}$ converges in distribution to the vector Y, where $Y_T = [S_T(r_1), S_T(r_2), \ldots, S_T(r_k)]'$ and $Y = [S(r_1), S(r_2), \ldots, S(r_k)]'$.
2. For each $\epsilon > 0$, $\mathbb{P}\left[\sup_{|r-s|<\delta} |S_T(r) - S_T(s)| > \epsilon\right] \to 0$ uniformly in T as $\delta \to 0$.
3. $\mathbb{P}[|S_T(0)| > \lambda] \to 0$ uniformly in T as $\lambda \to +\infty$.

Convergence in Probability for Random Functions

Convergence in Probability for Random Functions

Definition: Convergence in Probability for Random Functions

Let $\{S_T(\cdot)\}_{T=1}^{+\infty}$ and $\{V_T(\cdot)\}_{T=1}^{+\infty}$ denote sequences of continuous random functions. We say that

$$S_T(\cdot) \xrightarrow{p} V_T(\cdot)$$

if

$$\sup_{r\in[0,1]} |S_T(r) - V_T(r)| \xrightarrow{p} 0.$$

Useful Results

Useful Results

Suppose that $\{Y_t\}$ follows a random walk without drift,

$$Y_t = Y_{t-1} + \epsilon_t,$$

where $Y_0 = 0$ and $\{\epsilon_t\}$ is an i.i.d. sequence with mean zero and variance σ^2. Let $W(\cdot)$ be a standard Brownian Motion. Then,

1. $\frac{\sum_{t=1}^{T} \epsilon_t}{T^{1/2}} \xrightarrow{d} \sigma \cdot W(1)$

Useful Results

2. $$\frac{\sum_{t=1}^{T} Y_{t-1} \cdot \epsilon_t}{T} \xrightarrow{d} \frac{\sigma^2 \cdot \left\{ [W(1)]^2 - 1 \right\}}{2}$$

3. $$\frac{\sum_{t=1}^{T} t \cdot \epsilon_t}{T^{3/2}} \xrightarrow{d} \sigma \cdot W(1) - \sigma \int_0^1 W(r)\, dr$$

4. $$\frac{\sum_{t=1}^{T} Y_{t-1}}{T^{3/2}} \xrightarrow{d} \sigma \int_0^1 W(r)\, dr$$

5. $$\frac{\sum_{t=1}^{T} Y_{t-1}^2}{T^2} \xrightarrow{d} \sigma^2 \int_0^1 [W(r)]^2\, dr$$

Useful Results

6. $\dfrac{\sum_{t=1}^{T} t \cdot Y_{t-1}}{T^{5/2}} \xrightarrow{d} \sigma \int_0^1 r \cdot W(r)\ dr$

7. $\dfrac{\sum_{t=1}^{T} t \cdot Y_{t-1}^2}{T^{3}} \xrightarrow{d} \sigma^2 \int_0^1 r \cdot [W(r)]^2\ dr$

8. $\dfrac{\sum_{t=1}^{T} t^{v}}{T^{(v+1)}} \rightarrow \dfrac{1}{v+1}$

Asymptotic Properties of an AR(1) when $\rho = 1$: Dickey-Fuller Tests

Outline

1. Motivation

2. Useful Definitions and Results

3. Asymptotic Properties of an AR(1) when $\rho = 1$: Dickey-Fuller Tests

4. Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

No Constant Term or Time Trend in the Regression, True Process is a Random Walk

True Process: $Y_t = \rho \cdot Y_{t-1} + \epsilon_t$ where $\rho = 1$, $\{\epsilon_t\}$ is i.i.d. with mean zero and variance σ^2.

Estimating Equation: $Y_t = \rho \cdot Y_{t-1} + \epsilon_t$

$$T \cdot (\hat{\rho}_T - 1) \xrightarrow{d} \frac{(1/2) \cdot \left\{ [W(1)]^2 - 1 \right\}}{\int_0^1 [W(r)]^2 \, dr}$$

$$\frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\rho_T}} \xrightarrow{d} \frac{(1/2) \cdot \left\{ [W(1)]^2 - 1 \right\}}{\left\{ \int_0^1 [W(r)]^2 \, dr \right\}^{1/2}}$$

Constant Term but No Time Trend in the Regression, True Process is a Random Walk

Constant Term but No Time Trend in the Regression, True Process is a Random Walk

True Process: $Y_t = \alpha + \rho \cdot Y_{t-1} + \epsilon_t$ where $\alpha = 0, \rho = 1$, $\{\epsilon_t\}$ is i.i.d. with mean zero and variance σ^2.

Estimating Equation: $Y_t = \alpha + \rho \cdot Y_{t-1} + \epsilon_t$

$$T \cdot (\hat{\rho}_T - 1) \xrightarrow{d} \frac{(1/2) \cdot \left\{ [W(1)]^2 - 1 \right\} - W(1) \cdot \int_0^1 W(r)\, dr}{\int_0^1 [W(r)]^2\, dr - \left[\int_0^1 W(r)\, dr \right]^2}$$

$$\frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\rho_T}} \xrightarrow{d} \frac{(1/2) \cdot \left\{ [W(1)]^2 - 1 \right\} - W(1) \cdot \int_0^1 W(r)\, dr}{\left\{ \int_0^1 [W(r)]^2\, dr - \left[\int_0^1 W(r)\, dr \right]^2 \right\}^{1/2}}$$

Constant Term but No Time Trend in the Regression, True Process is a Random Walk with Drift

Constant Term but No Time Trend in the Regression, True Process is a Random Walk with Drift

True Process: $Y_t = \alpha + \rho \cdot Y_{t-1} + \epsilon_t$ where $\alpha \neq 0, \rho = 1$, $\{\epsilon_t\}$ is i.i.d. with mean zero and variance σ^2.

Estimating Equation: $Y_t = \alpha + \rho \cdot Y_{t-1} + \epsilon_t$

$$\begin{bmatrix} T^{1/2}\left(\hat{\alpha}_T - \alpha\right) \\ T^{3/2}\left(\hat{\rho}_T - 1\right) \end{bmatrix} \xrightarrow{d} N\left(0, \sigma^2 \cdot \begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha^2/3 \end{bmatrix}^{-1}\right)$$

Intuition: The time trend generated by the drift term dominates the asymptotic distribution.

Constant Term and Time Trend in the Regression, True Process is a Random Walk with Drift or Without Drift

Constant Term and Time Trend in the Regression, True Process is a Random Walk with Drift or Without Drift

True Process: $Y_t = \alpha + \rho \cdot Y_{t-1} + \epsilon_t$ where $\rho = 1$, $\{\epsilon_t\}$ is i.i.d. with mean zero and variance σ^2.

Estimating Equation: $Y_t = \alpha + \delta \cdot t + \rho \cdot Y_{t-1} + \epsilon_t$.

Asymptotic Distribution of $\hat{\rho}_T$ is slightly more complicated than in the previous cases. See Case 4 by Hamilton (1994, p. 497).

Hamilton (1994, p. 501) has an interesting discussion on how to choose the best Dickey-Fuller Test for each context.

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

Outline

1. Motivation

2. Useful Definitions and Results

3. Asymptotic Properties of an AR(1) when $\rho = 1$: Dickey-Fuller Tests

4. Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

Let $\{Y_t\}$ be a $AR(p)$ stochastic process possibly with a trend term, i.e.,

$$Y_t = \alpha + \delta \cdot t + \phi_1 \cdot Y_{t-1} + \phi_2 \cdot Y_{t-2} + \ldots + \phi_p \cdot Y_{t-p} + \epsilon_t.$$

We want to test whether $\{Y_t\}$ has one unit root, i.e., if the polynomial function $1 - \phi_1 \cdot z - \phi_2 \cdot z^2 - \ldots - \phi_p \cdot z^p$ has one (and only one) unit root.

Why do we test for unit roots? It is usually a pre-test for a more complicated model, such as a VAR. If we find a unit root, we use the first difference of the variable instead of its level.

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

To test for the presence of one unit root, we implement the Augment Dickey-Fuller test. It uses the following estimating regressions:

- No Drift and No Deterministic Time Trend

$$\Delta Y_t = \gamma \cdot Y_{t-1} + \sum_{i=2}^{p} \beta_i \cdot \Delta Y_{t-i+1} + \epsilon_t$$

$$\rightarrow (\tau 1) \quad H_0 : \rho = 1$$

Test $(\tau 1)$: Presence of a unit root under the null.

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

OBS 1: By construction, $\gamma = \rho - 1$. The ADF tests the null that $\gamma = 0$ against the alternative that $\gamma < 0$. To connect this test with the unit root for the level of Y_t, I wrote the null as $\rho = 1$.

The same is true for the next two tests.

OBS 2: My notation here differs slightly from Hamilton's notation. I described the regression that is run by function `ur.df` in R. Hamilton uses the following notation: $Y_t = \rho \cdot Y_{t-1} + \sum_{i=2}^{p} \beta_i \cdot \Delta Y_{t-i+1} + \epsilon_t$.

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

- With Drift but No Deterministic Time Trend

$$\Delta Y_t = \gamma \cdot Y_{t-1} + \alpha + \sum_{i=2}^{p} \beta_i \cdot \Delta Y_{t-i+1} + \epsilon_t$$

$$\rightarrow (\phi 1) \quad H_0 : \rho = 1 \quad \& \quad \alpha = 0$$

$$\rightarrow (\tau 2) \quad H_0 : \rho = 1$$

Test ($\phi 1$): Presence of a unit root and absence of drift under the null.

Test ($\tau 2$): Presence of a unit root under the null.

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

- With Drift and a Deterministic Time Trend

$$\Delta Y_t = \gamma \cdot Y_{t-1} + \delta \cdot t + \alpha + \sum_{i=2}^{p} \beta_i \cdot \Delta Y_{t-i+1} + \epsilon_t$$

$$\rightarrow (\phi 2) \quad H_0 : \rho = 1 \quad \& \quad \delta = 0 \quad \& \quad \alpha = 0$$

$$\rightarrow (\phi 3) \quad H_0 : \rho = 1 \quad \& \quad \delta = 0$$

$$\rightarrow (\tau 3) \quad H_0 : \rho = 1$$

Test ($\phi 2$): Presence of unit root, absence of drift and absence of time trend under the null.
Test ($\phi 3$): Presence of a unit root and absence of time trend under the null.
Test ($\tau 3$): Presence of a unit root under the null.

Testing for Unit Roots in an AR(p): Augmented Dickey-Fuller Tests

Choosing which regression to run:

- Constant term is usually included.
- When a time series show a trend, we just include a trend term.
- In any one of those test, we have to choose the order p. To do so, we can use BIC.

In R, we can use function `ur.df` in package `urca` to implement the Augmented Dickey-Fuller test.

See file `code02_testing-unit-root.R` for explanation on how to interpret it.

If you want the technical details behind this test, check Chapter 17.7 by Hamilton.

Thank you!

Contact Information:

Vitor Possebom

E-mail: vitor.possebom@fgv.br

Website: sites.google.com/site/vitorapossebom/

References