3.2.2 Phillips-Perron (PP) Test

The model is given by:

$$y_t = \phi_1 y_{t-1} + u_t, \qquad u_t = \sum_{s=0}^{\infty} \psi_s \epsilon_{t-s}, \qquad \epsilon_t \sim \text{ iid}(0, \sigma_\epsilon^2),$$

where $\psi_0 = 0$ and $\sum_{s=0}^{\infty} s|\psi_s| < \infty$.

Note that the errors are serially correlated and heteroskedastic.

The autocovariance function of u_t is:

$$\gamma(\tau) = \mathrm{E}(u_t u_{t-\tau}) = \sigma_\epsilon^2 \sum_{s=0}^{\infty} \psi_s \psi_{s+\tau}, \qquad \tau = 0, 1, 2, \cdots.$$

Define the long-run variance of u_t as:

$$\lambda^2 = \lim_{T\to\infty} \frac{1}{T}\mathrm{E}((\sum_{t=1}^{T} u_t)^2) = \sum_{\tau=-\infty}^{\infty} \gamma(\tau) = \gamma(0) + 2\sum_{\tau=1}^{\infty} \gamma(\tau) = \sigma_\epsilon^2(\sum_{j=0}^{\infty} \psi_j)^2.$$

The PP test statistic $\tilde{t}_T$ is:

$$\tilde{t}_T = \left(\frac{\gamma(0)}{\lambda^2}\right)^{1/2} t_T - \frac{1}{2\lambda}\frac{T s_\phi}{s_T}(\lambda^2 - \gamma(0)),$$

where

t_T denotes the t statistic of $\hat{\phi}_1$, s_ϕ is the standard error of $\hat{\phi}_1$, and
$s_T^2 = \dfrac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{\phi}_1 y_{t-1})^2.$

Estimate λ by:

$$\hat{\lambda} = \hat{\gamma}(0) + 2\sum_{\tau=1}^{q} k_1(\frac{\tau}{q+1})\hat{\gamma}(\tau),$$

which is called **Newey-West estimator**, where $k_1(x) = 1-|x|$ for $x \leq 1$ and $k_1(x) = 0$ for $x > 1$, which is called **Bartlett kernel**, or

$$\hat{\lambda} = \hat{\gamma}(0) + 2\sum_{\tau=1}^{q} k_2(\frac{\tau}{q+1})\hat{\gamma}(\tau),$$

where $k_2(x) = 1 - 6x^2 + 6x^3$ for $0 \leq x \leq \frac{1}{2}$, $k_2(x) = 2(1-x)^3$ for $\frac{1}{2} \leq x \leq 1$ and $k_2(x) = 0$ for $x > 1$, which is called **Parzen kernel**, or

$$\hat{\lambda} = \frac{T}{T-1}\left(\hat{\gamma}(0) + \sum_{\tau=1}^{T-1} k_3(\frac{\tau}{q+1})\hat{\gamma}(\tau)\right),$$

where $k_3(x) = \dfrac{3}{(6\pi x/5)^2}\left(\dfrac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right)$, which is called the **second-order spectrum kernel**.

We need to choose the bandwidth q.

Use the same statistical tables as before to test $H_0: \phi_1 = 1$ against $H_1: \phi_1 < 1$.

Some Formulas:

For proof, we use following formulas.

Let $u_t = \psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, where $\sum_{j=0}^{\infty} j|\psi_j| < \infty$ and $\{\epsilon_t\}$ is an i.i.d. sequence with mean zero, variance σ^2 and finite fourth moment.

Define:

$\gamma(j) = \mathrm{E}(u_t u_{t-j}) = \sigma^2 \sum_{s=0}^{\infty} \psi_s \psi_{s+j} \quad$ for $j = 0, 1, 2, \cdots$,

$\lambda = \sigma \sum_{j=0}^{\infty} \psi_j = \sigma\psi(1)$,

$\xi_t = \sum_{i=1}^{t} u_i$ for $t = 1, 2, \cdots, T \quad$ and $\quad \xi_0 = 0$.

Then,

1. $T^{-1/2} \sum_{t=1}^{T} u_t \longrightarrow \lambda W(1)$

2. $T^{-1/2} \sum_{t=1}^{T} u_{t-j}\epsilon_t \longrightarrow N(0, \sigma^2\gamma(0)), \quad \text{for } j = 1, 2, \cdots$

3. $T^{-1} \sum_{t=1}^{T} u_t u_{t-j} \longrightarrow \gamma(j), \quad \text{for } j = 1, 2, \cdots$

4. $T^{-1} \sum_{t=1}^{T} \xi_{t-1}\epsilon_t \longrightarrow \frac{1}{2}\sigma\lambda(W(1)^2 - 1)$

5. $$T^{-1}\sum_{t=1}^{T}\xi_{t-1}u_{t-j} \longrightarrow \begin{cases} \frac{1}{2}(\lambda^2 W(1)^2 - \gamma(0)), & \text{for } j = 0, \\ \frac{1}{2}(\lambda^2 W(1)^2 - \gamma(0)) + \sum_{i=0}^{j-1}\gamma(i), & \text{for } j = 1, 2, \cdots \end{cases}$$

6. $$T^{-3/2}\sum_{t=1}^{T}\xi_{t-1} \longrightarrow \lambda\int_0^1 W(r)\mathrm{d}r$$

7. $$T^{-3/2}\sum_{t=1}^{T}tu_{t-j} \longrightarrow \lambda\left(W(1) - \int_0^1 W(r)\mathrm{d}r\right), \quad \text{for } j = 0, 1, 2, \cdots$$

8. $$T^{-2}\sum_{t=1}^{T}\xi_{t-1}^2 \longrightarrow \lambda^2\int_0^1 (W(r))^{-2}\mathrm{d}r$$

9. $T^{-5/2} \sum_{t=1}^{T} t\xi_{t-1} \longrightarrow \lambda \int_0^1 rW(r)\mathrm{d}r$

10. $T^{-3} \sum_{t=1}^{T} t\xi_{t-1} \longrightarrow \lambda^2 \int_0^1 r(W(r))^2\mathrm{d}r$

11. $T^{-(\mu-1)} \sum_{t=1}^{T} t^{\mu} \longrightarrow \dfrac{1}{\mu+1}, \quad \text{for } \mu = 0, 1, 2, \cdots$

3.3 Cointegration (共和分)

1. For a scalar y_t, when $\Delta y_t = y_t - y_{t-1}$ is a white noise (i.e., iid), we write $\Delta y_t \sim I(1)$.

2. **Definition of Cointegration:**

   Suppose that each series in a $g \times 1$ vector y_t is $I(1)$, i.e., each series has unit root, and that a linear combination of each series (i.e, $a'y_t$ for a nonzero vector a) is $I(0)$, i.e., stationary.

   Then, we say that y_t has a cointegration.

3. **Example:**

Suppose that $y_t = (y_{1,t},\ y_{2,t})'$ is the following vector autoregressive process:

$$y_{1,t} = \phi_1 y_{2,t} + \epsilon_{1,t},$$
$$y_{2,t} = y_{2,t-1} + \epsilon_{2,t}.$$

Then,

$$\Delta y_{1,t} = \phi_1 \epsilon_{2,t} + \epsilon_{1,t} - \epsilon_{1,t-1}, \quad \text{(MA(1) process)},$$
$$\Delta y_{2,t} = \epsilon_{2,t},$$

where both $y_{1,t}$ and $y_{2,t}$ are $I(1)$ processes.

The linear combination $y_{1,t} - \phi_1 y_{2,t}$ is $I(0)$.

In this case, we say that $y_t = (y_{1,t},\ y_{2,t})'$ is cointegrated with $a = (1,\ -\phi_1)$.

$a = (1,\ -\phi_1)$ is called the cointegrating vector, which is not unique.

Therefore, the first element of a is set to be one.

4. Suppose that $y_t \sim I(1)$ and $x_t \sim I(1)$.

   For the regression model $y_t = x_t\beta + u_t$, OLS does not work well if we do not have the β which satisfies $u_t \sim I(0)$.

   $\Longrightarrow$ **Spurious regression (見せかけの回帰)**

5. Suppose that $y_t \sim I(1)$, y_t is a $g \times 1$ vector and $y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}$.

$y_{2,t}$ is a $k \times 1$ vector, where $k = g - 1$.

Consider the following regression model:

$$y_{1,t} = \alpha + \gamma' y_{2,t} + u_t, \qquad t = 1, 2, \cdots, T.$$

OLSE is given by:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} T & \sum y_{2,t}' \\ \sum y_{2,t} & \sum y_{2,t} y_{2,t}' \end{pmatrix}^{-1} \begin{pmatrix} \sum y_{1,t} \\ \sum y_{1,t} y_{2,t} \end{pmatrix}.$$

Next, consider testing the null hypothesis $H_0: \; R\gamma = r$, where R is a $m \times k$ matrix ($m \leq k$) and r is a $m \times 1$ vector.

The F statistic, denoted by F_T, is given by:

$$F_T = \frac{1}{m}(R\hat{\gamma} - r)' \left(s_T^2 \begin{pmatrix} 0 & R \end{pmatrix} \begin{pmatrix} T & \sum y_{2,t}' \\ \sum y_{2,t} & \sum y_{2,t} y_{2,t}' \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ R' \end{pmatrix} \right)^{-1} (R\hat{\gamma} - r),$$

where

$$s_T^2 = \frac{1}{T-g} \sum_{t=1}^{T} (y_{1,t} - \hat{\alpha} - \hat{\gamma}' y_{2,t})^2.$$

When we have the γ such that $y_{1,t} - \gamma y_{2,t}$ is stationary, OLSE of γ, i.e., $\hat{\gamma}$, is not statistically equal to zero.

When the sample size T is large enough, H_0 is rejected by the F test.

6. Phillips, P.C.B. (1986) "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, Vol.33, pp.95 – 131.

Consider a $g \times 1$ vector y_t whose first difference is described by:

$$\Delta y_t = \Psi(L)\epsilon_t = \sum_{s=0}^{\infty} \Psi_s \epsilon_{t-s},$$

for ϵ_t an i.i.d. $g \times 1$ vector with mean zero , variance $\mathrm{E}(\epsilon_t \epsilon_t') = PP'$, and finite fourth moments and where $\{s\Psi_s\}_{s=0}^{\infty}$ is absolutely summable.

Let $k = g - 1$ and $\Lambda = \Psi(1)P$.

Partition y_t as $y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}$ and $\Lambda\Lambda'$ as $\Lambda\Lambda' = \begin{pmatrix} \Sigma_{11} & \Sigma_{21}' \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where $y_{1,t}$ and Σ_{11} are scalars, $y_{2,t}$ and Σ_{21} are $k \times 1$ vectors, and Σ_{22} is a $k \times k$ matrix.

Suppose that $\Lambda\Lambda'$ is nonsingular,and define $\sigma_1^{*2} = \Sigma_{11} - \Sigma_{21}'\Sigma_{22}^{-1}\Sigma_{21}$.

Let L_{22} denote the Cholesky factor of Σ_{22}^{-1}, i.e., L_{22} is the lower triangular matrix satisfying $\Sigma_{22}^{-1} = L_{22}L_{22}'$.

Then, (a) – (c) hold.

(a) OLSEs of α and γ in the regression model $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$, denoted by $\hat{\alpha}_T$ and $\hat{\gamma}_T$, are characterized by:

$$\begin{pmatrix} T^{-1/2}\hat{\alpha}_T \\ \hat{\gamma}_T - \Sigma_{22}^{-1}\Sigma_{21} \end{pmatrix} \longrightarrow \begin{pmatrix} \sigma_1^* h_1 \\ \sigma_1^* L_{22} h_2 \end{pmatrix},$$

where

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} 1 & \int_0^1 W_2^*(r)'\mathrm{d}r \\ \int_0^1 W_2^*(r)\mathrm{d}r & \int_0^1 W_2^*(r)W_2^*(r)'\mathrm{d}r \end{pmatrix}^{-1} \begin{pmatrix} \int_0^1 W_1^*(r)\mathrm{d}r \\ \int_0^1 W_2^*(r)W_1^*(r)\mathrm{d}r \end{pmatrix},$$

where $W_1^*(r)$ and $W_2^*(r)$ denote scalar and g-dimensional standard Brownian motions, and $W_1^*(r)$ is independent of $W_2^*(r)$.

(b) The sum of squared residuals, denoted by $\mathrm{RSS}_T = \sum_{t=1}^{T} \hat{u}_t^2$, satisfies

$$T^{-2}\mathrm{RSS}_T \longrightarrow \sigma_1^{*2} H,$$

where

$$H = \int_0^1 (W_1^*(r))^2 \mathrm{d}r - \left(\begin{pmatrix} \int_0^1 W_1^*(r)\mathrm{d}r \\ \int_0^1 W_2^*(r)W_1^*(r)\mathrm{d}r \end{pmatrix}' \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right)^{-1}.$$

(c) The F_T test satisfies:

$$\begin{aligned}T^{-1}F_T \ \longrightarrow\ & \frac{1}{m}(\sigma_1^* R^* h_2 - r^*)' \\ & \times\left(\sigma_1^{*2} H \begin{pmatrix} 0 & R^* \end{pmatrix} \begin{pmatrix} 1 & \int_0^1 W_2^*(r)' \mathrm{d}r \\ \int_0^1 W_2^*(r) \mathrm{d}r & \int_0^1 W_2^*(r) W_2^*(r)' \mathrm{d}r \end{pmatrix}^{-1} \begin{pmatrix} 0 & R^* \end{pmatrix}'\right)^{-1} \\ & \times(\sigma_1^* R^* h_2 - r^*),\end{aligned}$$

where $R^* = RL_{22}$ and $r^* = r - R\Sigma_{22}^{-1}\Sigma_{21}$.

(a) indicates that OLSE $\hat{\gamma}_T$ is not consistent.

(b) indicates that $s_T^2 = \dfrac{1}{T-g}\displaystyle\sum_{t=1}^{T}\hat{u}_t^2$ diverges.

(c) indicates that F_T diverges.

$\Longrightarrow$ **Spurious regression (見せかけの回帰)**

7. **Resolution for Spurious Regression:**

Suppose that $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$ is a spurious regression.

(1) Estimate $y_{1,t} = \alpha + \gamma' y_{2,t} + \phi y_{1,t-1} + \delta y_{2,t-1} + u_t$.

Then, $\hat{\gamma}_T$ is $\sqrt{T}$-consistent, and the t test statistic goes to the standard normal distribution under $H_0: \ \gamma = 0$.

(2) Estimate $\Delta y_{1,t} = \alpha + \gamma' \Delta y_{2,t} + u_t$. Then, $\hat{\alpha}_T$ and $\hat{\beta}_T$ are $\sqrt{T}$-consistent, and the t test and F test make sense.

(3) Estimate $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$ by the Cochrane-Orcutt method, assuming that u_t is the first-order serially correlated error.

Usually, choose (2).

However, there are two exceptions.

(i) The true value of ϕ is not one, i.e., less than one.

(ii) $y_{1,t}$ and $y_{2,t}$ are the cointegrated processes.

In these two cases, taking the first difference leads to the misspecified regression.

8. **Cointegrating Vector:**

   Suppose that each element of y_t is $I(1)$ and that $a'y_t$ is $I(0)$.

   a is called a **cointegrating vector (共和分ベクトル)**, which is not unique.

   Set $z_t = a'y_t$, where z_t is scalar, and a and y_t are $g \times 1$ vectors.

For $z_t \sim I(0)$ (i.e., stationary),

$$T^{-1} \sum_{t=1}^{T} z_t^2 = T^{-1} \sum_{t=1}^{T} (a' y_t)^2 \longrightarrow \mathrm{E}(z_t^2).$$

For $z_t \sim I(1)$ (i.e., nonstationary, i.e., a is not a cointegrating vector),

$$T^{-2} \sum_{t=1}^{T} (a' y_t)^2 \longrightarrow \lambda^2 \int_0^1 (W(r))^2 \, \mathrm{d}r,$$

where $W(r)$ denotes a standard Brownian motion and λ^2 indicates variance of $(1 - L)z_t$.

If a is not a cointegrating vector, $T^{-1} \sum_{t=1}^{T} z_t^2$ diverges.

$\Longrightarrow$ We can obtain a consistent estimate of a cointegrating vector by minimizing $\sum_{t=1}^{T} z_t^2$ with respect to a, where a normalization condition on a has to be imposed.

The estimator of the a including the normalization condition is super-consistent (T-consistent).