Econometrics II's Final Exam.
Solution

Ryo Sakamoto
Royou Kiku

1 Consider the following regression model:

$$y = X\beta + u \quad u \sim \mathcal{N}(0, \sigma^2 I_n),$$

where y and u denote $n \times 1$ vectors, β indicates a $k \times 1$ vector, and X represents a $n \times k$ matrix. The explanatory variable X is assumed to be independent of the error term u. Answer the following questions.

(1) We estimate β by OLS (ordinary least squares method). Set up th eoptimization problem and derive the OLS estimator of β.

Solution:

The optimization problem is given by

$$\max_{\beta} \; S(\beta),$$

where $S(\beta) = (y - X\beta)'(y - X\beta)$. We denote the OLS estimator by $\widehat{\beta}$. Then, the first order condition is:

$$\begin{aligned} & \nabla_\beta S(\widehat{\beta}) = 0 \\ \iff \quad & 2X'(y - X\widehat{\beta}) = 0. \end{aligned}$$

Solving this equation, we have the OLS estimator:

$$\widehat{\beta} = (X'X)^{-1}X'y.$$

(2) We estimate β by MLE (maximization likelihood estimation method). Obtain the likelihood function and derive the ML estimator of β.

Solution:

The assumption $u \sim \mathcal{N}(0, \sigma^2 I_n)$ implies that the error term u_i is independently and identically distributed for $i = 1, \cdots, n$. Then, the joint density of $u_i, \; i = 1, \cdots, n$ is given by

$$\begin{aligned} f(u_i, i = 1, \cdots, n) &= \prod_{i=1}^{n} f(u_i) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{u_i^2}{2\sigma^2}\right). \end{aligned}$$

Using the change of variables method, we obtain the likelihood function $L(\theta)$:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_i - x_i\beta)^2\right\},$$

where $\theta := (\beta', \sigma^2)' \in \mathbb{R}^{k+1}$ indicates the parameter vector and x_i is a $1 \times k$ vector. Taking a logarithm, we have the log-likelihood function:

$$\log L(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i\beta)^2.$$

Then, we consider the following optimization problem:

$$\max_{\theta} \ \log L(\theta).$$

Denoting the ML estimator by $\tilde{\theta} := (\tilde{\beta}', \tilde{\sigma}^2)$, the first order conditions are:

$$\nabla_\theta \log L(\tilde{\theta}) = \begin{pmatrix} \nabla_\beta \log L(\tilde{\theta}) \\ \nabla_{\sigma^2} \log L(\tilde{\theta}) \end{pmatrix} = 0.$$

Solving these conditions, we have the ML estimator: [1]

$$\tilde{\beta} = (X'X)^{-1}X'y,$$
$$\tilde{\sigma}^2 = \frac{1}{n}(y - X\tilde{\beta})'(y - X\tilde{\beta}).$$

(3) We estimate β by MM (method of moment). Set up the problem and derive the MM estimator of β.

Solution:

Since X is independent of u, we have the following orthogonality condition:

$$\mathbb{E}[X'u] = 0.$$

Thus, the MM estimator $\overline{\beta}$ satisfies the following condition:

$$\frac{1}{n}\sum_{i=1}^{n} x_i' u_i = 0 \iff \frac{1}{n}\sum_{i=1}^{n} x_i'(y_i - x_i\overline{\beta}) = 0.$$

Arranging this expression, we have the MM estimator:

$$\overline{\beta} = (X'X)^{-1}X'y.$$

[1] It is sufficient to obtain the estimator of β. However, the first order conditions are the system of equations that the estimator must satisfy. Therefore, you need to show the first order conditions in terms of β and σ^2.

2 Consider the following regression model:

$$y = X\beta + u \quad u \sim \mathcal{N}(0, \sigma^2 I_n),$$

where y and u denote $n \times 1$ vectors, β indicates a $k \times 1$ vector, and X represents a $n \times k$ matrix. The explanatory variable X is assumed to be correlated with error term u. Answer the following questions.

(4) Show that the OLS estimator, denoted by $\widehat{\beta}$, is inconsistent.

Solution:

The OLS estimator is expressed as follows:

$$\begin{aligned}\widehat{\beta} &= \beta + (X'X)^{-1}X'u \\ &= \beta + \left(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i' u_i\right).\end{aligned}$$

By the Law of Large Numbers,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} x_i' x_i &\xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}[x'x] =: M_{xx}, \\ \frac{1}{n}\sum_{i=1}^{n} x_i' u_i &\xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}[x'u] =: M_{xu} \neq 0,\end{aligned}$$

where $M_{xu} \neq 0$ since x_i is correlated with u_i. By the continuous mapping theorem, we have

$$\widehat{\beta} = \beta + \left(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i' u_i\right) \xrightarrow[n\to\infty]{\mathbb{P}} \beta + M_{xx}^{-1} M_{xu} \neq \beta,$$

which concludes that $\widehat{\beta}$ is inconsistent.

(5) Suppose that Z (i.e., $n \times r$ matrix for $r > k$) is independent of u but it is highly correlated with X. Using the instrumental variable$Z$, derive the GMM (generalized method of moment) estimator, denoted by $\tilde{\beta}$.

Solution:

Since Z is independent of u, we have the orthogonality condition:

$$\mathbb{E}[Z'u] = 0.$$

And, its empirical counterpart is:

$$\frac{1}{n}\sum_{i=1}^{n} z_i'(y_i - x_i\beta) = 0.$$

Since this is the case of an over identification, i.e., $r > k$, we solve the following minimization problem:

$$\begin{aligned}&\min_{\beta} \left[\frac{1}{n}\sum_{i=1}^{n} z_i'(y_i - x_i\beta)\right]' W \left[\frac{1}{n}\sum_{i=1}^{n} z_i'(y_i - x_i\beta)\right] \\ &= \min_{\beta}\, (y - X\beta)' Z W Z' (y - X\beta),\end{aligned}$$

where W is the inverse matrix of the variance-covariance matrix of $Z'(y - X\beta) = Z'u$, which is given by

$$\begin{aligned} Var(Z'u) &= \mathbb{E}[Z'uu'Z] \\ &= Z'\mathbb{E}[uu']Z \\ &= \sigma^2 Z'Z. \end{aligned}$$

Therefore, the problem becomes:

$$\min_{\beta} \ (y - X\beta)'Z(Z'Z)^{-1}Z'(y - X\beta),$$

where we ignore σ^2 since it is a constant and does not affect a solution. We define as $S(\beta) := (y - X\beta)'Z(Z'Z)^{-1}Z'(y - X\beta)$, then the first order condition is:

$$\nabla_\beta S(\tilde{\beta}) = 0.$$

Solving this condition, we obtain the GMM estimator:

$$\tilde{\beta} = \left[X'Z(Z'Z)^{-1}Z'X\right]^{-1} X'Z(Z'Z)^{-1}Z'y.$$

(6) Show that the GMM estimator $\tilde{\beta}$ is consistent and asymptotically normal.

Solution:

The GMM estimator $\tilde{\beta}$ is expressed as

$$\tilde{\beta} = \beta + \left[\frac{1}{n}\sum_{i=1}^n x_i' z_i \left(\frac{1}{n}\sum_{i=1}^n z_i' z_i\right)^{-1} \frac{1}{n}\sum_{i=1}^n z_i' x_i\right]^{-1} \frac{1}{n}\sum_{i=1}^n x_i' z_i \left(\frac{1}{n}\sum_{i=1}^n z_i' z_i\right)^{-1} \frac{1}{n}\sum_{i=1}^n z_i' u_i$$

By the Law of Large Numbers, we have:

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n x_i' z_i &\xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}[x'z] =: M_{xz}, \\ \frac{1}{n}\sum_{i=1}^n z_i' z_i &\xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}[z'z] =: M_{zz}, \\ \frac{1}{n}\sum_{i=1}^n z_i' u_i &\xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}[z'u] = 0. \end{aligned}$$

By the continuous mapping theorem,

$$\tilde{\beta} \xrightarrow[n\to\infty]{\mathbb{P}} \beta + \left(M_{xz} M_{zz}^{-1} M_{xz}'\right)^{-1} M_{xz} M_{zz}^{-1} \cdot 0 = \beta,$$

which concludes that the GMM estimator $\tilde{\beta}$ is a consistent estimator of β. Next, we will show the asymptotic normality of $\tilde{\beta}$. Arranging the expression above, we have:

$$\sqrt{n}\left(\tilde{\beta} - \beta\right) = \left[\frac{1}{n}\sum_{i=1}^n x_i' z_i \left(\frac{1}{n}\sum_{i=1}^n z_i' z_i\right)^{-1} \frac{1}{n}\sum_{i=1}^n z_i' x_i\right]^{-1} \frac{1}{n}\sum_{i=1}^n x_i' z_i \left(\frac{1}{n}\sum_{i=1}^n z_i' z_i\right)^{-1} \frac{1}{\sqrt{n}}\sum_{i=1}^n z_i' u_i.$$

Note that the expectation and the variance of $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i' u_i$ are 0 and $\sigma^2 Z'Z$, respectively. Thus, by the Central Limit Theorem,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i' u_i \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \sigma^2 M_{zz}).$$

Therefore, using the Law of Large Numbers, the continuous mapping theorem, and the Slutsky theorem, we obtain the asymptotic normality of $\tilde{\beta}$:

$$\sqrt{n}\left(\tilde{\beta}-\beta\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \sigma^2 (M_{xz}M_{zz}^{-1}M_{xz}')^{-1}\right),$$

where the asymptotic variance-covariance matrix of $\sqrt{n}\left(\tilde{\beta}-\beta\right)$ is obtained as follows:

$$\begin{aligned} Var\left(\sqrt{n}\left(\tilde{\beta}-\beta\right)\right) &\xrightarrow[n\to\infty]{\mathbb{P}} (M_{xz}M_{zz}^{-1}M_{xz}')^{-1}M_{xz}M_{zz}^{-1}(\sigma^2 M_{zz})M_{zz}^{-1}M_{xz}'(M_{xz}M_{zz}^{-1}M_{xz}')^{-1} \\ &= \sigma^2 (M_{xz}M_{zz}^{-1}M_{xz}')^{-1}. \end{aligned}$$

(7) We need to choose either OLS or GMM. Explain how we choose one of the estimators.

Solution:

To decide which estimators we use, we need to test whether the orthogonality condition, i.e., $\mathbb{E}[Z'u] = 0$ is correct. The null and alternative hypotheses are:

$$\begin{cases} H_0 : \mathbb{E}[Z'u] = 0; \\ H_1 : \mathbb{E}[Z'u] \neq 0. \end{cases}$$

Since the number of equations is r and that of parameter is k, the statistic below asymptotically follows a χ^2 distribution with $r-k$ degrees of freedom.

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i'\widehat{u}_i\right)'\left[\widehat{Var}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i'\widehat{u}_i\right)\right]^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i'\widehat{u}_i\right) \xrightarrow[n\to\infty]{d} \chi^2(r-k),$$

where $\widehat{u}_i = y_i - X\tilde{\beta}$ and $\widehat{Var}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i'\widehat{u}_i\right)$ is the estimator of $Var\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i' u_i\right)$. If we do not reject H_0, we then choose the GMM estimator since the orthogonality condition is likely to be correct.

3 Consider the AR(1) model:

$$y_t = \phi y_{t-1} + \epsilon_t \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2),$$

for $t = 1, 2, \cdots, T$, where $\epsilon_1, \epsilon_2, \cdots, \epsilon_T$ are mutually independent. Moreover, for simplicity, $y_0 = 0$ is assumed. Answer the following questions.

(8) For $|\phi| < 1$, obtain the likelihood function in the innovation form.

Solution:

Using the Bayes' rule, the joint distribution of $y_1, y_2, \cdots, y_T$ is written as

$$\begin{aligned} f(y_1, y_2, \cdots, y_T) &= f(y_T | y_{T-1}, \cdots, y_1) f(y_1, y_2, \cdots, y_{T-1}) \\ &\vdots \\ &= f(y_1) \prod_{t=2}^{T} f(y_t | y_{t-1}, \cdots, y_1), \end{aligned}$$

where $f(y_1)$ denotes an unconditional distribution of y_1 and $f(y_t | y_{t-1}, \cdots, y_1)$ is a conditional one of y_t. Firstly, let us focus on the unconditional distribution. Using the initial condition $y_0 = 0$, we have $y_1 = \epsilon_1$. Since $\epsilon_1 \sim \mathcal{N}(0, \sigma^2)$, we apply the change of variables methods to obtain: [2]

$$f(y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y_1^2}{2\sigma^2}\right).$$

Then, we turn to consider the conditional distribution. Again, by the change of variables method,

$$f(y_t | y_{t-1} \cdot \cdots, y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y_t - \phi y_{t-1})^2}{2\sigma^2}\right\}.$$

Therefore, we obtain the likelihood function:

$$L(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y_1^2}{2\sigma^2}\right) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y_t - \phi y_{t-1})^2}{2\sigma^2}\right\},$$

where $\theta := (\phi, \sigma^2)' \in \mathbb{R}^2$ is the parameter vector.

(9) For $|\phi| < 1$, obtain the variance-covariance matrix of $y = (y_1, y_2, \cdots, y_T)'$. Next, obtain the likelihood function of y, based on the variance-covariance matrix of y.

Solution:

[2] Without the assumption $y_0 = 0$, the unconditional distribution of y_1 would be given by

$$f(y_1) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{1-\phi^2}}} \exp\left\{-\frac{y_1^2}{2\frac{\sigma^2}{1-\phi^2}}\right\}.$$

Firstly, we will derive the variance-covariance matrix of y. The variance of $y_t,\ t \in \{2, \cdots, T\}$, denoted by $\gamma(0)$, is:

$$
\begin{aligned}
\gamma(0) &= Var(y_t) \\
&= Var(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots) \\
&= \sigma^2(1 + \phi^2 + \phi^4 + \cdots) \\
&= \frac{\sigma^2}{1-\phi^2}.
\end{aligned}
$$

Here, notice that we have the initial condition $y_0 = 0$, which implies $y_1 = \epsilon_1$. Thus, for $t = 1$, we have:

$$
Var(y_1) = Var(\epsilon_1) = \sigma^2.
$$

The autocovarinance, denoted by $\gamma(\tau)$ for $\tau = 1, 2, \cdots$, is given by

$$
\begin{aligned}
\gamma(\tau) &= \mathbb{E}[(y_t - \mu)(y_{t-\tau} - \mu)] \\
&= \mathbb{E}[y_t y_{t-\tau}] \\
&= \mathbb{E}[(\phi^\tau y_{t-\tau} + \epsilon_t + \phi\epsilon_{t-1} + \cdots + \phi^{\tau-1}\epsilon_{t-\tau+1})y_{t-\tau}] \\
&= \phi^\tau\gamma(0) \\
&= \frac{\sigma^2\phi^\tau}{1-\phi^2},
\end{aligned}
$$

where μ denotes the mean of y_t and $\mu = 0$ for all t. Therefore, the variance-covariance matrix of y is:

$$
\Sigma := Var(y) = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} 1-\phi^2 & \phi & \phi^2 & \cdots & \phi^{T-1} \\ \phi & 1 & \phi & \cdots & \phi^{T-2} \\ \phi^2 & \phi & 1 & \cdots & \phi^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{T-1} & \phi^{T-2} & \phi^{T-3} & \cdots & 1 \end{pmatrix}
$$

Using this matrix, we can define the likelihood function as follows:

$$
L(\theta) = \frac{1}{(2\pi)^{\frac{T}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}y'\Sigma^{-1}y\right),
$$

where Σ denotes the variance-covariance matrix of y derived above.

(10) For $|\phi| < 1$, show the following equality:

$$
\begin{pmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{T-1} \\ \phi & 1 & \ddots & \ddots & \vdots \\ \phi^2 & \ddots & \ddots & \ddots & \phi^2 \\ \vdots & \ddots & \ddots & \ddots & \phi \\ \phi^{T-1} & \cdots & \phi^2 & \phi & 1 \end{pmatrix} = \left(\begin{pmatrix} \sqrt{1-\phi^2} & & & & 0 \\ 1 & -\phi & & & \\ & 1 & -\phi & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & -\phi \end{pmatrix}'\begin{pmatrix} \sqrt{1-\phi^2} & & & & 0 \\ 1 & -\phi & & & \\ & 1 & -\phi & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & -\phi \end{pmatrix}\right)^{-1}
$$

Solution:

The left hand side can be transformed as follows:

$$\begin{pmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{T-1} \\ \phi & 1 & \ddots & \ddots & \vdots \\ \phi^2 & \ddots & \ddots & \ddots & \phi^2 \\ \vdots & \ddots & \ddots & \ddots & \phi \\ \phi^{T-1} & \cdots & \phi^2 & \phi & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\phi & 0 & \cdots & 0 \\ -\phi & 1+\phi^2 & -\phi & \ddots & \vdots \\ 0 & -\phi & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\phi \\ 0 & \cdots & 0 & -\phi & 1 \end{pmatrix}^{-1},$$

which is equal to the right hand side.

(11) For $\phi = 1$, derive the autocovariance between y_t and $y_{t-\tau}$.

Solution:

When $\phi = 1$, this is the case of a random walk process. Then, we have

$$\begin{aligned} y_t &= y_{t-1} + \epsilon_t \\ &= \epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_1. \end{aligned}$$

Thus, the autocovariance between y_t and $y_{t-\tau}$, denoted by $\gamma(\tau)$, is:

$$\begin{aligned} \gamma(\tau) &= \mathbb{E}[(y_t - \mu)(y_{t-\tau} - \mu)] \\ &= \mathbb{E}[y_t y_{t-\tau}] \\ &= \mathbb{E}[(\epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_{t-\tau} + \epsilon_{t-\tau-1} + \cdots + \epsilon_1)(\epsilon_{t-\tau} + \epsilon_{t-\tau-1} + \cdots + \epsilon_1)] \\ &= \mathbb{E}[\epsilon_{t-\tau}^2] + \mathbb{E}[\epsilon_{t-\tau-1}^2] + \cdots + \mathbb{E}[\epsilon_1^2] \\ &= \sigma^2(t-\tau). \end{aligned}$$

(12) For $\phi = 1$, derive the asymptotic distribution of $T(\widehat{\phi} - 1)$.

Solution:

The OLS estimator of the model $y_t = \phi y_{t-1} + \epsilon_t$ is given by

$$\widehat{\phi} = \phi + \frac{\sum_{t=1}^T y_{t-1}\epsilon_t}{\sum_{t=1}^T y_{t-1}^2},$$

which is arranged as follows:

$$\begin{aligned} (\widehat{\phi} - \phi) &= \frac{\frac{1}{T}\sum_{t=1}^T y_{t-1}\epsilon_t}{\frac{1}{T}\sum_{t=1}^T y_{t-1}^2} \\ \iff \quad T(\widehat{\phi} - \phi) &= \frac{\frac{1}{T}\sum_{t=1}^T y_{t-1}\epsilon_t}{\frac{1}{T^2}\sum_{t=1}^T y_{t-1}^2}. \end{aligned}$$

We will derive the asymptotic distribution of the numerator $\frac{1}{T}\sum_{t=1}^T y_{t-1}\epsilon_t$ and the denominator $\frac{1}{T^2}\sum_{t=1}^T y_{t-1}^2$.

(a) First, let us consider the numerator. Since $y_t = \phi y_{t-1} + \epsilon_t$ with $\phi = 1$, we have:

$$\begin{aligned} y_t^2 &= (y_{t-1} + \epsilon_t)^2 \\ \iff \quad y_{t-1}\epsilon_t &= \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2). \end{aligned}$$

Taking into account $y_0 = 0$,

$$\begin{aligned} \sum_{t=1}^{T} y_{t-1}\epsilon_t &= \frac{1}{2}\sum_{t=1}^{T}(y_t^2 - y_{t-1}^2 - \epsilon_t^2) \\ &= \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2. \end{aligned}$$

Divided by $\sigma^2 T$ on both sides, we have:

$$\frac{1}{\sigma^2}\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma\sqrt{T}}\right)^2 - \frac{1}{2\sigma^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2.$$

Since $y_t \sim \mathcal{N}(0, \sigma^2 t)$, we obtain: [3]

$$\left(\frac{y_T}{\sigma\sqrt{T}}\right)^2 \sim \chi^2(1).$$

Moreover, by the ergodicity, we have:

$$\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2 \xrightarrow[T\to\infty]{\mathbb{P}} \mathbb{E}[\epsilon_t^2] = \sigma^2.$$

Therefore, by the continuous mapping theorem and the Slutsky theorem, we have the asymptotic distribution of the numerator:

$$\frac{1}{\sigma^2}\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t \xrightarrow[T\to\infty]{d} \frac{1}{2}\left(\chi^2(1) - 1\right).$$

(b) Second, we will derive the asymptotic distribution of the denominator $\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2$. We define $X_T(r)$ as

$$X_T(r) = \begin{cases} 0 & 0 \le r < \frac{1}{T}, \\ \frac{\epsilon_1}{T} & \frac{1}{T} \le r < \frac{2}{T}, \\ \frac{\epsilon_1+\epsilon_2}{T} & \frac{2}{T} \le r < \frac{3}{T}, \\ \vdots & \vdots \\ \frac{\epsilon_1+\cdots+\epsilon_T}{T} & r = 1. \end{cases}$$

[3] In case of $\phi = 1$, the expectation of y_t is zero and the variance of it is given by

$$Var(y_t) = Var(\epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_1) = \sigma^2 t.$$

Then, $y_t \sim \mathcal{N}(0, \sigma^2 t)$.

Let $[Tr]$ be the largest integer which is less than or equal to $T \times r$. For instance, if $r = \frac{2.8}{T}$, then $[Tr] = [2.8] = 2$. Using this operator, we can express $X_T(r)$ as follows:

$$X_T(r) = \frac{1}{T} \sum_{t=1}^{[Tr]} \epsilon_t$$
$$\iff \quad \sqrt{T} X_T(r) = \frac{[Tr]}{T} \sqrt{\frac{T}{[Tr]}} \frac{1}{\sqrt{[Tr]}} \sum_{t=1}^{[Tr]} \epsilon_t.$$

For each part of them, we have

$$\frac{[Tr]}{T} \xrightarrow[T \to \infty]{} r,$$
$$\sqrt{\frac{T}{[Tr]}} \xrightarrow[T \to \infty]{} \frac{1}{\sqrt{r}},$$
$$\frac{1}{\sqrt{[Tr]}} \sum_{t=1}^{[Tr]} \epsilon_t \xrightarrow[T \to \infty]{d} \mathcal{N}(0, \sigma^2),$$

where we use the Central Limit Theorem for the third one. Therefore, by the Slutsky theorem, we have:

$$\sqrt{T} X_T(r) \xrightarrow[T \to \infty]{d} \mathcal{N}(0, r\sigma^2) = \sigma W(r), \tag{1}$$

where $W(r)$ denotes a standard Brownian motion.
Since $y_t = \epsilon_t + \cdots + \epsilon_1$, $X_T(r)$ can be expressed as follows:

$$X_T(r) = \begin{cases} 0 & 0 \le r < \frac{1}{T}, \\ \frac{y_1}{T} & \frac{1}{T} \le r < \frac{2}{T}, \\ \frac{y_2}{T} & \frac{2}{T} \le r < \frac{3}{T}, \\ \vdots & \vdots \\ \frac{y_{T-1}}{T} & \frac{T-1}{T} \le r < 1, \\ \frac{y_T}{T} & r = 1. \end{cases}$$

In addition, we define $S_T(r)$ as follows:

$$S_T(r) = \begin{cases} 0 & 0 \le r < \frac{1}{T}, \\ \frac{y_1^2}{T} & \frac{1}{T} \le r < \frac{2}{T}, \\ \frac{y_2^2}{T} & \frac{2}{T} \le r < \frac{3}{T}, \\ \vdots & \vdots \\ \frac{y_{T-1}^2}{T} & \frac{T-1}{T} \le r < 1, \\ \frac{y_T^2}{T} & r = 1. \end{cases}$$

To obtain $\int_0^1 X_T(r)dr$ and $\int_0^1 S_T(r)dr$, we compute a sum of rectangulars as follows:

$$\begin{aligned} \int_0^1 X_T(r)dr &\simeq \frac{y_1}{T}\left(\frac{2}{T} - \frac{1}{T}\right) + \cdots + \frac{y_{T-1}}{T}\left(\frac{T}{T} - \frac{T-1}{T}\right) \\ &= \frac{y_1}{T^2} + \cdots + \frac{y_{T-1}}{T^2} \\ &\simeq \frac{1}{T^2} \sum_{t=1}^{T} y_t, \end{aligned}$$

and similarly,

$$\int_0^1 S_T(r)dr \simeq \frac{1}{T^2}\sum_{t=1}^{T} y_t^2 \simeq \frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2.$$

Using equation (1) and the continuous mapping theorem,

$$\int_0^1 \sqrt{T}X_T(r)dr \xrightarrow[T\to\infty]{d} \sigma \int_0^1 W(r)dr.$$

Since $S_T(r) = \left(\sqrt{T}X_T(r)\right)^2$, using the continuous mapping theorem, we obtain:

$$\begin{aligned} S_T(r) &\xrightarrow[T\to\infty]{d} \sigma^2(W(r))^2 \\ \iff \quad \int_0^1 S_T(r)dr &\xrightarrow[T\to\infty]{d} \sigma^2 \int_0^1 (W(r))^2 dr, \end{aligned}$$

which implies

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

Thus, we have obtained the asymptotic distribution of the denominator.

Therefore, from the discussion of (a) and (b), and using $\phi = 1$, the continuous mapping theorem and the Slutsky theorem, we have the asymptotic distribution of $T(\widehat{\phi}-1)$ as follows:

$$T(\widehat{\phi}-1) = \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2} \xrightarrow[T\to\infty]{d} \frac{\frac{1}{2}\left[(W(r))^2 - 1\right]}{\int_0^1 (W(r))^2 dr}.$$

(13) We estimate $\Delta y_t = \rho y_{t-1} + \epsilon_t$ using the conventional OLS command in econometric softwares. The following results are obtained.

$$\begin{aligned} \Delta y_t = -0.09\ & y_{t-1} \\ (1.90)& \end{aligned}$$

where the value in the parenthesis represents the t-value.

We want to test whether y_t has a unit root. Show the null hypothesis and the alternative one. Test whether y_t has a unit root at 5 % significance level, where $T = 1,000$.

Solution:

We use the augmented Dickey-Fuller test. Consider the model:

$$y_t = \phi y_{t-1} + \epsilon_t.$$

If the model has a unit root, then

$$\begin{aligned} y_t - y_{t-1} &= \epsilon_t \\ \iff \quad \Delta y_t &= \epsilon_t, \end{aligned}$$

where Δ is a difference operator. This implies that y_{t-1} does not have any effects on Δy_t when $\phi = 1$. Therefore, we consider the model:

$$\Delta y_t = \rho y_{t-1} + \epsilon_t,$$

and test whether $\rho = 0$ or not. Then, the null hypothesis and the alternative one are:

$$\begin{cases} H_0 : \rho = 0; \\ H_1 : \rho < 0. \end{cases}$$

The test statistic is given by

$$t = \frac{\widehat{\rho}}{SE(\widehat{\rho})},$$

where $SE(\widehat{\rho})$ denotes the standard error of $\widehat{\rho}$. Referring to the t statistic table of the Dickey-Fuller test, the rejecting area is $\{t : t < -2.86\}$ when $T = 1,000$ and the significance level is 5%. Since the t value is 1.90, we do not reject the null hypothesis that y_t has a unit root at 5% significance level.

4 When y_i^* is unobservable and y_i is observed, consider the following model:

$$y_i^* = x_i\beta + u_i$$
$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

where x_i is not correlated with u_i for $i = 1, 2, \cdots, n$. We assume that u_i for $i = 1, 2, \cdots, n$ are mutually independently and normally distributed as $u_i \sim \mathcal{N}(0, \sigma^2)$.
Answer the following questions.

(14) What is the probability that y_i^* is greater than zero? What is the probability that y_i^* is less than or equal to zero?

Solution:

The probability that y_i^* is greater than zero is:

$$\begin{aligned} \mathbb{P}(y_i^* > 0) &= \mathbb{P}(x_i\beta + u_i > 0) \\ &= \mathbb{P}(u_i > -x_i\beta) \\ &= \mathbb{P}(u_i^* > -x_i\beta^*) \\ &= 1 - F(-x_i\beta^*) \\ &= F(x_i\beta^*), \end{aligned}$$

where $u_i^* := u_i/\sigma$ and $\beta^* := \beta/\sigma$. $F(\cdot)$ denotes the cumulative distribution function of u_i^*, which is given by

$$F(x_i\beta^*) = \int_{-\infty}^{x_i\beta^*} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz,$$

since $u_i \sim \mathcal{N}(0, \sigma^2)$. The last equality holds because of the symmetricity of $F(\cdot)$. The probability that y_i^* is less than or equal to zero is:

$$\mathbb{P}(y_i^* \leq 0) = 1 - \mathbb{P}(y_i^* > 0) = 1 - F(x_i\beta^*).$$

(15) Derive the joint distribution of $y_1, y_2, \cdots, y_n$.

Solution:

Since y_i is a binary random variable, y_i is considered to follow the Bernoulli distribution:

$$\begin{aligned} f(y_i) &= [\mathbb{P}(y_i = 1)]^{y_i} [1 - \mathbb{P}(y_i = 1)]^{1-y_i} \\ &= [\mathbb{P}(y_i^* > 0)]^{y_i} [1 - \mathbb{P}(y_i^* > 0)]^{1-y_i} \\ &= [F(x_i\beta^*)]^{y_i} [1 - F(x_i\beta^*)]^{1-y_i}. \end{aligned}$$

We assume that $y_1, y_2, \cdots, y_n$ are mutually independent. Then, the joint distribution of them is given by

$$\begin{aligned} f(y_1, y_2, \cdots, y_n) &= \prod_{i=1}^{n} f(y_i) \\ &= \prod_{i=1}^{n} [F(x_i\beta^*)]^{y_i} [1 - F(x_i\beta^*)]^{1-y_i}. \end{aligned}$$

(16) We want to test the null hypothesis $H_0 : \beta_j = 0$ and the alternative one $H_1 : \beta_j \neq 0$, where β_j denotes the jth element of β. We can consider several testing procedures. Explain one of them.

Solution:

The null hypothesis and the alternative one are:

$$\begin{cases} H_0 : \beta_j = 0; \\ H_1 : \beta_j \neq 0. \end{cases}$$

The t statistic is given by

$$t = \frac{\widehat{\beta}_j}{SE(\widehat{\beta}_j)},$$

where $SE(\widehat{\beta}_j)$ denotes the standard error of $\widehat{\beta}_j$. The rejecting area is given by $\{t : |t| > 1.96\}$ when the significance level is 5%. Therefore, if the realized t value is in the rejecting area, we reject H_0 and conclude that jth element of the explanatory variables has an effect on y_i.