「Econometrics II」

Homework # 2 (Due: January 20, 2026, AM10:20)

- The answer should be written in English or Japanese.
- Your name and student ID number should be included in your answer sheet.
- Bring your answer sheet in this classroom and hand it to me.
- DO NOT send me the answer sheet by email.

$\boxed{1}$ Consider the following regression model:

$$y_i = x_i\beta + u_i$$

for $i = 1, 2, \cdots, n$. x_i and β are $1 \times k$ and $k \times 1$ vectors. u_i is mutually independently distributed with mean zero and variance σ^2.

(1) When x_i is correlated with u_i, show that the OLS estimator $\hat{\beta}$ is inconsistent.

(2) When x_i is not correlated with u_i, show that $\sqrt{n}(\hat{\beta} - \beta) \longrightarrow N(0, \sigma^2 M_{xx}^{-1})$, where the OLS estimator is $\hat{\beta}$, and $M_{xx} = \lim_{n\to\infty} \dfrac{1}{n} \sum_{i=1}^{n} x_i' x_i$.

(3) Suppose that we have another $1 \times k$ variable z_i, which is not correlated with u_i. Using z_i, construct a consistent estimator of β, denoted by $\tilde{\beta}$.

(4) Show that $\sqrt{n}(\tilde{\beta} - \beta) \longrightarrow N(0, \sigma^2 M)$. Obtain M, utilizing the followings:

$$M_{zx} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} z_i' x_i \qquad M_{zz} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} z_i' z_i$$

(5) We consider testing whether x_i is correlated with u_i. Explain the testing procedure for choice of either $\hat{\beta}$ or $\tilde{\beta}$.

$\boxed{2}$ Consider the following regression:

$$y_{it} = x_{it}\beta + v_i + u_{it} \qquad i = 1, 2, \cdots, n \text{ and } t = 1, 2, \cdots, T$$

where v_i is the individual effect and the error term u_{it} is mutually independently distributed as $u_{it} \sim N(0, \sigma_u^2)$.

(1) Suppose that $v_i \sim N(0, \sigma_v^2)$ is independently distributed with each other. When x_{it} is not correlated with v_i and u_{it}, construct the likelihood function of y_{it} for $i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T$. Obtain the maximum likelihood estimator of β, denoted by $\tilde{\beta}$.

(2) When v_i is correlated with x_{it}, we want to obtain a consistent estimator of β, which is denoted by $\hat{\beta}$. Derive $\hat{\beta}$.

(3) We need to choose $\tilde{\beta}$ or $\hat{\beta}$. Which estimator should be chosen? Explain the procedure.

$\boxed{3}$ Consider the following regression:

$$y_i = x_i\beta + u_i \qquad i = 1, 2, \cdots, n$$

where the error term u_i is correlated with x_i. Assume that u_i is independent of u_j for $i \neq j$.

(1) Consider another regression model: $x_i = z_i\Gamma + v_i$, where x_i, z_i, Γ and v_i are $1 \times k$, $1 \times p$, $p \times k$ and $1 \times k$ vectors or matrices. Suppose that z_i is not correlated with u_i and v_i. Estimate Γ by OLS, which is denoted by $\hat{\Gamma}$. Utilizing $\hat{\Gamma}$, obtain a consistent estimator of β.

(2) Derive the asymptotic distribution of the consistent estimator of β given by (1).

$\boxed{4}$ We want to know an average effect of new educational system. Carry out the new educational system to some students, which are chosen at random, and the traditional one to the other students. The outcome of the new system is denoted by Y^1 while that of the traditional one is represented by Y^0. For the ith individual, we can observe one of the two outcomes, which are denoted by Y_i^1 and Y_i^0. That is, Y_i^1 is observed and Y_i^0 is unobserved, if the ith student takes the new educational system.

Let D_i be the dummy variable which takes 1 when the ith individual take the new educational system and 0 otherwise. X_i represents all the possible information (i.e., explanatory variables) of the ith student. $\pi(X_i)$ denotes the conditional probability of $D_i = 1$ given X_i, which is called the propensity score.

Then, we have the following theorems:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{D_iY_i}{\pi(X_i)} - \frac{(1-D_i)Y_i}{1-\pi(X_i)}\right) \ \xrightarrow{p} \ \mu \tag{1}$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\left(\frac{D_iY_i}{\pi(X_i)} - \frac{(1-D_i)Y_i}{1-\pi(X_i)}\right) - \mu\right) \ \xrightarrow{d} \ N(0,\ \sigma^2) \tag{2}$$

where $\mu = E(Y^1 - Y^0)$ and $\sigma^2 = V(Y^1 - Y^0)$.

(1) We have to assume:

    (i) D_i, Y_i and X_i are identically distributed, and

    (ii) D_i is independent of Y_i, given X_i, for all i.

Showing where we use these assumptions, prove Equation (1).

(2) To construct the confidence interval of μ or test whether μ is zero (i.e., no effect of the new education system), practically, Equation (2) is utilized as:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{D_iY_i}{\pi(X_i)} - \frac{(1-D_i)Y_i}{1-\pi(X_i)}\right) \ \sim \ N\left(\mu,\ \frac{\hat{\sigma}^2}{n}\right)$$

for large n, where $\hat{\sigma}^2$ denotes the estimate of σ^2.

Explain how to obtain $\hat{\sigma}^2$.