「Econometrics II」 Homework No.1

Send your answer to tanizaki [at] econ.osaka-u.ac.jp by December 20, 2021.

1 Suppose that y_i takes 1, 2, $\cdots$, n for $i = 1, 2, \cdots, T$, which is a discrete random variable. X_i is a vector of exogenous variables. n is assumed to be known.

(1) We assume that y_i is distributed as a binomial random variable with parameter p_i for $i = 1, 2, \cdots, T$. Consider estimating β with $p_i = F(X_i\beta)$, where $F(\cdot)$ has to be specified by a researcher. Representatively, a standard normal distribution function or a logistic distribution function is taken for $F(\cdot)$. Construct the likelihood function.

Note that the binomial distribution is given by:

$$f(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

(2) Obtain the first order condition for the maximum likelihood estimator of β.

2 Suppose that y_i takes one when the ith person says Yes in a questionnaire, while it takes zero when the ith person says No. X_i is a vector of exogenous variables. Using OLS, we consider estimating the following linear regression model:

$$y_i = X_i\beta + u_i,$$

for $i = 1, 2, \cdots, n$.

(3) Focus on the left hand side, i.e., y_i. Show that $P(y_i = 1) = \mathrm{E}(y_i)$, where $P(y_i = 1)$ denotes the probability that y_i takes one.

(4) Focus on the right hand side, i.e., $X_i\beta + u_i$. Consider $\mathrm{E}(X_i\beta + u_i)$. From the regression model, $\mathrm{E}(y_i) = \mathrm{E}(X_i\beta + u_i)$ should hold. However, you feel something strange. What is the problem of the above regression model?

(5) Explain how we should estimate the regression model when we have the binary dependent variable.

3 Consider that n random variables X_1, X_2, $\cdots$, X_n are mutually independently and identically distributed as the density function $f(x;\theta)$, where θ is an unknown parameter to be estimated. For simplicity, θ is a scalar. Let $s(X)$ be an unbiased estimator of θ for $X=(X_1, X_2, \cdots, X_n)$.

(6) Suppose that $L(\theta;x)=\prod_{i=1}^{n} f(x_i;\theta)$, which is called the likelihood function. Show the following two equalities:

$$\mathrm{E}\left(\frac{\partial \log L(\theta;X)}{\partial \theta}\right)=0.$$

$$-\mathrm{E}\left(\frac{\partial^2 \log L(\theta;X)}{\partial \theta^2}\right)=\mathrm{E}\left(\left(\frac{\partial \log L(\theta;X)}{\partial \theta}\right)^2\right)=\mathrm{V}\left(\frac{\partial \log L(\theta;X)}{\partial \theta}\right).$$

(7) Show that the following inequality holds.

$$\mathrm{V}(s(X)) \geq (I(\theta))^{-1},$$

where $I(\theta)=-\mathrm{E}\left(\frac{\partial^2 \log L(\theta;X)}{\partial \theta^2}\right)$, which is called Fisher's information matrix.

(8) Suppose that $s(X)=\tilde{\theta}$ is a maximum likelihood estimator. Then, show that

$$\sqrt{n}(\tilde{\theta}-\theta) \longrightarrow N(0,\sigma^2),$$

where $\sigma^2=\lim_{n\to\infty}\left(\frac{I(\theta)}{n}\right)^{-1}$.