Econometrics 2 (2018) TA session 3*

Kenji Hatakenaka †

18 October 2018

Aim(目的)

- Review the previous lectures with some additional contents.
- Introduce an empirical example of probit.

2 Qualitative Variable or Quantitative Variable

In most cases, dependent variable is continuous or assumed to be continuous. For example,

- temperature
- individual income

$\rightarrow y_i$ is continuous and $-\infty < y_i < \infty$.
However, we may encounter some variables which takes only several value.
For example,

- male or female
- smoking or not smoking

$\rightarrow y_i = 0$ or 1.

---

* All comments welcome!
† E-mail: u626530i@ecs.osaka-u.ac.jp, Room 501

In this section, we learn the case these variables are dependent variable.

2.1 Discrete Choice Model(離散選択モデル)

2.1.1 Binary Choice Model

This model is represented by

$$y_i^* = X_i\beta + u_i \quad u_i \sim (0, \sigma^2), \quad i = 1, 2, \ldots n. \tag{1}$$

where y^* is unobserved, but y_i is observed as 0 or 1,

$$y_i = \begin{cases} 1, & \text{if} \quad y_i^* > 0, \\ 0, & \text{if} \quad y_i^* \leq 0. \end{cases} \tag{2}$$

Consider the probability that y_i takes 1, i.e.

$$\begin{aligned} P(y_i = 1) &= P(y_i^* > 0) = P(u_i > -X_i\beta) \\ &= P\left(\frac{u_i}{\sigma} > -X_i\frac{\beta}{\sigma}\right) = P(u_i^* > -X_i\beta^*) \\ &= 1 - P\left(u_i^* \leq -X_i\beta^*\right) = 1 - F(-X_i\beta^*) = F(X_i\beta^*) \end{aligned} \tag{3}$$

where u_i^* and $\beta^*$are defined as

$$u_i^* = \frac{u_i}{\sigma}, \quad \beta^* = \frac{\beta}{\sigma} \tag{4}$$

The last equality of (3) comes from the **symmetry of distribution** u_i^*. Note that we estimate β^*, but we cannot estimate β and σ separately. The distribution function is defined as follows:

$$F(x) = \int_{-\infty}^{x} f(z)\mathrm{d}z \tag{5}$$

2.1.2 Difference between Probit and Logit

Here, we shortly review the difference between probit and logit. The difference is below:

- Probit $\to u_i^*$ is standard normal distribution, i.e., $u_i^* \sim N(0,1)$

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) \mathrm{d}z, \quad f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right). \tag{6}$$

- Logit $\to u_i^*$ is logistic distribution,

$$F(x) = \frac{1}{1+\exp(-x)}, \quad f(x) = \frac{\exp(-x)}{(1+\exp(-x))^2} \tag{7}$$

We can also consider the other distribution function for u_i^*. For example, we sometime encounter Gumbel distribution.

2.2 Likelihood Function

y_i is the following Bernouli distribution $f(y_i)$ as follows:

$$f(y_i) = (P(y_i = 1))^{y_i}(1 - P(y_i = 0))^{1-y_i}, \quad y_i = 0, 1, \tag{8}$$

$$= (F(X_i\beta^*))^{y_i}(1 - F(X_i\beta^*))^{1-y_i} \tag{9}$$

Then we obtain the likelihood function as follows:

$$L(\beta^*) = f(y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} (F(X_i\beta^*))^{y_i}(1 - F(X_i\beta^*))^{1-y_i}. \tag{10}$$

Then log-likelihood function is :

$$\log L(\beta^*) = \sum_{i=1}^{n} \{y_i \log F(X_i\beta^*) + (1 - y_i)\log(1 - F(X_i\beta^*))\} \tag{11}$$

Solving the maximization problem of $\log L(\beta^*)$ with respect to β^*, the FOC is:

$$\frac{\partial L(\beta^*)}{\partial \beta^*} = \sum_{i=1}^{n} \left(\frac{y_i X_i' f(X_i\beta^*)}{F(X_i\beta^*)} - \frac{(1 - y_i) X_i' f(X_i\beta^*)}{1 - F(X_i\beta^*)} \right) \tag{12}$$

$$= \sum_{i=1}^{n} \frac{X_i' f(X_i\beta^*)(y_i - F(X_i\beta^*))}{F(X_i\beta^*)(1 - F(X_i\beta^*))} = 0 \tag{13}$$

SOC is

$$\frac{\partial^2 \log L(\beta^*)}{\partial\beta^*\partial\beta^{*\prime}} = \sum_{i=1}^{n} \frac{X_i' X_i f_i'(y_i - F_i)}{F_i(1-F_i)} - \sum_{i=1}^{n} \frac{X_i' X_i f_i^2}{F_i(1-F_i)} \tag{14}$$

$$- \sum_{i=1}^{n} X_i' f_i (y_i - F_i) \frac{X_i f_i (1-2F_i)}{(F_i(1-F_i))^2} \tag{15}$$

(15) is negative definite. When we adopt Logit model, we can calculate further.

$$\frac{\partial L(\beta^*)}{\partial\beta^*} = \sum_{i=1}^{n} \frac{X_i' f(X_i\beta^*)(y_i - F(X_i\beta^*))}{F(X_i\beta^*)(1-F(X_i\beta^*))} \tag{16}$$

$$= \sum_{i=1}^{n} \frac{X_i' \frac{\exp(-X_i\beta^*)}{(1+\exp(-X_i\beta^*))^2}\left(y_i - \frac{1}{1+\exp(-X_i\beta^*)}\right)}{\frac{1}{1+\exp(-X_i\beta^*)}\frac{\exp(-X_i\beta)}{1+\exp(-X_i\beta^*)}} \tag{17}$$

$$= \sum_{i=1}^{n} X_i' \left(y_i - \frac{1}{\exp(-X_i\beta^*)}\right) = 0. \tag{18}$$

For maximization, the method of scoring is given by

$$\beta^{*(j+1)} = \beta^{*(j)} + \left(\sum_{i=1}^{n} \frac{X_i' X_i (f_i^{(j)})^2}{F_i^{(j)}(1-F_i^{(j)})}\right)^{-1} \sum_{i=1}^{n} \frac{X_i' f_i^{(j)}(y_i - F_i^{(j)})}{F_i^{(j)}(1-F_i^{(j)})}. \tag{19}$$

Variance of MLE $\hat{\beta}^*$ is $I(\hat{\beta}^*)^{-1}$ where

$$I(\hat{\beta}^*) = -E\left[\frac{\partial^2 \log L(\hat{\beta}^*)}{\partial\beta^*\partial\beta^{*\prime}}\right] = \sum_{i=1}^{n} \frac{X_i' X_i \hat{f}_i^2}{\hat{F}_i(1-\hat{F}_i)}. \tag{20}$$

We can estimate $\hat{\beta}^*$ and test the significance of $\hat{\beta}^*$.

2.3 Another Interpretation

This maximization problem is equivalent to the nonlinear least squares estimation problem from the following regression model:

$$y_i = F(X_i\beta^*) + u_i \tag{21}$$

where

$$u_i = y_i - F_i = \begin{cases} 1 - F_i & \text{w.p. } P(y_i = 1) \\ 0 - F_i & \text{w.p. } P(y_i = 0) \end{cases} \tag{22}$$

Therefore, the mean and variance of u_i are:

$$E(u_i) = \underbrace{(1 - F_i)}_{\text{value}} \underbrace{F_i}_{\text{prob.}} + \underbrace{(-F_i)}_{\text{value}} \underbrace{(1 - F_i)}_{\text{prob.}} = 0 \tag{23}$$

$$\sigma_i^2 = V(u_i) = E[u_i^2 - E(u_i)^2] = E(u_i^2) \tag{24}$$

$$= \underbrace{(1 - F_i)^2}_{\text{value}} \underbrace{F_i}_{\text{prob.}} + \underbrace{(-F_i)^2}_{\text{value}} \underbrace{(1 - F_i)}_{\text{prob.}} = F_i(1 - F_i). \tag{25}$$

Then the weighted least squares method solves the following minimization problem:

$$\min_{\beta^*} \quad \sum_{i=1}^{n} \frac{(y_i - F(X_i\beta^*))^2}{\sigma_i^2} \tag{26}$$

Then, the first order condition becomes:

$$\sum_{i=1}^{n} \frac{2X_i' f(X_i\beta^*)(y_i - F(X_i\beta^*))}{\sigma_i^2} = 0 \tag{27}$$

This is equivalent to the first order condition of MLE.

2.4 example1 and 2

2.4.1 Random Utility Model

$$U_{1,i} = X_i\beta_{1,i} + \epsilon_{1,i}, \quad U_{2,i} = X_i\beta_{2,i} + \epsilon_{2,i} \tag{28}$$

We can observe

$$y_i = \begin{cases} 1 & \text{if } U_{1,i} > U_{2,i} \\ 0 & \text{if } U_{1,i} \leq U_{2.i} \end{cases} \tag{29}$$

The latent equation is

$$y_i^* = X_i\beta + \epsilon_i \tag{30}$$

where

$$y_i^* \equiv U_{1,i} - U_{2,i} \tag{31}$$
$$\beta \equiv \beta_1 - \beta_2 \tag{32}$$

2.4.2 questionnaire(アンケート)

$$y_i = F(X_i\beta) + \epsilon_i \tag{33}$$
$$y_i = \begin{cases} 1 & \text{yes} \\ 0 & \text{no} \end{cases} \tag{34}$$

$F(\cdot)$ is a distribution function of ϵ_i(ex. normal dist, logistic dist) where

$$E[y_i] = P(y_i = 1) = F(X_i\beta) \tag{35}$$

Likelihood function is

$$L(\beta) = \prod_{i=1}^{n} f(y_i) \tag{36}$$
$$= \prod_{i=1}^{n} F_i^{y_i}(1 - F_i)^{1-y_i} \tag{37}$$

where $f(y_i)$ is (35). $\rightarrow$ MLE

2.4.3 Ordered Probit or Logit Model

Ordered Probit or Logit Model is the case y_i is observed as $1, 2, \ldots, m$.

$$y_i = \begin{cases} 1 & y_i^* < a_1 \\ 2 & a_1 \leq y_i^* < a_2 \\ \vdots & \\ m & a_{m-1} \leq y_i^* \end{cases} \tag{38}$$

Probability density function of y_i is

$$f(y_i) = (P(y_i = 1))^{I_{\{i=1\}}}(P(y_i = 2))^{I_{\{i=2\}}} \cdots (P(y_i = m))^{I_{\{i=m\}}},$$

where

$$\begin{aligned} P(y_i = j) &= P(a_{j-1} \leq y_i^* < a_j) \\ &= P(y_i^* < a_j) - P(y_i^* < a_{j-1}) \\ &= P(u_i < a_j - X_i\beta) - P(u_i < a_{j-1} - X_i\beta) \\ &= F(a_j - X_i\beta) - F(a_{j-1} - X_i\beta) \end{aligned}$$

$$I_{\{i=j\}} = \begin{cases} 1 & \text{if } y_i = j \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, 2, \cdots, m$. $a_0 = -\infty, a_m = \infty$.
So Likelihood function:

$$L(\beta) = \prod_{i=1}^{n} f(y_i) \tag{39}$$

can be constructed. $\rightarrow$ MLE

2.4.4 Multinomial logit model

This model is unorderd choice model(ex. $y_i = 1(\text{menial}), 2(\text{blue collar}), 3(\text{white collar}, 4(\text{professional}))$. The individual has $m + 1$ choices, i.e. $j = 0, 1, 2, \cdots, m$

$$P(y_i = j) = \frac{\exp(X_i\beta_j)}{\sum_{j=0}^{m} \exp(X_i\beta_j)} = P_{ij} \tag{40}$$

for $\beta_0 = 0$(The case of $m = 1$ correspond to bivariate logit model).
Note that

$$\log\frac{P_{ij}}{P_{i0}} = X_i\beta_j. \tag{41}$$

The log likelihood function is

$$\log L(\beta_1, \beta_2, \cdots, \beta_m) = \sum_{i=1}^{n}\sum_{j=0}^{m} I_{\{y_i=j\}} \log P_{ij} \tag{42}$$

2.5 Marginal effect("When x_i increase by 1%, how much y_i would increase?")

When we employ OLS method ($y = x_1\beta_1 + \cdots + x_k\beta_k + \epsilon$), marginal effect is "$\beta_i$%". When we conduct probit or logit estimation, the result is not straightforward. The model is represented as follows:

$$y_i^* = P(y_i = 1) + u_i = F(X_i\beta^*) + u_i. \tag{43}$$

By differentiating this equation by X_{ik}, kth independent variable of individual i, we obtain

$$\frac{\mathrm{d}P(y_i = 1)}{\mathrm{d}X_{ik}} = \frac{\mathrm{d}F(X_i\beta^*)}{\mathrm{d}X_{ik}} = \beta_k^* f(X_i\beta^*). \tag{44}$$

Where $f(X_i\beta^*)$ is probability density function of $X_i\beta$.