Homework (Due: December 19, 2019, AM10:20)

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

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

where X_i is assumed to be exogenous and nonstochastic, and u_1, u_2, $\cdots$, u_n are mutually independent errors.

Let $f(x)$ be the density function of u_i and $F(x)$ be the cumulative distribution function of u_i, i.e., $F(x) = \int_{-\infty}^{x} f(z)\mathrm{d}z$.

(a) Let us define:

$$y_i = \begin{cases} 1, & \text{if } y_i^* > 0, \\ 0, & \text{if } y_i^* \le 0, \end{cases}$$

i.e., y_i^* is not observed and we know the sign of y_i^* (i.e., positive or negative). $y_i$is assigned to be one when $y_i^* > 0$, while it is zero when $y_i \le 0$.

(1) What is $\mathrm{E}(y_i)$?
(2) Obtain the likelihood function.
(3) Assuming that the density function of u_i is $f(\cdot)$, derive the first-order condition.
(4) Discuss how to estimate β and σ^2.
(5) What is the asymptotic distribution of the maximum likelihood estimator of $\beta^* = \dfrac{\beta}{\sigma}$?

(b) Let us define:

$$y_i^* = y_i, \qquad \text{if } y_i > 0,$$

i.e., y_t^* is not observed when $y_t \le 0$ and $y_t^* = y_t$ is observed when $y_t > 0$.

(6) What is $\mathrm{E}(y_i | y_i > 0)$?
(7) Obtain the likelihood function.
(8) Assuming that the density function of u_i is $f(\cdot)$, derive the first-order condition.
(9) Discuss how to estimate β and σ^2.
(10) What are the asymptotic distributions of the maximum likelihood estimators of β and σ^2?

(c) Let us define:

$$y_i^* = \begin{cases} y_i, & \text{if } y_i > 0, \\ 0, & \text{if } y_i \leq 0, \end{cases}$$

i.e., $y_t^* = 0$ is observed when $y_t \leq 0$ and $y_t^* = y_t$ is observed when $y_t > 0$.

(11) Obtain the likelihood function.

(12) Assuming that the density function of u_i is $f(\cdot)$, derive the first-order condition.

(13) Discuss how to estimate β and σ^2.

(14) What are the asymptotic distributions of the maximum likelihood estimators of β and σ^2?

$\boxed{2}$ Suppose that the probability function of y_i is Poisson with parameter λ_i for $i = 1, 2, \cdots, n$. Note that the Poisson distribution with parameter λ is:

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!},$$

for $x = 0, 1, 2, \cdots$.

(15) What is $\mathrm{E}(y_i)$?

(16) Assuming $\lambda_i = \exp(X_i\beta)$, obtain the likelihood function, where β is a unknown parameter vector to be estimated.

(17) Derive the first-order condition for maximization of the log-likelihood function.

(18) Discuss how to estimate β.

(19) What are the asymptotic distribution of the maximum likelihood estimator of β?

$\boxed{3}$ We estimate the following regression model:

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

for $i = 1, 2, \cdots, n$. u_1, u_2, $\cdots$, u_n are assumed to be mutually independent with mean zero and variance σ^2. 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.

(20) 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.

(21) 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?

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