Censored Regression Model or Tobit Model:

$$y_i = \begin{cases} X_i\beta + u_i, & \text{if } y_i > a, \\ a, & \text{otherwise.} \end{cases}$$

The probability which y_i takes a is given by:

$$P(y_i = a) = P(y_i \le a) = F(a) \equiv \int_{-\infty}^{a} f(x)\mathrm{d}x,$$

where $f(\cdot)$ and $F(\cdot)$ denote the density function and cumulative distribution function of y_i, respectively.

Therefore, the likelihood function is:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} F(a)^{I(y_i=a)} \times f(y_i)^{1-I(y_i=a)},$$

where $I(y_i = a)$ denotes the indicator function which takes one when $y_i = a$ or zero otherwise.

When $u_i \sim N(0, \sigma^2)$, the likelihood function is:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} \Big(\int_{-\infty}^{a} (2\pi\sigma^2)^{-1/2} \exp(-\frac{1}{2\sigma^2}(y_i - X_i\beta)^2) \mathrm{d}y_i \Big)^{I(y_i=a)} \times \Big((2\pi\sigma^2)^{-1/2} \exp(-\frac{1}{2\sigma^2}(y_i - X_i\beta)^2) \Big)^{1-I(y_i=a)},$$

which is maximized with respect to β and σ^2.

2.3 Count Data Model (計数データモデル)

Poisson distribution:

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

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

In the case of Poisson random variable X, the expectation of X is:

$$E(X) = \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} \lambda\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda\sum_{x'=0}^{\infty} \frac{e^{-\lambda}\lambda^{x'}}{x'!} = \lambda.$$

Remember that $\sum_x f(x) = 1$, i.e., $\sum_{x=0}^{\infty} e^{-\lambda}\lambda^x/x! = 1$.

Therefore, the probability function of the count data y_i is taken as the Poisson distribution with parameter λ_i.

In the case where the explained variable y_i takes 0, 1, 2, $\cdots$ (discrete numbers), assuming that the distribution of y_i is Poisson, the logarithm of λ_i is specified as a linear function, i.e.,

$$\mathrm{E}(y_i) = \lambda_i = \exp(X_i\beta).$$

Note that λ_i should be positive.

Therefore, it is better to avoid the specification: $\lambda = X_i\beta$.

The joint distribution of $y_1, y_2, \cdots, y_n$ is:

$$f(y_1, y_2, \cdots, y_n) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} = L(\beta),$$

where $\lambda_i = \exp(X_i\beta)$.

The log-likelihood function is:

$$\begin{aligned} \log L(\beta) &= -\sum_{i=1}^{n} \lambda_i + \sum_{i=1}^{n} y_i \log \lambda_i - \sum_{i=1}^{n} y_i! \\ &= -\sum_{i=1}^{n} \exp(X_i\beta) + \sum_{i=1}^{n} y_i X_i\beta - \sum_{i=1}^{n} y_i!. \end{aligned}$$

The first-order condition is:

$$\frac{\partial \log L(\beta)}{\partial \beta} = -\sum_{i=1}^{n} X_i' \exp(X_i\beta) + \sum_{i=1}^{n} X_i' y_i = 0.$$

$\Longrightarrow$ Nonlinear optimization procedure

[Review] Nonlinear Optimization Procedures:

Note that the Newton-Raphson method (one of the nonlinear optimization procedures) is:

$$\beta^{(j+1)} = \beta^{(j)} - \left(\frac{\partial^2 \log L(\beta^{(j)})}{\partial \beta \partial \beta'}\right)^{-1} \frac{\partial \log L(\beta^{(j)})}{\partial \beta},$$

which comes from the first-order Taylor series expansion around $\beta = \beta^*$:

$$0 = \frac{\partial \log L(\beta)}{\partial \beta} \approx \frac{\partial \log L(\beta^*)}{\partial \beta} + \frac{\partial^2 \log L(\beta^*)}{\partial \beta \partial \beta'}(\beta - \beta^*),$$

and β and β^* are replaced by $\beta^{(j+1)}$ and $\beta^{(j)}$, respectively.

An alternative nonlinear optimization procedure is known as the method of scoring, which is shown as:

$$\beta^{(j+1)} = \beta^{(j)} - \left(\mathrm{E}\Big(\frac{\partial^2 \log L(\beta^{(j)})}{\partial\beta\partial\beta'} \Big) \right)^{-1} \frac{\partial \log L(\beta^{(j)})}{\partial\beta},$$

where $\left(\dfrac{\partial^2 \log L(\beta^{(j)})}{\partial\beta\partial\beta'} \right)$ is replaced by $\mathrm{E}\left(\dfrac{\partial^2 \log L(\beta^{(j)})}{\partial\beta\partial\beta'} \right)$.

[End of Review]

In this case, we have the following iterative procedure:

$$\beta^{(j+1)} = \beta^{(j)} - \left(-\sum_{i=1}^{n} X_i' X_i \exp(X_i\beta^{(j)}) \right)^{-1} \left(-\sum_{i=1}^{n} X_i' \exp(X_i\beta^{(j)}) + \sum_{i=1}^{n} X_i' y_i \right).$$

The Newton-Raphson method is equivalent to the scoring method in this count model, because any random variable is not included in the expectation.

Zero-Inflated Poisson Count Data Model: In the case of too many zeros, we have to modify the estimation procedure.

Suppose that the probability of $y_i = j$ is decomposed of two regimes.

$\longrightarrow$ We have the case of $y_i = j$ and Regime 1, and that of $y_i = j$ and Regime 2.

Consider $P(y_i = 0)$ and $P(y_i = j)$ separately as follows:

$$P(y_i = 0) = P(y_i = 0|\text{Regime 1})P(\text{Regime 1}) + P(y_i = 0|\text{Regime 2})P(\text{Regime 2})$$

$$P(y_i = j) = P(y_i = j|\text{Regime 1})P(\text{Regime 1}) + P(y_i = j|\text{Regime 2})P(\text{Regime 2}),$$

for $j = 1, 2, \cdots$.

Assume:

- $P(y_i = 0|\text{Regime 1}) = 1$ and $P(y_i = j|\text{Regime 1}) = 0$ for $j = 1, 2, \cdots$,

- $P(\text{Regime 1}) = F_i$ and $P(\text{Regime 2}) = 1 - F_i$,

- $P(y_i = j|\text{Regime 2}) = \dfrac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$ for $j = 0, 1, 2, \cdots$,

where $F_i = F(Z_i\alpha)$ and $\lambda_i = \exp(X_i\beta)$. $\Longrightarrow$ w_i and X_i are exogenous variables.

Under the first assumption, we have the following equations:

$$P(y_i = 0) = P(\text{Regime 1}) + P(y_i = 0|\text{Regime 2})P(\text{Regime 2})$$

$$P(y_i = j) = P(y_i = j|\text{Regime 2})P(\text{Regime 2}),$$

for $j = 1, 2, \cdots$.

Combining the above two equations, we obtain the following:

$$P(y_i = j) = P(\text{Regime 1})I_i + P(y_i = j|\text{Regime 2})P(\text{Regime 2}),$$

for $j = 0, 1, 2, \cdots$,

where the indicator function I_i is given by $I_i = 1$ for $y_i = 0$ and $I_i = 0$ for $y_i \neq 0$.

F_i denotes a cumulative distribution function of $Z_i\alpha$. $\Longrightarrow$ We have to assume F_i.

Including the other two assumptions, we obtain the distribution of y_i as follows:

$$P(y_i = j) = F_i I_i + \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}(1 - F_i), \qquad j = 0, 1, 2, \cdots$$

where $F_i \equiv F(Z_i\alpha)$, $\lambda_i = \exp(X_i\beta)$, and the indicator function I_i is given by $I_i = 1$ for $y_i = 0$ and $I_i = 0$ for $y_i \neq 0$.

Therefore, the log-likelihood function is:

$$\log L(\alpha,\beta) = \sum_{i=1}^{n} \log P(y_i = j) = \sum_{i=1}^{n} \log\left(F_i I_i + \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}(1 - F_i)\right),$$

where $F_i \equiv F(Z_i\alpha)$ and $\lambda_i = \exp(X_i\beta)$.

$\log L(\alpha,\beta)$ is maximized with respect to α and β.

$\Longrightarrow$ The Newton-Raphson method or the method of scoring is utilized for maximization.

3 Panel Data

3.1 GLS — Review

Regression model:

$$y = X\beta + u, \qquad\qquad u \sim N(0, \Omega),$$

where y, X, β, u, 0 and Ω are $n \times 1$, $n \times k$, $k \times 1$, $n \times 1$, $n \times 1$, and $n \times n$, respectively.

We solve the following minimization problem:

$$\min_{\beta} (y - X\beta)'\Omega^{-1}(y - X\beta).$$

Let b be a solution of the above minimization problem.

GLS estimator of β is given by:

$$b = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.$$

3.2 Panel Model Basic

Model:

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

where i indicates individual and t denotes time.

There are n observations for each t.

u_{it} indicates the error term, assuming that $\mathrm{E}(u_{it}) = 0$, $\mathrm{V}(u_{it}) = \sigma_u^2$ and $\mathrm{Cov}(u_{it}, u_{js}) = 0$ for $i \neq j$ and $t \neq s$.

v_i denotes the individual effect, which is fixed or random.