ECO475H1 S
Applied Econometrics II

Tianyu Du

January 24, 2019

`Github Page https://github.com/TianyuDu/Spikey_UofT_Notes`
`Note Page TianyuDu.com/notes`

Contents

1 Lecture 3. Jan. 24 2019 **2**

1.1 Two Side Censoring MLE 2

1.2 Two Side Truncated MLE 3

1 Lecture 3. Jan. 24 2019

1.1 Two Side Censoring MLE

Consider the latent dependent variable

$$Y^* = \mathbf{x}'\boldsymbol{\beta} + \epsilon \tag{1.1}$$

where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$.
Therefore, given fixed $\mathbf{x}$,

$$Y^* \sim \mathcal{N}(\mathbf{x}'\boldsymbol{\beta}, \sigma^2) \tag{1.2}$$

Define parameter set

$$\boldsymbol{\theta} \equiv (\boldsymbol{\beta}, \sigma) \tag{1.3}$$

The observable variable is

$$Y = \begin{cases} U & \text{if } Y^* \geq U \\ Y^* & \text{if } Y^* \in (L, U) \\ L & \text{if } Y^* \leq L \end{cases} \tag{1.4}$$

Let $f_Y(y|\mathbf{x}, \boldsymbol{\beta}) : [L, U] \to [0, 1]$ be the probability measure of Y.
Let $y \in [L, U]$,

$$f_Y(y|\mathbf{x}, \boldsymbol{\beta}) = \begin{cases} \mathbb{P}(Y^* \geq U|\mathbf{x}, \boldsymbol{\beta}) & \text{if } y \geq U \\ f_{Y^*}(y|\mathbf{x}, \boldsymbol{\beta}) & \text{if } y \in (L, U) \\ \mathbb{P}(Y^* \leq L|\mathbf{x}, \boldsymbol{\beta}) & \text{if } y \leq L \end{cases} \tag{1.5}$$

$$= \begin{cases} 1 - F_{Y^*}(U|\mathbf{x}, \boldsymbol{\beta}) & \text{if } y \geq U \\ f_{Y^*}(y|\mathbf{x}, \boldsymbol{\beta}) & \text{if } y \in (L, U) \\ F_{Y^*}(L|\mathbf{x}, \boldsymbol{\beta}) & \text{if } y \leq L \end{cases} \tag{1.6}$$

Define indicator $(d_1(y), d_2(y), d_3(y))$ as

$$d_1(y) \equiv \mathcal{I}(y \geq U) \tag{1.7}$$

$$d_2(y) \equiv \mathcal{I}(y \in (L, U)) \tag{1.8}$$

$$d_3(y) \equiv \mathcal{I}(y \leq L) \tag{1.9}$$

Then the probability measure of Y can be expressed as

$$f_Y(y|\mathbf{x}, \boldsymbol{\beta}) = (1 - F_{Y^*}(U|\mathbf{x}, \boldsymbol{\beta}))^{d_1} \times f_{Y^*}(y|\mathbf{x}, \boldsymbol{\beta})^{d_2} \times F_{Y^*}(L|\mathbf{x}, \boldsymbol{\beta})^{d_3} \tag{1.10}$$

Suppose samples are i.i.d., the joint density is

$$f_{Y_1,\dots,Y_N}(y_1,\dots,y_N|\mathbf{X},\boldsymbol{\beta}) = \prod_{i=1}^{N} f_Y(y_i|\mathbf{x}_i,\boldsymbol{\beta}) \tag{1.11}$$

The log-likelihood is

$$\mathcal{L}_N(\boldsymbol{\theta}|\mathbf{X}) = \sum_{i=1}^{N} \Big\{ d_{1,i} \times \ln(1 - F_{Y^*}(U|\mathbf{x}_i,\boldsymbol{\beta})) + d_{2,i} \times \ln(f_{Y^*}(y|\mathbf{x}_i,\boldsymbol{\beta})) + d_{3,i} \times \ln(F_{Y^*}(L|\mathbf{x}_i,\boldsymbol{\beta})) \Big\} \tag{1.12}$$

Finally, solving

$$\hat{\boldsymbol{\theta}}_{MLE} = (\hat{\boldsymbol{\beta}}_{MLE}, \hat{\sigma}_{MLE}) = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmax}}\, \mathcal{L}_N(\boldsymbol{\theta}) \tag{1.13}$$

1.2 Two Side Truncated MLE

Suppose the observations are truncated with lower and upper bounds L and U.
Let the latent dependent variable be

$$Y^* = \mathbf{x}'\boldsymbol{\beta} + \epsilon \tag{1.14}$$

and

$$\epsilon \sim \mathcal{N}(0, \sigma^2) \tag{1.15}$$

which implies, for given $\mathbf{x}$,

$$Y^* \sim \mathcal{N}(\mathbf{x}'\boldsymbol{\beta}, \sigma^2) \tag{1.16}$$

Define parameter set

$$\boldsymbol{\theta} \equiv \{\boldsymbol{\beta}, \sigma\} \tag{1.17}$$

Observable random variable Y is

$$Y = \begin{cases} Y^* & \text{if } Y^* \in (L, U) \\ -- & \text{if } Y^* \notin (L, U) \end{cases} \tag{1.18}$$

Constructing the distribution for Y, note that F_Y is only defined on $y \in (L, U)$,

$$F_Y(y|\mathbf{x}, \boldsymbol{\theta}) = \mathbb{P}(Y < y|\mathbf{x}, \boldsymbol{\theta}) \tag{1.19}$$

$$= \frac{\mathbb{P}(Y^* < y \wedge Y^* \in (L, U)|\mathbf{x}, \boldsymbol{\theta})}{\mathbb{P}(Y^* \in (L, U)|\mathbf{x}, \boldsymbol{\theta})} \tag{1.20}$$

$$= \frac{\mathbb{P}(Y^* \in (L, y)|\mathbf{x}, \boldsymbol{\theta})}{\mathbb{P}(Y^* \in (L, U)|\mathbf{x}, \boldsymbol{\theta})} \tag{1.21}$$

$$= \frac{F_{Y^*}(y|\mathbf{x}, \boldsymbol{\theta}) - F_{Y^*}(L|\mathbf{x}, \boldsymbol{\theta})}{F_{Y^*}(U|\mathbf{x}, \boldsymbol{\theta}) - F_{Y^*}(L|\mathbf{x}, \boldsymbol{\theta})} \tag{1.22}$$

Then construct the density of Y

$$f_Y(y|\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial F_Y(y|\mathbf{x}, \boldsymbol{\theta})}{\partial y} \tag{1.23}$$

$$= \frac{f_{Y^*}(y|\mathbf{x}, \boldsymbol{\theta})}{F_{Y^*}(U|\mathbf{x}, \boldsymbol{\theta}) - F_{Y^*}(L|\mathbf{x}, \boldsymbol{\theta})} \tag{1.24}$$

The sample log-likelihood is

$$\mathcal{L}_N(\boldsymbol{\theta}) = \sum_{i=1}^{N} \ln(f_{Y^*}(y_i|\mathbf{x}_i, \theta)) - \ln(F_{Y^*}(U|\mathbf{x}_i, \boldsymbol{\theta}) - F_{Y^*}(L|\mathbf{x}_i, \boldsymbol{\theta})) \tag{1.25}$$

and the estimator is given by

$$\hat{\boldsymbol{\theta}}_{MLE} = \{\hat{\boldsymbol{\beta}}_{MLE}, \hat{\sigma}_{MLE}\} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} \, \mathcal{L}_N(\boldsymbol{\theta}) \tag{1.26}$$