ECO 475 Homework 1

Tianhao Wang 1003907061

February 20, 2020

1

Given the truncation model as

$$y = \begin{cases} \boldsymbol{x}^T\boldsymbol{\beta} + \epsilon, & \text{if } L < y^* < U \\ NA, & \text{otherwise} \end{cases}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, we can estimate $\boldsymbol{\beta}$ and σ^2 as follows.

Recall that

$$\begin{aligned} F_\epsilon(\epsilon = t | L < \epsilon < U) &= Pr(\epsilon \le t | L < \epsilon < U) \\ &= \frac{Pr(\epsilon \le t, L < \epsilon < U)}{Pr(L < \epsilon < U)} = \frac{\Phi(\frac{t}{\sigma}) - \Phi(\frac{L}{\sigma})}{\Phi(\frac{U}{\sigma}) - \Phi(\frac{L}{\sigma})}, \quad \forall\ t \in (L, U), \end{aligned}$$

and 0 otherwise. Thus, we have

$$f_\epsilon(\epsilon = t | L < \epsilon < U) = F'_\epsilon(\epsilon = t | L < \epsilon < U) = \frac{\phi(\frac{t}{\sigma})}{\sigma[\Phi(\frac{U}{\sigma}) - \Phi(\frac{L}{\sigma})]} \quad \forall\ t \in (L, U).$$

Then, we have 2 methods to consistently estimate $\boldsymbol{\beta}$ and σ^2. [1].

1. *Method 1: NLLS.*

$$\begin{aligned} E[y|\boldsymbol{x}] &= E[y^*|\boldsymbol{x}, y^* \in (L, U)] \\ &= E[\boldsymbol{x}^T\boldsymbol{\beta} + \epsilon | \boldsymbol{x}, \boldsymbol{x}^T\boldsymbol{\beta} + \epsilon \in (L, U)] \\ &= \boldsymbol{x}^T\boldsymbol{\beta} + E[\epsilon | \epsilon \in (L - \boldsymbol{x}^T\boldsymbol{\beta}, U - \boldsymbol{x}^T\boldsymbol{\beta})]. \end{aligned}$$

[1] For grading purposes, please grade Method 1 only. For evaluation purposes, please evaluate Method 2, if you have extra time. Thank you!

Further, let $L^* = L - \boldsymbol{x}^T\boldsymbol{\beta}$ and $u^* = U - \boldsymbol{x}^T\boldsymbol{\beta}$ for notation convenience.

$$
\begin{aligned}
E[\epsilon|\epsilon \in (L^*, U^*)] &= \int_{L^*}^{U^*} \epsilon \cdot f_\epsilon(\epsilon|L^* < \epsilon < U^*) d\epsilon \\
&= \int_{L^*}^{U^*} \epsilon \cdot \frac{\phi(\frac{t}{\sigma})}{\sigma[\Phi(\frac{U^*}{\sigma}) - \Phi(\frac{L^*}{\sigma})]} d\epsilon \\
&= \frac{\sigma}{\Phi(\frac{U^*}{\sigma}) - \Phi(\frac{L^*}{\sigma})} \int_{L^*}^{U^*} \frac{\epsilon}{\sigma} \phi(\frac{\epsilon}{\sigma}) d\frac{\epsilon}{\sigma} \\
&= \frac{\sigma}{\Phi(\frac{U^*}{\sigma}) - \Phi(\frac{L^*}{\sigma}))} \left(\phi(\frac{\epsilon}{\sigma}) \right|_{\epsilon = U^*}^{L^*} = \sigma \frac{\phi(\frac{U^*}{\sigma}) - \phi(\frac{L^*}{\sigma})}{\Phi(\frac{U^*}{\sigma}) - \Phi(\frac{L^*}{\sigma})} \overset{def}{=} \sigma \lambda(\frac{L^*}{\sigma}, \frac{U^*}{\sigma})
\end{aligned}
$$

Therefore,

$$
E[y|\boldsymbol{x}] = E[y^*|\boldsymbol{x}, y^* \in (L, U)] = \boldsymbol{x}^T\boldsymbol{\beta} + \sigma\lambda(\frac{L^*}{\sigma}, \frac{U^*}{\sigma}).
$$

$$
(\widehat{\boldsymbol{\beta}}_{NLLS}, \widehat{\sigma}^2_{NLLS}) = \underset{\boldsymbol{b}, \sigma^2}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{x}_i^T \boldsymbol{\beta} + \sigma\lambda(\cdot) - y_i)^2.
$$

2. *Method 2: MLE.*

   Given $F_y(y = t) = F_\epsilon(\epsilon = t - \boldsymbol{x}^T\boldsymbol{\beta}) = Pr(\epsilon \le t - \boldsymbol{x}^T\boldsymbol{\beta}), \quad \forall\ t \in (L, U)$, and 0 otherwise, we have

$$
f(y = t|\boldsymbol{x}) = \frac{\phi(\frac{t - \boldsymbol{x}^T\boldsymbol{\beta}}{\sigma})}{\sigma[\Phi(\frac{U}{\sigma}) - \Phi(\frac{L}{\sigma})]} \quad \forall\ t \in (L, U).
$$

   Therefore, we can do MLE via

$$
(\widehat{\boldsymbol{\beta}}_{MLE}, \widehat{\sigma}^2_{MLE}) = \underset{\boldsymbol{b}, \sigma^2}{\arg\max} \frac{1}{n} \sum_{i=1}^{n} \ln f(y_i|\boldsymbol{x}_i).
$$

Clearly, neither method would give a explicit solution via FOC of Lagrangian optimization. Fortunately, the solution does exists, refer to the textbook Chap.16.3.

Therefore, the next step I would do is to estimate $\boldsymbol{\beta}$ and σ^2 via MATLAB etc, where function argmax and argmin are built in.

2

(a) The result of $\boldsymbol{\beta}_{probit}$ is as follows.

Under the model given by

$$
y_i = I(y_i^* \ge 0) = I(\boldsymbol{x}_i^T \boldsymbol{\beta}_{probit} \ge \epsilon_i),
$$

the coefficients above are unique only up to a multiplicative constant.

Therefore, given the normalized variance of ϵ_i as 1, the interpretation of the coefficients is that the coefficient β_j represent the effect on y^* of a one unit change in the value of x_j (or v.s. if x_j is discrete), which identify the direction and significant level of x_j affects on y via function $I(\cdot)$.

Meanwhile, given the p-value of β_j, we rejected the hypothesis $H_0 : \beta_j = 0$ at 95% confidence level for $j \in \{1, 2, 3, 4, 5, 6\}$, but failed to reject it for $j \in \{0\}$.

	estimate	Std. Error	t-statistic	p-value	Sig
β_0	-0.9246	0.4803	-1.925	0.05422	.
β_1	2.0117	0.5230	3.846	0.00012	***
β_2	-3.3239	0.7008	-4.743	2.10e-06	***
β_3	2.6569	0.5627	4.722	2.34e-06	***
β_4	0.9050	0.2630	3.441	0.00058	***
β_5	-2.3806	0.5334	-4.463	8.07e-06	***
β_6	0.3098	0.1505	2.059	0.03950	*

Table 1: Probit Model Regression Result

(b) On average, the change in probability of opening a new KFC store at the location from having a McDonald near by versus not having one is -0.4265715.

On average, the change in probability of opening a new KFC store at the location from increasing 1% distance to the nearest KFC distribution center is -0.3081857.

(c) The result of $\boldsymbol{\beta}_{logit}$ is as follows.

	estimate	Std. Error	t-statistic	p-value	Sig
β_0	-1.5653	0.8618	-1.816	0.069343	.
β_1	3.4151	0.9587	3.562	0.000367	***
β_2	-5.7684	1.3219	-4.364	1.28e-05	***
β_3	4.6095	1.0646	4.330	1.49e-05	***
β_4	1.5441	0.4804	3.214	0.001308	**
β_5	-4.1066	0.9933	-4.134	3.56e-05	***
β_6	0.5346	0.2733	1.956	0.050501	.

Table 2: Logit Model Regression Result

under the model given by

$$y_i = I(y_i^* \geq 0) = I(\boldsymbol{x}_i^T \boldsymbol{\beta}_{logit} \geq \epsilon_i).$$

The coefficients above are unique only up to a multiplicative constant.

Therefore, given the normalized variance of ϵ_i as $\frac{\pi^2}{3}$, the interpretation of the coefficients is that the coefficient β_j represent the effect on y^* of a one unit change in the value of x_j (or v.s. if x_j is discrete), which identify the direction and significant level of x_j affects on y via function $I(\cdot)$.

Meanwhile, given the p-value of β_j, we rejected the hypothesis $H_0 : \beta_j = 0$ at 95% confidence level for $j \in \{1, 2, 3, 4, 5\}$, but failed to reject it for $j \in \{0, 6\}$.

Average Marginal Effect:

On average, the change in probability of opening a new KFC store at the location from having a McDonald near by versus not having one is -0.4268979.

On average, the change in probability of opening a new KFC store at the location from increasing 1% distance to the nearest KFC distribution center is -0.3063677.

(d) They are not very different.

$$\frac{\widehat{\boldsymbol{\beta}}_{logit}}{||\widehat{\boldsymbol{\beta}}_{logit}||} = \begin{bmatrix} -0.16669782 \\ 0.36370490 \\ -0.61432465 \\ 0.49089839 \\ 0.16444449 \\ -0.43734583 \\ 0.05693174 \end{bmatrix} \approx \begin{bmatrix} -0.16997129 \\ 0.36981207 \\ -0.61102532 \\ 0.48841960 \\ 0.16636517 \\ -0.43763177 \\ 0.05695797 \end{bmatrix} = \frac{\widehat{\boldsymbol{\beta}}_{probit}}{||\widehat{\boldsymbol{\beta}}_{probit}||}$$

3 Appendix

- For code detail, refer to my Github.
- For STATA output of dofile, refer to the next page.

```
  ___  ____  ____  ____  ____(R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/
  Statistics/Data Analysis
```

```
  ___  ____  ____  ____  ____ (R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/   15.1   Copyright 1985-2017 StataCorp LLC
  Statistics/Data Analysis            StataCorp
                                      4905 Lakeway Drive
     Special Edition                  College Station, Texas 77845 USA
                                      800-STATA-PC        http://www.stata.com
                                      979-696-4600        stata@stata.com
                                      979-696-4601 (fax)

22-student Stata lab perpetual license:
       Serial number:  401506229836
         Licensed to:  Map and Data
                       University of Toronto

Notes:
      1.  Unicode is supported; see help unicode_advice.
      2.  Maximum number of variables is set to 5000; see help set_maxvar.

running C:\Users\MDW620-CAF\Documents\profile.do ...

1 . do "C:\Users\MDW620-CAF\Documents\Tianhao.do"

2 . use C:\Users\MDW620-CAF\Downloads/hw1data.dta, clear

3 .
4 . logit y i.x1 i.x2 i.x3 x4 x5 x6

Iteration 0:   log likelihood = -82.760511
Iteration 1:   log likelihood =  -30.57662
Iteration 2:   log likelihood = -28.750296
Iteration 3:   log likelihood = -28.672363
Iteration 4:   log likelihood = -28.672342
Iteration 5:   log likelihood = -28.672342

Logistic regression                             Number of obs     =        120
                                                LR chi2(6)        =     108.18
                                                Prob > chi2       =     0.0000
Log likelihood = -28.672342                     Pseudo R2         =     0.6536

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        1.x1 |   3.415127   .9586819     3.56   0.000     1.536145    5.294109
        1.x2 |  -5.768404   1.321964    -4.36   0.000    -8.359406   -3.177402
        1.x3 |   4.609453   1.064614     4.33   0.000     2.522848    6.696058
          x4 |   1.544106   .4804108     3.21   0.001      .602518    2.485694
          x5 |  -4.106603   .9933383    -4.13   0.000    -6.053511   -2.159696
          x6 |   .5345794   .2733479     1.96   0.051    -.0011727    1.070331
       _cons |  -1.565264   .8618573    -1.82   0.069    -3.254473    .1239449
------------------------------------------------------------------------------

5 . margin, dydx(i.x2)

Average marginal effects                        Number of obs     =        120
Model VCE    : OIM

Expression   : Pr(y), predict()
dy/dx w.r.t. : 1.x2
```

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.x2	**-.4268979**	**.0440885**	**-9.68**	**0.000**	**-.5133099**	**-.3404859**

Note: dy/dx for factor levels is the discrete change from the base level.

6 . margin, dydx(x5)

```
Average marginal effects                          Number of obs     =        120
Model VCE    : OIM

Expression   : Pr(y), predict()
dy/dx w.r.t. : x5
```

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x5	**-.3063677**	**.0407848**	**-7.51**	**0.000**	**-.3863045**	**-.226431**

7 .
8 . probit y i.x1 i.x2 i.x3 x4 x5 x6

```
Iteration 0:   log likelihood = -82.760511
Iteration 1:   log likelihood = -30.299996
Iteration 2:   log likelihood = -28.400517
Iteration 3:   log likelihood = -28.336254
Iteration 4:   log likelihood = -28.336104
Iteration 5:   log likelihood = -28.336104

Probit regression                                 Number of obs     =        120
                                                  LR chi2(6)        =     108.85
                                                  Prob > chi2       =     0.0000
Log likelihood = -28.336104                       Pseudo R2         =     0.6576
```

y	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.x1	**2.011708**	**.5337449**	**3.77**	**0.000**	**.9655868**	**3.057828**
1.x2	**-3.323863**	**.7132776**	**-4.66**	**0.000**	**-4.721861**	**-1.925864**
1.x3	**2.656909**	**.5679074**	**4.68**	**0.000**	**1.543831**	**3.769987**
x4	**.9049947**	**.2688554**	**3.37**	**0.001**	**.3780479**	**1.431941**
x5	**-2.380636**	**.550527**	**-4.32**	**0.000**	**-3.459649**	**-1.301623**
x6	**.3098398**	**.1558339**	**1.99**	**0.047**	**.004411**	**.6152687**
_cons	**-.9246096**	**.4787586**	**-1.93**	**0.053**	**-1.862959**	**.0137399**

Note: 8 failures and 3 successes completely determined.

9 . margin, dydx(i.x2)

```
Average marginal effects                          Number of obs     =        120
Model VCE    : OIM

Expression   : Pr(y), predict()
dy/dx w.r.t. : 1.x2
```

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.x2	**-.4265715**	**.0432721**	**-9.86**	**0.000**	**-.5113833**	**-.3417598**

Note: dy/dx for factor levels is the discrete change from the base level.

```
10 . margin, dydx(x5)

   Average marginal effects                          Number of obs     =        120
   Model VCE    : OIM

   Expression   : Pr(y), predict()
   dy/dx w.r.t. : x5
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
x5	**-.3081857**	**.0412259**	**-7.48**	**0.000**	**-.3889869**	**-.2273845**

```
11 .
   end of do-file

12 .
```