1. In class we discussed data truncation from below. Now consider the following truncated model from both above and below:

$$Y = \begin{cases} x'\beta + \epsilon, & \text{if } L < Y^* < U \\ \text{Not observed}, & \text{otherwise} \end{cases}$$

where $\epsilon \sim N(0, \sigma^2)$ and (L, U) are observed bounds. Propose a method to consistently estimate β and σ^2. Please be as detailed as possible.

By MLE:

Y^* has the normal distribution $N(x_i'\beta, \sigma^2)$

$$F_Y(y) = P(Y^* < y \mid L < Y^* < U) = \frac{P(Y^* < y,\ L < Y^* < U)}{P(L < Y^* < U)}$$

$$= \frac{P(L < Y^* < y)}{P(L < Y^* < U)}$$

$$= \frac{P(Y^* < y) - P(Y^* < L)}{P(Y^* < U) - P(Y^* < L)}$$

$$= \frac{\Phi\left(\frac{y - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{L - x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{U - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{L - x_i'\beta}{\sigma}\right)}$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{\frac{1}{\sigma}\phi\left(\frac{y - x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{U - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{L - x_i'\beta}{\sigma}\right)}$$

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} f_Y(y) = \prod_{i=1}^{n} \frac{\frac{1}{\sigma}\phi\left(\frac{y - x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{U - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{L - x_i'\beta}{\sigma}\right)}$$

$$\ell(\beta, \sigma^2) = \ln L(\beta, \sigma^2)$$

$$= \sum_{i=1}^{n} \ln \left(\frac{\frac{1}{\sigma}\phi\left(\frac{y - x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{U - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{L - x_i'\beta}{\sigma}\right)} \right)$$

$$= \sum_{i=1}^{n} \left[\ln \frac{1}{\sigma} + \ln\left(\phi\left(\frac{y - x_i'\beta}{\sigma}\right)\right) - \ln\left[\Phi\left(\frac{U - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{L - x_i'\beta}{\sigma}\right)\right]\right]$$

Then we can obtain the estimate β and σ^2 by setting the $\frac{d\ell(\beta, \sigma^2)}{d\beta} = 0$ and $\frac{d\ell(\beta, \sigma^2)}{d\sigma^2} = 0$

$$(\hat{\beta}, \hat{\sigma}^2) = \arg\max \ell(\beta, \sigma^2)$$

2. Consider a binary choice model with linear latent index $y_i^* = x_i'\beta + u_i$. Assume that $u_i/\sigma(x_i) \sim F$ for some continuous and strictly increasing distribution function F, where $\sigma(x)$ is a function of x. Assume that $F(0) = 0$.

(a) Show that $p(x) \equiv \mathbb{P}(y_i = 1 | x_i = x) = F(x'\beta/\sigma(x))$.

$$p(x) = P(y_i = 1 \mid x_i = x) = P(y^* > 0)$$

$$= P(x'\beta + u_i > 0)$$

$$= P(-u_i < x'\beta)$$

$$= P\left(\frac{-u_i}{\sigma(x)} < \frac{x'\beta}{\sigma(x)}\right)$$

Since $\frac{u_i}{\sigma(x)} \sim F$ for some continous and strictly increasing distribution function, $\frac{-u_i}{\sigma(x)} \sim F$

$$= F(x'\beta/\sigma(x))$$

(b) Show that there are infinite number of combinations of (F, σ) that can generate the same choice probability $p(x)$.

We have $P(x) = P(y=1|x) = F\left(\frac{x'\beta}{\sigma(x)}\right)$ from question a).

Since x is given, assume an arbitary $\alpha > 0$

Let $F^*\left(\frac{x'\beta}{\sigma'(x)}\right) = F\left(\alpha \cdot \frac{x'\beta}{\sigma(x)}\right)$

$\sigma^*(x) = \alpha \cdot \sigma(x)$

Thus, $F^*\left(\frac{x'\beta}{\sigma'(x)}\right) = F\left(\alpha \cdot \frac{x'\beta}{\alpha\sigma(x)}\right) = P(x)$

By selecting different α,
there are infinite number of Combination of (F, σ) that can generate the same choice probability $p(x)$.

(c) Show that assuming $F = \Phi$ (the standard normal cdf) is not restrictive, that is, for any choice probability, one can find a function $\sigma^*(x)$ such that $p(x) = \Phi(x'\beta/\sigma^*(x))$.

By MLE

$$P(Y=y|x) = F\left(\frac{x'\beta}{\sigma(x)}\right)^{I(y_i=1)} \left(1-F\left(\frac{x'\beta}{\sigma(x)}\right)\right)^{I(y_i=0)}$$

$$= \Phi\left(\frac{x'\beta}{\sigma(x)}\right)^{I(y_i=1)} \left(1-\Phi\left(\frac{x'\beta}{\sigma(x)}\right)\right)^{I(y_i=0)}$$

$$L(\beta, \sigma(x)) = \prod_{i=1}^{n} P(Y_i=y_i|x_i)$$

$$= \prod_{i=1}^{n} \Phi\left(\frac{x_i'\beta}{\sigma(x)}\right)^{I(y_i=1)} \left[1-\Phi\left(\frac{x_i'\beta}{\sigma(x)}\right)\right]^{I(y_i=0)}$$

$$\ln L = \frac{1}{n}\left[\sum_{i=1}^{n} I(y_i=1) \log \Phi\left(\frac{x_i'\beta}{\sigma(x)}\right) + \sum_{i=1}^{n} I(y_i=0) \log\left(1-\Phi\left(\frac{x_i'\beta}{\sigma(x)}\right)\right)\right]$$

MLE: $(\hat{\beta}, \hat{\sigma}^*(x)) = \arg\max \ln L(\beta, \sigma(x))$

Thus, given any $F = \Phi$, we can calculate a function $\sigma^*(x)$ by MLE estimators such that $p(x) = \Phi\left(\frac{x\beta}{\sigma^*(x)}\right)$.

(d) You already know that assuming $\sigma(x) = 1$ is restrictive. Intuitively discuss the contrast of assuming a specific functional form for F against assuming a specific functional form for σ.

From previous question, we have $p(x) = F(\frac{x'\beta}{\sigma(x)})$. Assume fix the specific functional form of F, when x changes, we can gain the same value of $p(x)$ by selecting different function of $\sigma^*(x)$ from MLE.

In comparison, assuming fix a specific functional form of $\sigma(x)$, i.e. $\sigma(x)=1$ is restrictive. We have $p(x) = F(x'\beta)$. As x changes, we can gain the same value of $p(x)$ by finding various F (if β is given).

3. Consider the data in hw1data.dta. It contains data about whether KFC open a store at 120 locations.

y: 1 if open a new store, 0 if not.
x1: 1 if there is a shopping mall near the location
x2: 1 if there is already a McDonald store near the location
x3: 1 if there is a subway station near the location
x4: log of pedestrian flow at the nearest major intersection (pedestrian flow measured in 10,000)
x5: log-distance to the nearest KFC distribution center (distance measured in 10km)
x6: population residence density of the location (in 10,000)

Answer the following questions.

(a) Compute probit estimates of a model in which the market entry decision y is the dependent variable and corresponding t-statistics, tabulate and interpret them.

```
. * Fit probit model
. probit y x1 x2 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]
x1	2.011708	.5337449	3.77	0.000	.9655868	3.057828
x2	-3.323863	.7132776	-4.66	0.000	-4.721861	-1.925864
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.
```

From the result, we noticed that $\hat{\beta}_1 = 2.01$, $\hat{\beta}_2 = -3.32$, $\hat{\beta}_3 = 2.66$, $\hat{\beta}_4 = 0.90$, $\hat{\beta}_5 = -2.38$, $\hat{\beta}_6 = 0.31$, $\hat{\beta}_0 = -0.92$.

We set the significance level of 0.1. From the table, the $\hat{\beta}_i$ (for $i = 0 \cdots 6$) are statistically significant since their P value < 0.1.

In general, X_1, X_3, X_4, X_6 have positive correlation with our dependent variable y. And X_2, X_5 have negative correlation with our dependent variable y.

$$P(Y=1|X) = E(Y|X) = F(X\beta)$$

$$= \Phi(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \hat{\beta}_4 X_4 + \hat{\beta}_5 X_5 + \hat{\beta}_6 X_6)$$

$$= \Phi(-0.92 + 2.01X_1 - 3.32X_2 + 2.66X_3 + 0.9X_4 - 2.38X_5 + 0.31X_6)$$

where $\Phi(\cdot)$ is CDF of standard normal.

(b) Compute the average marginal effect of x2 and x5, interpret your result.

```
. * Average marginal effects
. margin, dydx(x2 x5)

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

Expression   : Pr(y), predict()
dy/dx w.r.t. : x2 x5
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
x2	-.4302913	.0455122	-9.45	0.000	-.5194935	-.341089
x5	-.3081857	.0412259	-7.48	0.000	-.3889869	-.2273845

The average marginal effects is -0.4302913 for x_2 and -0.3081857 for x_5.

The result tells that, holding other factors constant, if there is already a McDonald store near the location, the probability of opening a KFC will decrease by 43.03% on average.

Similarly, holding other factors constant, if the log-distance to nearest KFC distribution error increased by 1 unit, the probability of opening a KFC will decrease by 30.81% on averge.

(c) Repeat part (a) and (b) using Logit model.

```
. * Fit logit model
. logit y x1 x2 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]
x1	3.415127	.9586819	3.56	0.000	1.536145	5.294109
x2	-5.768404	1.321964	-4.36	0.000	-8.359406	-3.177402
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

From the result, we noticed that $\hat{\beta}_1 = 3.42$, $\hat{\beta}_2 = -5.77$, $\hat{\beta}_3 = 4.61$, $\hat{\beta}_4 = 1.54$, $\hat{\beta}_5 = -4.11$, $\hat{\beta}_6 = 0.53$, $\hat{\beta}_0 = -1.57$

We set the significance level of 0.1. From the table, the $\hat{\beta}_i$ (for $i = 0 \cdots 6$) are statistically significant since their P value < 0.1.

In general, X_1, X_3, X_4, X_6 have positive correlation with our dependent variable y. And X_2, X_5 have negative correlation with our dependent variable y.

$P(Y=1|x) = E(Y|X) = F(XB)$

$= \Lambda(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \hat{\beta}_4 x_4 + \hat{\beta}_5 x_5 + \hat{\beta}_6 x_6)$

$= \Lambda(-1.57 + 3.42x_1 - 5.77x_2 + 4.61x_3 + 1.54x_4 - 4.11x_5 + 0.53x_6)$

where $\Lambda(\cdot)$ is cdf of logistic distribution.

```
. * Average marginal effects
. margin, dydx(x2 x5)

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

Expression   : Pr(y), predict()
dy/dx w.r.t. : x2 x5
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
x2	-.4303442	.0469477	-9.17	0.000	-.52236	-.3383285
x5	-.3063677	.0407848	-7.51	0.000	-.3863045	-.226431

The average marginal effects is -0.4303442 for x_2 and -0.3063677 for x_5.

The result tells that, holding other factors constant, if there is a McDonald store near the location, the probability of opening a KFC will decrease by 43.03% on average.

Similarly, holding other factors constant, if the log-distance to nearest KFC distribution error increased by 1 unit, the probability of opening a KFC will decrease by 30.64% on averge.

(d) Let $\hat{\beta}_P$ and $\hat{\beta}_L$ be the probit estimates and logit estimates. Let $\|\beta\|$ be norm of β, that is, the square root of the sum of squares of each element in the vector β. Compare $\hat{\beta}_P/\|\hat{\beta}_P\|$ and $\hat{\beta}_L/\|\hat{\beta}_L\|$. Are they very different?

At 0.01 significance level,

$$\hat{\beta}_P = (-0.9246096, 2.011708, -3.323863, 2.656909, 0.9049947, -2.380636, 0.3098398)^T$$

$$\hat{\beta}_L = (-1.565264, 3.415127, -5.768404, 4.609453, 1.544106, -4.166603, 0.5345794)^T$$

$$\|\hat{\beta}_P\| = \sqrt{0.9246096^2 + 2.011708^2 + 3.323863^2 + 2.656909^2 + 0.9049947^2 + 2.380636^2 + 0.3098398^2}$$

$$= 5.44$$

$$\|\hat{\beta}_L\| = \sqrt{1.565264^2 + 3.415127^2 + 5.768404^2 + 4.609453^2 + 1.544106^2 + 4.166603^2 + 0.5345794^2}$$
$$= 9.39$$

$$\frac{\hat{\beta}_P}{\|\hat{\beta}_P\|} = (-0.17, 0.37, -0.61, 0.49, 0.17, -0.44, 0.06)^T$$

$$\frac{\hat{\beta}_L}{\|\hat{\beta}_L\|} = (-0.17, 0.36, -0.61, 0.49, 0.16, -0.44, 0.06)^T$$

Compare the result in $\frac{\hat{\beta}_P}{\|\hat{\beta}_P\|}$ and $\frac{\hat{\beta}_L}{\|\hat{\beta}_L\|}$, we find that the each value are close/similar to each other, with the same direction in each term.

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

                                                            User: LIANGJIAYI WANG
                                                         Project: sad

      name:  <unnamed>
       log:  C:\Users\WLJY8\Desktop\Courses\YEAR 4\WINTER\ECO475\HW1\log file.smcl
  log type:  smcl
 opened on:  22 Feb 2021, 23:47:24

1 . do "C:\Users\WLJY8\AppData\Local\Temp\STD4e0c_000000.tmp"

2 . clear all

3 .
4 . use "C:\Users\WLJY8\Desktop\Courses\YEAR 4\WINTER\ECO475\HW1\hw1data.dta"

5 .
6 . *Q3
7 .
8 . * Fit probit model
9 . probit y x1 x2 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]
-------------+----------------------------------------------------------------
          x1 |   2.011708   .5337449     3.77   0.000     .9655868    3.057828
          x2 |  -3.323863   .7132776    -4.66   0.000    -4.721861   -1.925864
          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.

10 .
11 . * Average marginal effects
12 . margin, dydx(x2 x5)

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

Expression   : Pr(y), predict()
dy/dx w.r.t. : x2 x5

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |  -.4302913   .0455122    -9.45   0.000    -.5194935    -.341089
          x5 |  -.3081857   .0412259    -7.48   0.000    -.3889869   -.2273845
------------------------------------------------------------------------------
```

```
13 .
14 . * Fit logit model
15 . logit y x1 x2 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]
   -------------+----------------------------------------------------------------
             x1 |   3.415127   .9586819     3.56   0.000     1.536145    5.294109
             x2 |  -5.768404   1.321964    -4.36   0.000    -8.359406   -3.177402
             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
   ------------------------------------------------------------------------------

16 . * Average marginal effects
17 . margin, dydx(x2 x5)

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

   Expression   : Pr(y), predict()
   dy/dx w.r.t. : x2 x5

   ------------------------------------------------------------------------------
                |            Delta-method
                |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
   -------------+----------------------------------------------------------------
             x2 |  -.4303442   .0469477    -9.17   0.000      -.52236   -.3383285
             x5 |  -.3063677   .0407848    -7.51   0.000    -.3863045    -.226431
   ------------------------------------------------------------------------------

18 .
   end of do-file

19 . log close
         name:  <unnamed>
          log:  C:\Users\WLJY8\Desktop\Courses\YEAR 4\WINTER\ECO475\HW1\log file.smcl
     log type:  smcl
    closed on:  22 Feb 2021, 23:47:35
   ---------------------------------------------------------------------------------------------------
```