The model

According to the following model,

$$y_{ki} = \beta_{k0} + \beta_{k1} x_i + u_{ki} \qquad (1)$$

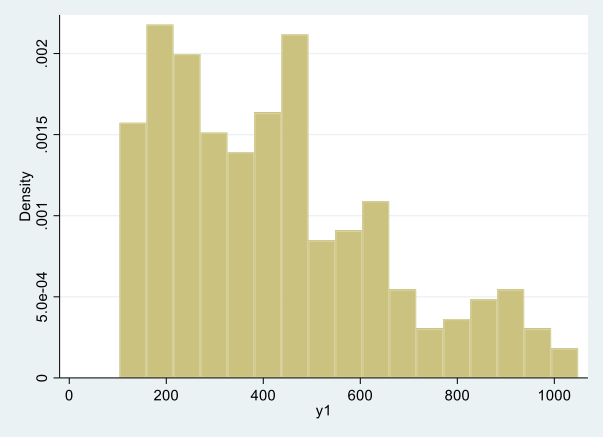
where: y_{ki} is Dependent variable, k=1, 2, 3.
x_i is Independent variable
u_i is Stochastic disturbance term

Requirements: (assign11.dta)

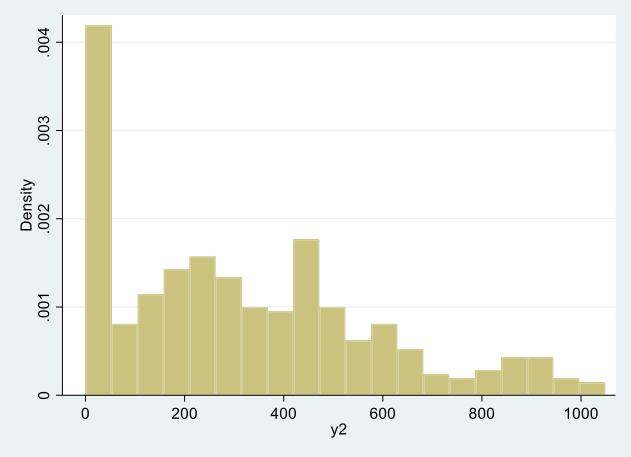
1 Plot histogram of y_{1i}, y_{2i}, y_{3i}, compute descriptive statistics of these three variables, then determine limitation of these three dependent variables.

```
use "C:\Users\Jilllin\OneDrive\Desktop\Thammasat\EE426\Data\assign11.dta"

. histogram y1
(bin=17, start=103.46626, width=55.638129)
```

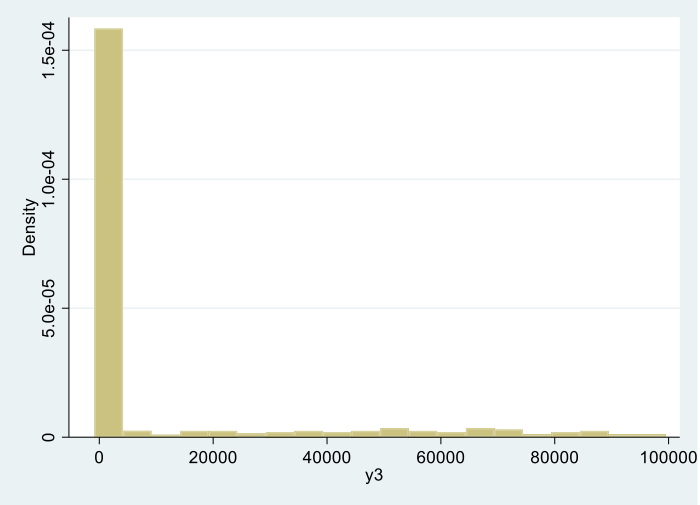



```
. histogram y2
(bin=20, start=0, width=52.465723)
```




```
. histogram y3
```

```
(bin=20, start=-866.33484, width=5018.7203)
```




```
. sum y1 y2 y3

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          y1 |        297    423.1683    227.8976   103.4663   1049.314
          y2 |        400    317.5914    266.3419          0   1049.314
          y3 |        400     10446.1    23186.57  -866.3348   99508.07
```

From the summary of the data, y1 has truncated problem. Y2 has censored problem (min of 0). Y3 has outlier problem.

2 Estimate the model (1) for y_{1i}, y_{2i}, y_{3i} using OLS, using truncated regression model, Tobit model, determine the most appropriated models for each y_{ki}.

```
*2

. reg y1 x

      Source |       SS           df       MS      Number of obs   =       297
-------------+----------------------------------   F(1, 295)       =     30.30
       Model |  1431905.59         1  1431905.59   Prob > F        =    0.0000
    Residual |  13941534.7       295  47259.4396   R-squared       =    0.0931
-------------+----------------------------------   Adj R-squared   =    0.0901
       Total |  15373440.3       296  51937.2982   Root MSE        =    217.39

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   123.1693   22.37637     5.50   0.000     79.13176    167.2069
       _cons |   69.01677   65.56422     1.05   0.293    -60.01611    198.0496
------------------------------------------------------------------------------

. est store y1ols

. sum y1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
```

```
-------------+--------------------------------------------------------
          y1 |        297    423.1683    227.8976   103.4663   1049.314

. scalar miny1=round(r(min))

. truncreg y1 x, ll(miny1) nolog
(note: 0 obs. truncated)

Truncated regression
Limit:   lower =        103                     Number of obs     =        297
         upper =       +inf                     Wald chi2(1)      =      23.38
Log likelihood = -1973.9423                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   249.6062   51.61679     4.84   0.000     148.4392    350.7733
       _cons |  -456.6913   178.5338    -2.56   0.011    -806.6112   -106.7714
-------------+----------------------------------------------------------------
      /sigma |   306.4488   26.46552    11.58   0.000     254.5774    358.3203
------------------------------------------------------------------------------

. est store y1trunc

. lrtest y1ols y1trunc, force

Likelihood-ratio test                                 LR chi2(1)  =     89.69
(Assumption: y1ols nested in y1trunc)                 Prob > chi2 =    0.0000

.
```

From LR-chi-squared test, H0 is rejected. We can conclude that truncated regression is more appropriate.

```
. reg y2 x

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(1, 398)       =     76.58
       Model |  4567087.34         1  4567087.34   Prob > F        =    0.0000
    Residual |  23737174.2       398  59641.1413   R-squared       =    0.1614
-------------+----------------------------------   Adj R-squared   =    0.1592
       Total |  28304261.6       399  70937.9989   Root MSE        =    244.22

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   179.8325   20.55046     8.75   0.000     139.4315    220.2335
       _cons |  -179.3679   58.08821    -3.09   0.002     -293.566    -65.1698
------------------------------------------------------------------------------

. est store y2ols

. tobit y2 x, ll(0)

Refining starting values:

Grid node 0:   log likelihood = -2400.6048
```

```
Fitting full model:

Iteration 0:   log likelihood = -2400.6048
Iteration 1:   log likelihood = -2389.6871
Iteration 2:   log likelihood = -2389.3078
Iteration 3:   log likelihood = -2389.3077

Tobit regression                                  Number of obs     =        400
                                                    Uncensored      =        328
Limits: lower = 0                                   Left-censored   =         72
        upper = +inf                                Right-censored  =          0

                                                  LR chi2(1)        =      73.17
                                                  Prob > chi2       =     0.0000
Log likelihood = -2389.3077                       Pseudo R2         =     0.0151

------------------------------------------------------------------------------
          y2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |    219.156   24.86468     8.81   0.000     170.2739    268.0382
       _cons |  -318.6258   70.93088    -4.49   0.000    -458.0707   -179.1808
-------------+----------------------------------------------------------------
    var(e.y2)|   81171.96   6591.132                      69195.6    95221.18
------------------------------------------------------------------------------

. est store y2tobit

. lrtest y2ols y2tobit, force

Likelihood-ratio test                                 LR chi2(1)  =    752.97
(Assumption: y2ols nested in y2tobit)                 Prob > chi2 =    0.0000

.
```

From LR-chi-squared test, Ho is rejected. We can conclude that tobit model is better than OLS.

```
reg y3 x

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(1, 398)       =     26.43
       Model |  1.3359e+10         1  1.3359e+10   Prob > F        =    0.0000
    Residual |  2.0115e+11       398   505402183   R-squared       =    0.0623
-------------+----------------------------------   Adj R-squared   =    0.0599
       Total |  2.1451e+11       399   537616802   Root MSE        =     22481

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   9726.042   1891.765     5.14   0.000     6006.941    13445.14
       _cons |  -16431.39   5347.288    -3.07   0.002    -26943.85   -5918.932
------------------------------------------------------------------------------

. est store y3ols

. tobit y3 x, ul(500) nolog
```

```
Tobit regression                                  Number of obs     =        400
                                                    Uncensored      =        314
Limits: lower = -inf                                Left-censored   =          0
        upper = 500                                 Right-censored  =         86

                                                  LR chi2(1)        =      67.20
                                                  Prob > chi2       =     0.0000
Log likelihood = -2305.4404                       Pseudo R2         =     0.0144

------------------------------------------------------------------------------
          y3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   216.5213   25.84758     8.38   0.000     165.7069    267.3358
       _cons |  -325.0374   72.09183    -4.51   0.000    -466.7647   -183.3101
-------------+----------------------------------------------------------------
    var(e.y3)|   85472.85   7109.987                      72578.21    100658.4
------------------------------------------------------------------------------

. est store y3tobit

. lrtest y3ols y3tobit
Mixed models are not nested
r(498);

. lrtest y3ols y3tobit, force

Likelihood-ratio test                                 LR chi2(1)  =   4538.61
(Assumption: y3ols nested in y3tobit)                 Prob > chi2 =    0.0000
```

From LR-chi-squared test, Ho is rejected. We can conclude that tobit model is better than OLS.