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

                                                        User: TianyuHW5_2

  ___  ____  ____  ____  ____ (R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/   15.0   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)

1150-user Stata network license expires 23 Jul 2019:
       Serial number:  401509213955
         Licensed to:  cuitianyu706599183@163.com
                       personal

Notes:
      1.  Unicode is supported; see help unicode_advice.
      2.  Maximum number of variables is set to 5000; see help set_maxvar.
      3.  New update available; type -update all-

 1 . doedit "C:\Users\cuiti\Master Study\Second Semester\econometrics\TIANYUCUI\ps5\ps2.do"

 2 . do "C:\Users\cuiti\AppData\Local\Temp\STD3898_000000.tmp"

 3 . ***set up for work place****
 4 . clear all

 5 . set more off, perm
   (set more preference recorded)

 6 . set scrollbufsize 2000000
   (set scrollbufsize will take effect the next time you launch Stata)

 7 . ****set working directory***
 8 . cd "C:\Users\cuiti\Master Study\Second Semester\econometrics\TIANYUCUI\ps5"
   C:\Users\cuiti\Master Study\Second Semester\econometrics\TIANYUCUI\ps5

 9 . ***set the observation***
10 . set obs 10000
   number of observations (_N) was 0, now 10,000

11 . *****exercise1: generate the variables to prepare for the regression*****
12 . generate X1 = runiform(1,3)

13 . generate X2 = runiform(0,1)

14 . generate X3=rbinomial(10000,0.3)

15 . gen eps = rnormal(2,1)

16 . gen Y = 0.5 + 1.2*X1 + -0.9*X2 + 0.1*X3 + eps
```

```
17 . * Basic Summary Stats
18 . su Y
```

Variable	Obs	Mean	Std. Dev.	Min	Max
Y	**10,000**	**304.4587**	**4.780371**	**283.3162**	**324.7488**

```
19 . egen Y_mean = mean(Y)

20 . gen Y_dum = (Y>Y_mean)

21 . tab Y_dum
```

Y_dum	Freq.	Percent	Cum.
0	**5,044**	**50.44**	**50.44**
1	**4,956**	**49.56**	**100.00**
Total	**10,000**	**100.00**	

```
22 .
23 . ****exercise 2: ols regression******
24 .  correlate Y X1
(obs=10,000)
```

	Y	X1
Y	**1.0000**	
X1	**0.1470**	**1.0000**

```
25 .  *the correlation of Y and X is 0.1416
26 .  display 0.1416-1.2
-1.0584

27 .  ***do the regression
28 .  reg Y X1 X2 X3
```

Source	SS	df	MS		
Model	**218184.148**	**3**	**72728.0493**	Number of obs	= **10,000**
Residual	**10312.4644**	**9,996**	**1.03165911**	F(3, 9996)	= **70496.20**
				Prob > F	= **0.0000**
Total	**228496.612**	**9,999**	**22.8519464**	R-squared	= **0.9549**
				Adj R-squared	= **0.9549**
				Root MSE	= **1.0157**

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X1	**1.186778**	**.0175235**	**67.72**	**0.000**	**1.152428**	**1.221128**
X2	**-.8714115**	**.0351358**	**-24.80**	**0.000**	**-.9402848**	**-.8025382**
X3	**.1001524**	**.0002204**	**454.36**	**0.000**	**.0997203**	**.1005845**
_cons	**2.055131**	**.6620921**	**3.10**	**0.002**	**.7572972**	**3.352965**

```
29 .
30 .
31 .  *****exercise 3&4: write the probit model and logit model****
32 .
33 . probit  Y_dum  X1 X2 X3

Iteration 0:   log likelihood = -6931.0846
Iteration 1:   log likelihood = -1531.2747
Iteration 2:   log likelihood = -1529.7943
Iteration 3:   log likelihood = -1529.7912
Iteration 4:   log likelihood = -1529.7912

Probit regression                                 Number of obs     =     10,000
                                                  LR chi2(3)        =   10802.59
                                                  Prob > chi2       =     0.0000
Log likelihood = -1529.7912                       Pseudo R2         =     0.7793

------------------------------------------------------------------------------
       Y_dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          X1 |    1.235118   .0520075    23.75   0.000     1.133185    1.337051
          X2 |   -.7780938   .0895126    -8.69   0.000    -.9535353   -.6026523
          X3 |    .0981449    .002319    42.32   0.000     .0935998    .1026901
       _cons |   -296.5425   7.003491   -42.34   0.000     -310.269   -282.8159
------------------------------------------------------------------------------
Note: 950 failures and 948 successes completely determined.

34 . logit Y_dum X1 X2 X3

Iteration 0:   log likelihood = -6931.0846
Iteration 1:   log likelihood = -1538.5927
Iteration 2:   log likelihood = -1532.8409
Iteration 3:   log likelihood = -1532.8404
Iteration 4:   log likelihood = -1532.8404

Logistic regression                               Number of obs     =     10,000
                                                  LR chi2(3)        =   10796.49
                                                  Prob > chi2       =     0.0000
Log likelihood = -1532.8404                       Pseudo R2         =     0.7788

------------------------------------------------------------------------------
       Y_dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          X1 |    2.227785   .0972505    22.91   0.000     2.037178    2.418393
          X2 |   -1.419549   .1621312    -8.76   0.000     -1.73732   -1.101777
          X3 |    .1775505   .0046667    38.05   0.000     .1684038    .1866971
       _cons |   -536.4411   14.09405   -38.06   0.000    -564.0649   -508.8172
------------------------------------------------------------------------------
Note: 149 failures and 151 successes completely determined.

35 .
36 . *****calculate the marginal effeccts of probit model and logit model****
37 . probit  Y_dum  X1 X2 X3

Iteration 0:   log likelihood = -6931.0846
Iteration 1:   log likelihood = -1531.2747
Iteration 2:   log likelihood = -1529.7943
Iteration 3:   log likelihood = -1529.7912
Iteration 4:   log likelihood = -1529.7912

Probit regression                                 Number of obs     =     10,000
                                                  LR chi2(3)        =   10802.59
                                                  Prob > chi2       =     0.0000
Log likelihood = -1529.7912                       Pseudo R2         =     0.7793
```

Y_dum	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
X1	**1.235118**	**.0520075**	**23.75**	**0.000**	**1.133185**	**1.337051**
X2	**-.7780938**	**.0895126**	**-8.69**	**0.000**	**-.9535353**	**-.6026523**
X3	**.0981449**	**.002319**	**42.32**	**0.000**	**.0935998**	**.1026901**
_cons	**-296.5425**	**7.003491**	**-42.34**	**0.000**	**-310.269**	**-282.8159**

Note: 950 failures and 948 successes completely determined.

```
38 . quietly probit $ylist $xlist

39 . margins, dydx(*) atmeans
```

Conditional marginal effects Number of obs = **10,000**
Model VCE : **OIM**

Expression : **Pr(Y_dum), predict()**
dy/dx w.r.t. : **X1 X2 X3**
at : X1 = **2.001611** (mean)
X2 = **.5030039** (mean)
X3 = **3000.092** (mean)

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
X1	**.492662**	**.0207467**	**23.75**	**0.000**	**.4519993**	**.5333248**
X2	**-.3103649**	**.0357046**	**-8.69**	**0.000**	**-.3803446**	**-.2403852**
X3	**.0391479**	**.0009249**	**42.32**	**0.000**	**.037335**	**.0409608**

```
40 . margins, dydx(*)
```

Average marginal effects Number of obs = **10,000**
Model VCE : **OIM**

Expression : **Pr(Y_dum), predict()**
dy/dx w.r.t. : **X1 X2 X3**

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
X1	**.1049494**	**.0036532**	**28.73**	**0.000**	**.0977893**	**.1121096**
X2	**-.0661155**	**.0074482**	**-8.88**	**0.000**	**-.0807137**	**-.0515174**
X3	**.0083395**	**.0000116**	**717.99**	**0.000**	**.0083167**	**.0083623**

```
41 .
42 . logit Y_dum X1 X2 X3

Iteration 0:   log likelihood = -6931.0846
Iteration 1:   log likelihood = -1538.5927
Iteration 2:   log likelihood = -1532.8409
Iteration 3:   log likelihood = -1532.8404
Iteration 4:   log likelihood = -1532.8404
```

Logistic regression Number of obs = **10,000**
LR chi2(**3**) = **10796.49**
Prob > chi2 = **0.0000**
Log likelihood = **-1532.8404** Pseudo R2 = **0.7788**

Y_dum	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
X1	**2.227785**	**.0972505**	**22.91**	**0.000**	**2.037178**	**2.418393**
X2	**-1.419549**	**.1621312**	**-8.76**	**0.000**	**-1.73732**	**-1.101777**
X3	**.1775505**	**.0046667**	**38.05**	**0.000**	**.1684038**	**.1866971**
_cons	**-536.4411**	**14.09405**	**-38.06**	**0.000**	**-564.0649**	**-508.8172**

Note: 149 failures and 151 successes completely determined.

```
43 . quietly logit $ylist $xlist

44 . margins, dydx(*) atmeans
```

Conditional marginal effects Number of obs = **10,000**
Model VCE : **OIM**

Expression : **Pr(Y_dum), predict()**
dy/dx w.r.t. : **X1 X2 X3**
at : X1 = **2.001611** (mean)
X2 = **.5030039** (mean)
X3 = **3000.092** (mean)

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
X1	**.5568354**	**.0243146**	**22.90**	**0.000**	**.5091796**	**.6044911**
X2	**-.3548164**	**.0405264**	**-8.76**	**0.000**	**-.4342466**	**-.2753862**
X3	**.0443788**	**.0011665**	**38.05**	**0.000**	**.0420926**	**.046665**

```
45 . margins, dydx(*)
```

Average marginal effects Number of obs = **10,000**
Model VCE : **OIM**

Expression : **Pr(Y_dum), predict()**
dy/dx w.r.t. : **X1 X2 X3**

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
X1	**.1045903**	**.0036465**	**28.68**	**0.000**	**.0974432**	**.1117374**
X2	**-.0666451**	**.0074084**	**-9.00**	**0.000**	**-.0811654**	**-.0521248**
X3	**.0083357**	**.0000122**	**685.37**	**0.000**	**.0083118**	**.0083595**

```
46 .
end of do-file

47 .
```