Stata Textbook Examples

Introductory Econometrics: A Modern Approach by Jeffrey M. Wooldridge (1st & 2d eds.)

Chapter 17 - Limited Dependent Variable Models and Sample Selection Corrections

Example 17.1: Married Woman's Labor Force Participation

可以直接输入这个命令引用数据。

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ, clear
```

```
regress inlf nwifeinc educ exper expersq age kidslt6 kidsge6

      Source |       SS       df       MS              Number of obs =     753
-------------+------------------------------           F(  7,   745) =   38.22
       Model |  48.8080578     7  6.97257968           Prob > F      =  0.0000
    Residual |  135.919698   745  .182442547           R-squared     =  0.2642
-------------+------------------------------           Adj R-squared =  0.2573
       Total |  184.727756   752  .245648611           Root MSE      =  .42713

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0034052   .0014485    -2.35   0.019    -.0062488   -.0005616
        educ |   .0379953    .007376     5.15   0.000      .023515    .0524756
       exper |   .0394924   .0056727     6.96   0.000     .0283561    .0506287
     expersq |  -.0005963   .0001848    -3.23   0.001    -.0009591   -.0002335
         age |  -.0160908   .0024847    -6.48   0.000    -.0209686    -.011213
     kidslt6 |  -.2618105   .0335058    -7.81   0.000    -.3275875   -.1960335
     kidsge6 |   .0130122    .013196     0.99   0.324    -.0128935    .0389179
       _cons |   .5855192    .154178     3.80   0.000     .2828442    .8881943
------------------------------------------------------------------------------
```

```
logit inlf nwifeinc educ exper expersq age kidslt6 kidsge6

Iteration 0:   log likelihood =  -514.8732
Iteration 1:   log likelihood = -406.94123
Iteration 2:   log likelihood = -401.85151
Iteration 3:   log likelihood = -401.76519
Iteration 4:   log likelihood = -401.76515

Logit estimates                                   Number of obs   =        753
                                                  LR chi2(7)      =     226.22
                                                  Prob > chi2     =     0.0000
Log likelihood = -401.76515                       Pseudo R2       =     0.2197

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0213452   .0084214    -2.53   0.011    -.0378509   -.0048394
        educ |   .2211704   .0434396     5.09   0.000     .1360303    .3063105
```

```
       exper |   .2058695   .0320569     6.42   0.000     .1430391    .2686999
     expersq |  -.0031541   .0010161    -3.10   0.002    -.0051456   -.0011626
         age |  -.0880244    .014573    -6.04   0.000     -.116587   -.0594618
     kidslt6 |  -1.443354   .2035849    -7.09   0.000    -1.842373   -1.044335
     kidsge6 |   .0601122   .0747897     0.80   0.422     -.086473    .2066974
       _cons |   .4254524   .8603696     0.49   0.621    -1.260841    2.111746
------------------------------------------------------------------------------
```

probit inlf nwifeinc educ exper expersq age kidslt6 kidsge6

```
Iteration 0:   log likelihood =  -514.8732
Iteration 1:   log likelihood = -405.78215
Iteration 2:   log likelihood = -401.32924
Iteration 3:   log likelihood = -401.30219
Iteration 4:   log likelihood = -401.30219

Probit estimates                                  Number of obs   =        753
                                                  LR chi2(7)      =     227.14
                                                  Prob > chi2     =     0.0000
Log likelihood = -401.30219                       Pseudo R2       =     0.2206

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.       z     P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0120237   .0048398    -2.48   0.013    -.0215096   -.0025378
        educ |   .1309047   .0252542     5.18   0.000     .0814074     .180402
       exper |   .1233476   .0187164     6.59   0.000     .0866641    .1600311
     expersq |  -.0018871      .0006    -3.15   0.002     -.003063   -.0007111
         age |  -.0528527   .0084772    -6.23   0.000    -.0694678   -.0362376
     kidslt6 |  -.8683285   .1185223    -7.33   0.000    -1.100628    -.636029
     kidsge6 |    .036005   .0434768     0.83   0.408     -.049208    .1212179
       _cons |   .2700768    .508593     0.53   0.595    -.7267472    1.266901
------------------------------------------------------------------------------
```

Changes in probability if kidslt6 changes

mfx compute, at(mean kidslt6=1)

```
Marginal effects after probit
      y  = Pr(inlf) (predict)
         =  .32416867
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
nwifeinc |   -.004323      .00175   -2.48   0.013  -.007744 -.000902   20.1290
    educ |    .047065      .00912    5.16   0.000   .029187  .064943   12.2869
   exper |   .0443479      .00704    6.30   0.000    .03055  .058146   10.6308
 expersq |  -.0006785      .00022   -3.11   0.002  -.001106 -.000251   178.039
     age |  -.0190025      .00284   -6.69   0.000  -.024568 -.013437   42.5378
```

```
 kidslt6 |  -.3121957      .03077  -10.15   0.000  -.372509 -.251882   1.00000
 kidsge6 |   .0129451       .0157    0.82   0.410  -.017829   .04372   1.35325
------------------------------------------------------------------------------
```

mfx compute, at(mean kidslt6=1.5)

```
Marginal effects after probit
      y  = Pr(inlf) (predict)
         =   .1866692
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
nwifeinc |  -.0032274      .00136   -2.37   0.018  -.005892 -.000563   20.1290
    educ |   .0351375      .00789    4.46   0.000   .019683  .050592   12.2869
   exper |    .033109      .00683    4.85   0.000   .019731  .046487   10.6308
 expersq |  -.0005065      .00018   -2.88   0.004  -.000851 -.000162   178.039
     age |  -.0141867      .00232   -6.12   0.000  -.018733  -.00964   42.5378
 kidslt6 |  -.2330773      .01067  -21.84   0.000  -.253993 -.212162   1.50000
 kidsge6 |   .0096645      .01189    0.81   0.416  -.013647  .032976   1.35325
------------------------------------------------------------------------------
```

Example 17.2: Married Women's Annual Labor Supply

use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ, clear

regress hours nwifeinc educ exper expersq age kidslt6 kidsge6

```
      Source |       SS       df       MS              Number of obs =     753
-------------+------------------------------           F(  7,   745) =   38.50
       Model |  151647606     7  21663943.7           Prob > F      =  0.0000
    Residual |  419262118   745  562767.944           R-squared     =  0.2656
-------------+------------------------------           Adj R-squared =  0.2587
       Total |  570909724   752  759188.463           Root MSE      =  750.18

------------------------------------------------------------------------------
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -3.446636      2.544    -1.35   0.176    -8.440898    1.547626
        educ |   28.76112   12.95459     2.22   0.027     3.329284    54.19297
       exper |   65.67251   9.962983     6.59   0.000     46.11365    85.23138
     expersq |  -.7004939   .3245501    -2.16   0.031    -1.337635   -.0633524
         age |  -30.51163   4.363868    -6.99   0.000    -39.07858   -21.94469
     kidslt6 |  -442.0899    58.8466    -7.51   0.000    -557.6148    -326.565
     kidsge6 |  -32.77923   23.17622    -1.41   0.158    -78.2777     12.71924
       _cons |   1330.482   270.7846     4.91   0.000     798.8906    1862.074
------------------------------------------------------------------------------
```

tobit hours nwifeinc educ exper expersq age kidslt6 kidsge6, ll(0)

```
Tobit estimates                                   Number of obs   =        753
                                                  LR chi2(7)      =     271.59
                                                  Prob > chi2     =     0.0000
Log likelihood = -3819.0946                       Pseudo R2       =     0.0343

------------------------------------------------------------------------------
       hours |      Coef.   Std. Err.       t     P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -8.814243   4.459096    -1.98   0.048    -17.56811   -.0603725
        educ |   80.64561   21.58322     3.74   0.000     38.27453    123.0167
       exper |   131.5643   17.27938     7.61   0.000     97.64231    165.4863
     expersq |  -1.864158   .5376615    -3.47   0.001    -2.919667   -.8086479
         age |  -54.40501   7.418496    -7.33   0.000    -68.96862    -39.8414
     kidslt6 |  -894.0217   111.8779    -7.99   0.000    -1113.655   -674.3887
     kidsge6 |    -16.218   38.64136    -0.42   0.675    -92.07675    59.64075
       _cons |   965.3053   446.4358     2.16   0.031     88.88531    1841.725
-------------+----------------------------------------------------------------
         _se |   1122.022   41.57903           (Ancillary parameter)
------------------------------------------------------------------------------

  Obs. summary:        325  left-censored observations at hours<=0
                       428     uncensored observations
```

Changes in probability

**** fixup for expersq : take square of mean rather than mean of square per JMW***

summ exper,meanonly

local exp2=r(mean)^2

mfx compute, at(mean expersq=`exp2') predict(ystar(0,.))

```
Marginal effects after tobit
      y  = E(hours*|hours>0) (predict, ystar(0,.))
         =  687.31745
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
nwifeinc |  -5.687381     2.87788   -1.98   0.048  -11.3279 -.046836   20.1290
    educ |   52.03649       13.82    3.77   0.000   24.9495  79.1234   12.2869
   exper |   84.89173      12.398    6.85   0.000    60.593   109.19   10.6308
 expersq |  -1.202846      .36661   -3.28   0.001  -1.92139 -.484297   113.014
     age |  -35.10478     4.66947   -7.52   0.000  -44.2568 -25.9528   42.5378
 kidslt6 |  -576.8666       70.93   -8.13   0.000  -715.887 -437.847   .237716
 kidsge6 |  -10.46465       24.94   -0.42   0.675  -59.3456  38.4163   1.35325
```

```
------------------------------------------------------------------------------

* marginal effects conditional on positive hours

mfx compute, at(mean expersq=`exp2') predict(e(0,.))

Marginal effects after tobit
      y  = E(hours|hours>0) (predict, e(0,.))
         =  1065.1973
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
nwifeinc |  -3.987413     2.01764   -1.98   0.048  -7.94192 -.032909   20.1290
    educ |   36.48269     9.68927    3.77   0.000   17.4921  55.4733   12.2869
   exper |   59.51744     8.68378    6.85   0.000   42.4975  76.5373   10.6308
 expersq |   -.843313      .25692   -3.28   0.001  -1.34686 -.339765   113.014
     age |   -24.6119     3.27362   -7.52   0.000  -31.0281 -18.1957   42.5378
 kidslt6 |  -404.4402      49.722   -8.13   0.000  -501.893 -306.987   .237716
 kidsge6 |  -7.336744      17.485   -0.42   0.675   -41.607  26.9335   1.35325
------------------------------------------------------------------------------
```

Example 17.3: Poisson Regression for Number of Arrests

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/CRIME1, clear

reg narr86 pcnv avgsen tottime ptime86 qemp86 inc86 black hispan born60

      Source |       SS       df       MS              Number of obs =    2725
-------------+------------------------------           F(  9,  2715) =   23.57
       Model |  145.702778     9  16.1891976           Prob > F      =  0.0000
    Residual |  1864.64438  2715  .686793509           R-squared     =  0.0725
-------------+------------------------------           Adj R-squared =  0.0694
       Total |  2010.34716  2724  .738012906           Root MSE      =  .82873

------------------------------------------------------------------------------
      narr86 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        pcnv |   -.131886   .0404037    -3.26   0.001    -.2111112   -.0526609
      avgsen |  -.0113316   .0122413    -0.93   0.355    -.0353348    .0126717
     tottime |   .0120693   .0094364     1.28   0.201     -.006434    .0305725
     ptime86 |  -.0408735    .008813    -4.64   0.000    -.0581544   -.0235925
      qemp86 |  -.0513099   .0144862    -3.54   0.000     -.079715   -.0229047
       inc86 |  -.0014617    .000343    -4.26   0.000    -.0021343   -.0007891
       black |   .3270097   .0454264     7.20   0.000     .2379359    .4160835
      hispan |   .1938094   .0397156     4.88   0.000     .1159335    .2716853
      born60 |   -.022465   .0332945    -0.67   0.500    -.0877502    .0428202
```

```
       _cons |    .576566   .0378945    15.22   0.000      .502261    .6508711
------------------------------------------------------------------------------
```

poisson narr86 pcnv avgsen tottime ptime86 qemp86 inc86 black hispan born60

```
Iteration 0:   log likelihood = -2249.0104
Iteration 1:   log likelihood = -2248.7614
Iteration 2:   log likelihood = -2248.7611
Iteration 3:   log likelihood = -2248.7611

Poisson regression                                Number of obs   =       2725
                                                  LR chi2(9)      =     386.32
                                                  Prob > chi2     =     0.0000
Log likelihood = -2248.7611                       Pseudo R2       =     0.0791

------------------------------------------------------------------------------
      narr86 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        pcnv |  -.4015713   .0849712    -4.73   0.000    -.5681117   -.2350308
      avgsen |  -.0237723    .019946    -1.19   0.233    -.0628658    .0153212
     tottime |   .0244904   .0147504     1.66   0.097    -.0044199    .0534006
     ptime86 |  -.0985584   .0206946    -4.76   0.000    -.1391192   -.0579977
      qemp86 |  -.0380187   .0290242    -1.31   0.190    -.0949051    .0188677
       inc86 |  -.0080807    .001041    -7.76   0.000     -.010121   -.0060404
       black |   .6608376   .0738342     8.95   0.000     .5161252      .80555
      hispan |   .4998133   .0739267     6.76   0.000     .3549196     .644707
      born60 |  -.0510286   .0640518    -0.80   0.426    -.1765678    .0745106
       _cons |  -.5995888   .0672501    -8.92   0.000    -.7313966    -.467781
------------------------------------------------------------------------------
```

Change in expected arrests if pcnv changes by .10

display "Change in expected arrests if pcnv changes by .10 is " _b[pcnv]*.10

```
Change in expected arrests if pcnv changes by .10 is -.04015713
```

Example 17.4: Duration of Recidivism

use http://fmwww.bc.edu/ec-p/data/wooldridge/RECID, clear

cnreg ldurat workprg priors tserved felon alcohol drugs black married educ age, censored(cens)

```
Censored normal regression                        Number of obs   =       1445
```

```
                                                  LR chi2(10)      =     166.74
                                                  Prob > chi2      =     0.0000
Log likelihood =  -1597.059                       Pseudo R2        =     0.0496

------------------------------------------------------------------------------
      ldurat |      Coef.   Std. Err.       t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     workprg |  -.0625715   .1200369    -0.52   0.602    -.2980382    .1728951
      priors |  -.1372529   .0214587    -6.40   0.000    -.1793466   -.0951592
     tserved |  -.0193305   .0029779    -6.49   0.000    -.0251721    -.013489
       felon |   .4439947   .1450865     3.06   0.002     .1593903    .7285991
     alcohol |  -.6349093   .1442166    -4.40   0.000    -.9178072   -.3520113
       drugs |  -.2981602   .1327356    -2.25   0.025    -.5585367   -.0377836
       black |  -.5427179   .1174428    -4.62   0.000    -.7730958     -.31234
     married |   .3406837   .1398431     2.44   0.015      .066365    .6150024
        educ |   .0229196   .0253974     0.90   0.367    -.0269004    .0727395
         age |   .0039103   .0006062     6.45   0.000     .0027211    .0050994
       _cons |   4.099386   .3475351    11.80   0.000     3.417655    4.781117
-------------+----------------------------------------------------------------
         _se |    1.81047   .0623022          (Ancillary parameter)
------------------------------------------------------------------------------

  Obs. summary:        552     uncensored observations
                       893 right-censored observations
```

Change in durat if a man serves for a felony

mfx compute, nose

```
Marginal effects after cnreg
      y  = Fitted values (predict)
         =  4.8341597
-------------------------------------------------------------------------------
                         variable |          dy/dx                 X
---------------------------------+---------------------------------------------
                          workprg*|      -.0625715             .465052
                           priors |      -.1372529             1.43183
                          tserved |      -.0193305             19.1820
                            felon*|       .4439947             .314187
                          alcohol*|      -.6349093             .209689
                            drugs*|      -.2981602             .241522
                            black*|      -.5427179             .485121
                          married*|       .3406837             .255363
                             educ |       .0229196             9.70242
                              age |       .0039103             345.436
-------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

mat pct=e(Xmfx_dydx)

matmap pct pct, m(100*(exp(@)-1))

mat list pct

```
pct[1,10]
       workprg      priors     tserved       felon     alcohol       drugs
r1  -6.0654125  -12.825026  -1.9144899   55.892217  -47.001643  -25.781754

         black     married        educ         age
r1  -41.883343   40.590851   2.3184231   .39179407
```

Example 17.5: Wage Offer Equation for Married Women

use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ, clear

reg lwage educ exper expersq

```
      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  3,   424) =   26.29
       Model |  35.0223023     3  11.6741008           Prob > F      =  0.0000
    Residual |  188.305149   424  .444115917           R-squared     =  0.1568
-------------+------------------------------           Adj R-squared =  0.1509
       Total |  223.327451   427  .523015108           Root MSE      =  .66642

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1074896   .0141465     7.60   0.000     .0796837    .1352956
       exper |   .0415665   .0131752     3.15   0.002     .0156697    .0674633
     expersq |  -.0008112   .0003932    -2.06   0.040    -.0015841   -.0000382
       _cons |  -.5220407   .1986321    -2.63   0.009    -.9124668   -.1316145
------------------------------------------------------------------------------
```

heckman lwage educ exper expersq, sel(inlf = nwifeinc educ exper expersq age kidslt6 kidsge6) twostep

```
Heckman selection model -- two-step estimates   Number of obs      =       753
(regression model with sample selection)        Censored obs       =       325
                                                Uncensored obs     =       428

                                                Wald chi2(6)       =    180.10
```

```
                                                  Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lwage        |
        educ |   .1090655    .015523     7.03   0.000     .0786411      .13949
       exper |   .0438873   .0162611     2.70   0.007     .0120163    .0757584
     expersq |  -.0008591   .0004389    -1.96   0.050    -.0017194    1.15e-06
       _cons |  -.5781033   .3050062    -1.90   0.058    -1.175904    .0196979
-------------+----------------------------------------------------------------
inlf         |
    nwifeinc |  -.0120237   .0048398    -2.48   0.013    -.0215096   -.0025378
        educ |   .1309047   .0252542     5.18   0.000     .0814074     .180402
       exper |   .1233476   .0187164     6.59   0.000     .0866641    .1600311
     expersq |  -.0018871      .0006    -3.15   0.002     -.003063   -.0007111
         age |  -.0528527   .0084772    -6.23   0.000    -.0694678   -.0362376
     kidslt6 |  -.8683285   .1185223    -7.33   0.000    -1.100628    -.636029
     kidsge6 |    .036005   .0434768     0.83   0.408     -.049208    .1212179
       _cons |   .2700768    .508593     0.53   0.595    -.7267472    1.266901
-------------+----------------------------------------------------------------
mills        |
      lambda |   .0322619   .1336246     0.24   0.809    -.2296376    .2941613
-------------+----------------------------------------------------------------
         rho |    0.04861
       sigma |  .66362876
      lambda |  .03226186   .1336246
------------------------------------------------------------------------------
```

This page prepared by Oleksandr Talavera (revised 9 Nov 2002)

Send your questions/comments/suggestions to Kit Baum at **baum@bc.edu**
These pages are maintained by the Faculty Micro Resource Center's **GSA Program**,
a unit of Boston College **Academic Technology Services**