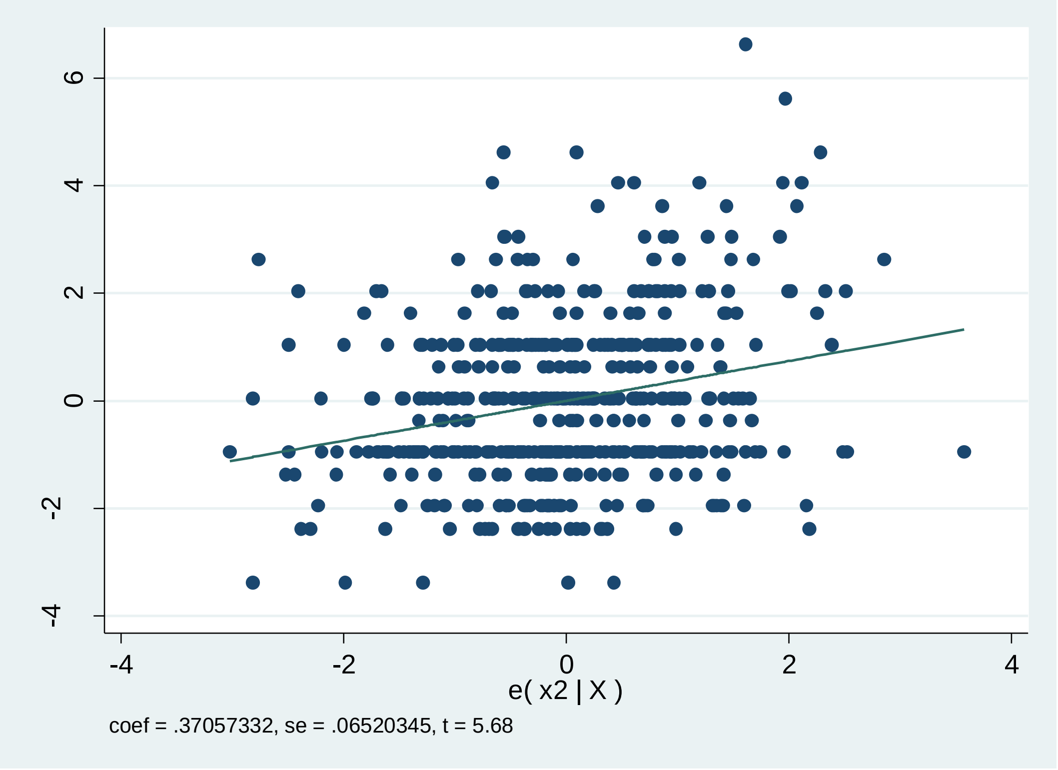
Poisson Regression Model

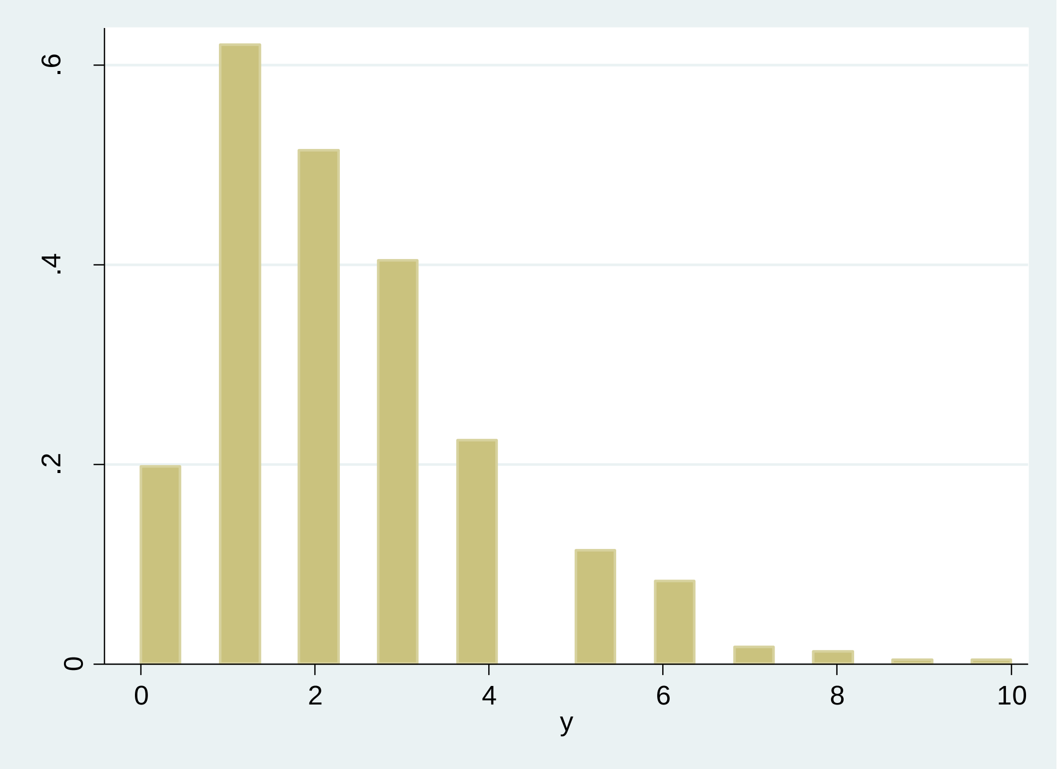
Simulated Data

```
set obs 500
set seed 12345
g x1=(runiform()>0.7)
g x2=rnormal(0,1)
g i=0.7+0.5*x1+0.2*x2
g mu=exp(i)
g y=rpoisson(mu)
```

```
. reg y x1 x2

      Source |       SS           df       MS      Number of obs   =       500
-------------+----------------------------------   F(2, 497)       =     59.63
       Model |  274.451872         2  137.225936   Prob > F        =    0.0000
    Residual |  1143.74813       497  2.30130408   R-squared       =    0.1935
-------------+----------------------------------   Adj R-squared   =    0.1903
       Total |      1418.2       499  2.84208417   Root MSE        =     1.517

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.465322   .1526516     9.60   0.000       1.1654    1.765244
          x2 |   .3705733   .0652034     5.68   0.000     .2424649    .4986817
       _cons |   1.925688   .0796609    24.17   0.000     1.769174    2.082202
------------------------------------------------------------------------------

. avplot x2, recast(scatter)
```



```
. histogram y
```




```
. poisson y x1 x2, nolog

Poisson regression                                Number of obs   =        500
                                                  LR chi2(2)      =     110.76
                                                  Prob > chi2     =     0.0000
Log likelihood = -871.26736                       Pseudo R2       =     0.0598

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .5638447   .0599408     9.41   0.000     .4463629    .6813266
          x2 |   .1567641   .0280064     5.60   0.000     .1018725    .2116557
       _cons |   .6449654   .0380948    16.93   0.000      .570301    .7196299
------------------------------------------------------------------------------

. poisson y x1 x2, ir nolog

Poisson regression                                Number of obs   =        500
                                                  LR chi2(2)      =     110.76
                                                  Prob > chi2     =     0.0000
Log likelihood = -871.26736                       Pseudo R2       =     0.0598

------------------------------------------------------------------------------
           y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.757416    .105341     9.41   0.000     1.562618    1.976498
          x2 |    1.16972   .0327597     5.60   0.000     1.107242    1.235722
       _cons |   1.905921   .0726057    16.93   0.000     1.768799    2.053673
------------------------------------------------------------------------------

. lincom x1, ir

 ( 1)  [y]x1 = 0

------------------------------------------------------------------------------
           y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   1.757416    .105341     9.41   0.000     1.562618    1.976498
------------------------------------------------------------------------------
```

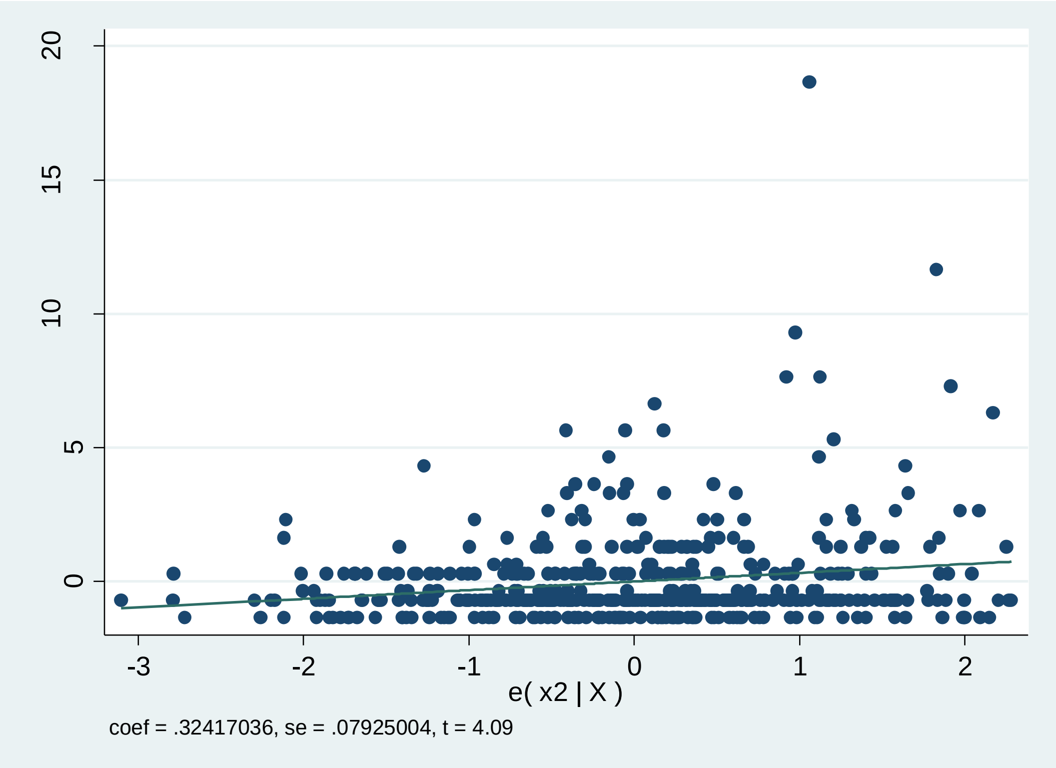
```
. lincom x2, ir

 ( 1)  [y]x2 = 0

------------------------------------------------------------------------------
           y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |    1.16972   .0327597     5.60   0.000     1.107242    1.235722
------------------------------------------------------------------------------

. estat gof

         Deviance goodness-of-fit =  517.6865
         Prob > chi2(497)         =    0.2520

         Pearson goodness-of-fit  =  462.5796
         Prob > chi2(497)         =    0.8636

. mfx

Marginal effects after poisson
      y  = Predicted number of events (predict)
         =  2.2366776
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
      x1*|   1.453223      .17306    8.40   0.000   1.11403  1.79241       .272
      x2 |   .3506308      .06189    5.67   0.000   .229336  .471926    .042486
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. margins , at(x1=(0 1)) atmeans

Adjusted predictions                              Number of obs     =        500
Model VCE    : OIM

Expression   : Predicted number of events, predict()

1._at        : x1              =           0
               x2              =    .0424864 (mean)

2._at        : x1              =           1
               x2              =    .0424864 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   1.918658   .0727248    26.38   0.000      1.77612    2.061196
          2  |    3.37188   .1576509    21.39   0.000      3.06289     3.68087
------------------------------------------------------------------------------
```

Negative Binomial Regression Model

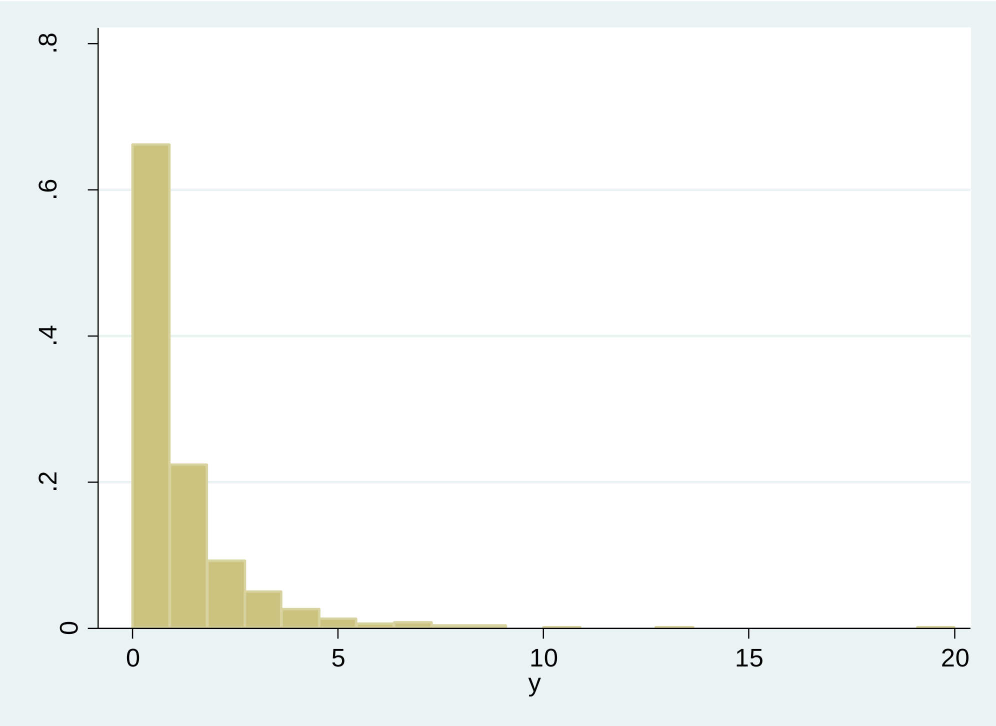
Simulated Data

```
set obs 500
set seed 5555
g x1=(runiform()>0.7)
g x2=rnormal(0,1)
g e=rgamma(0.5,0.8)
g u=ln(e)
g i=0.5+0.7*x1+0.3*x2+u
g mu=exp(i)
g y=rpoisson(mu)
```

```
. reg y x1 x2

      Source |       SS       df       MS              Number of obs =     500
-------------+------------------------------           F(2, 497)     =   15.95
       Model |  100.462651     2  50.2313254           Prob > F      =  0.0000
    Residual |  1565.66535   497  3.15023209           R-squared     =  0.0603
-------------+------------------------------           Adj R-squared =  0.0565
       Total |    1666.128   499  3.33893387           Root MSE      =  1.7749

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .6181471   .1705049     3.63   0.000      .283148    .9531463
          x2 |   .3241704     .07925     4.09   0.000     .1684639    .4798768
       _cons |   .7024186   .0962587     7.30   0.000     .5132945    .8915428
------------------------------------------------------------------------------

. avplot x2, recast(scatter)
```



```
. histogram y
```




```
. poisson y x1 x2, nolog

Poisson regression                                Number of obs   =        500
                                                  LR chi2(2)      =     106.47
                                                  Prob > chi2     =     0.0000
Log likelihood =  -762.5963                       Pseudo R2       =     0.0653

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    .608821   .0941509     6.47   0.000     .4242886    .7933533
          x2 |   .3558912   .0474573     7.50   0.000     .2628766    .4489057
       _cons |  -.4139544   .0666517    -6.21   0.000    -.5445893   -.2833195
------------------------------------------------------------------------------

. estat gof

         Deviance goodness-of-fit =  1026.931
         Prob > chi2(497)         =    0.0000

         Pearson goodness-of-fit  =  1320.657
         Prob > chi2(497)         =    0.0000

. nbreg y x1 x2, nolog

Negative binomial regression                      Number of obs   =        500
                                                  LR chi2(2)      =      34.30
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -626.96136                       Pseudo R2       =     0.0266

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    .578498   .1630038     3.55   0.000     .2590164    .8979795
          x2 |    .339791   .0793079     4.28   0.000     .1843504    .4952317
       _cons |  -.3991611   .0997148    -4.00   0.000    -.5945986   -.2037236
-------------+----------------------------------------------------------------
    /lnalpha |    .605714   .1379337                      .3353689     .876059
-------------+----------------------------------------------------------------
       alpha |    1.83256   .2527718                      1.398456    2.401417
------------------------------------------------------------------------------
LR test of alpha=0: chibar2(01) = 271.27               Prob >= chibar2 = 0.000
```

```
. nbreg y x1 x2, ir nolog

Negative binomial regression                      Number of obs     =        500
                                                  LR chi2(2)        =      34.30
Dispersion     = mean                             Prob > chi2       =     0.0000
Log likelihood = -626.96136                       Pseudo R2         =     0.0266

------------------------------------------------------------------------------
           y |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.783358   .2906941     3.55   0.000     1.295655    2.454639
          x2 |   1.404654   .1114002     4.28   0.000     1.202437    1.640878
       _cons |   .6708826    .066897    -4.00   0.000      .551784    .8156878
-------------+----------------------------------------------------------------
    /lnalpha |    .605714   .1379337                      .3353689     .876059
-------------+----------------------------------------------------------------
       alpha |    1.83256   .2527718                      1.398456    2.401417
------------------------------------------------------------------------------
LR test of alpha=0: chibar2(01) = 271.27             Prob >= chibar2 = 0.000

. mfx

Marginal effects after nbreg
      y  = Predicted number of events (predict)
         =  .81733998
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
     x1*|   .5320673      .17083    3.11   0.002   .197249  .866885        .32
     x2 |   .2777248      .06556    4.24   0.000   .149225  .406224    .036321
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. margins , at(x1=(0 1)) atmeans

Adjusted predictions                              Number of obs     =        500
Model VCE    : OIM

Expression   : Predicted number of events, predict()

1._at        : x1              =           0
               x2              =    .0363213 (mean)

2._at        : x1              =           1
               x2              =    .0363213 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .6792137   .0675899    10.05   0.000       .54674    .8116874
          2  |   1.211281   .1575355     7.69   0.000     .9025171    1.520045
------------------------------------------------------------------------------
```

Example 1

The model $\Pr[Y=y]=\dfrac{e^{-\mu}\mu^{y}}{y!}$ and $\mu_i=\exp(x_i'\beta)$

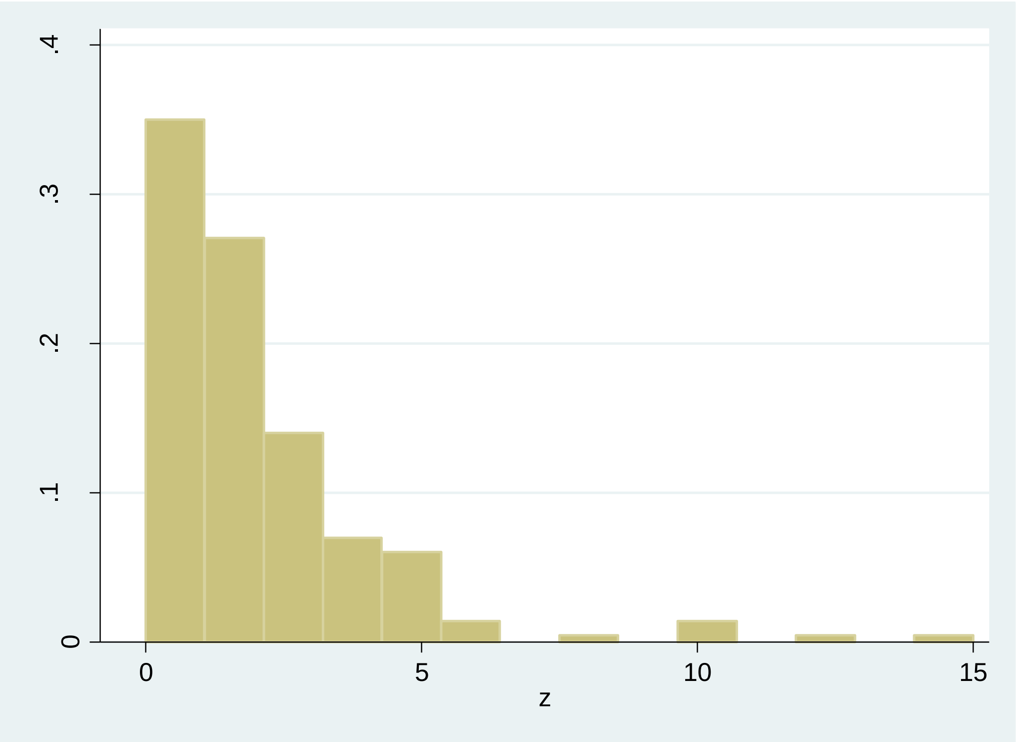
where:

y = Number of times respondent going out to watch movie at the theater each month.

and x_i is matrix of independent variables include:

dinc is dummy variable = 1 if audience's salary is less than 10,000 Baht.
dsex is dummy variable = 1 if audience is female
dedu is dummy variable = 1 if audience's education is lower than bachelor degree
dsta is dummy variable = 1 if audience is single.
form is perception of audience on the movie.

```
. histogram z
(bin=14, start=0, width=1.0714286)
```




```
. reg z dinc dsex dedu dsta form

      Source |       SS       df       MS              Number of obs =     200
-------------+------------------------------           F(  5,   194) =    3.89
       Model |  72.1606046     5  14.4321209           Prob > F      =  0.0022
    Residual |  719.119395   194  3.70680101           R-squared     =  0.0912
-------------+------------------------------           Adj R-squared =  0.0678
       Total |      791.28   199  3.97628141           Root MSE      =  1.9253

------------------------------------------------------------------------------
           z |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dinc |  -.9274258   .2916378    -3.18   0.002    -1.502614   -.3522381
        dsex |  -.4259839   .2881238    -1.48   0.141    -.9942411    .1422733
        dedu |  -.6331845   .4504681    -1.41   0.161    -1.521628    .2552591
        dsta |   .8493393   .3582055     2.37   0.019     .1428622    1.555816
        form |   .2355107   .1379162     1.71   0.089     -.036497    .5075184
```

```
       _cons |   2.511279   .3957344     6.35   0.000     1.730785    3.291774
------------------------------------------------------------------------------

. poisson z dinc dsex dedu dsta form

Iteration 0:   log likelihood =  -365.2093
Iteration 1:   log likelihood =  -365.2093

Poisson regression                                Number of obs   =        200
                                                  LR chi2(5)      =      29.80
                                                  Prob > chi2     =     0.0000
Log likelihood =  -365.2093                       Pseudo R2       =     0.0392

------------------------------------------------------------------------------
           z |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dinc |  -.3793193   .0978647    -3.88   0.000    -.5711307    -.187508
        dsex |  -.1710311   .0936989    -1.83   0.068    -.3546775    .0126152
        dedu |  -.2672999   .1585634    -1.69   0.092    -.5780785    .0434787
        dsta |   .3685056   .1296941     2.84   0.004     .1143099    .6227014
        form |   .0969804    .046308     2.09   0.036     .0062184    .1877424
       _cons |   .8719971   .1389398     6.28   0.000     .5996802    1.144314
------------------------------------------------------------------------------

. estat gof

         Goodness-of-fit chi2  =  210.5588
         Prob > chi2(194)      =    0.1973

. mfx

Marginal effects after poisson
      y  = Predicted number of events (predict)
         =  2.3658003
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   dinc*|  -.8841442      .22276   -3.97   0.000  -1.32074 -.447545       .445
   dsex*|  -.4156465      .23338   -1.78   0.075  -.873064  .041771        .65
   dedu*|  -.5714722      .30478   -1.88   0.061  -1.16884  .025895        .11
   dsta*|   .7821038      .24474    3.20   0.001   .302415  1.26179        .81
    form |   .2294362      .10909    2.10   0.035   .015627  .443245   4.9e-08
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Example 2

The model $\Pr\left[Y=y\right]=\dfrac{e^{-\mu}\mu^{y}}{y!}$ and $\mu_i = \exp(x_i'\beta)$

where:

y = Number of times student going out at night a year.

and x_i is matrix of independent variables include:

program = Program of study = 1 for international program and 0 for Thai program.
sex = Gender of student.
grade = Grade point average of student.
sta = Status of student's family.
dom = Dummy variable = 1 if stay in dormitory and 0 otherwise.
ex = Year of experience in drinking alcohol.
belief = Level of religion belief.
inc = Annual income of student.
pv = Average spending on alcohol beverage each time.
ttc = Average travelling expenditure each time.
ttt = Average travelling time each time (hour).
tst = Average time spending each time (hour).
costt = Annual gross consumption expenditure of student.
cbeer = Annual alcohol spending of student.
zother2 = Annual other leisure spending.
tother2 = Annual other entertainment spending.

Linear regression model

```
. reg y program sex grade sta dom ex belief inc pv ttc t_tt t_st costt c_beer
z_other2 t_other2

      Source |       SS       df       MS              Number of obs =     312
-------------+------------------------------           F( 16,   295) =    6.76
       Model |  26413.1369    16  1650.82106           Prob > F      =  0.0000
    Residual |  71997.2092   295  244.058336           R-squared     =  0.2684
-------------+------------------------------           Adj R-squared =  0.2287
       Total |  98410.3462   311  316.431981           Root MSE      =  15.622

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     program |  -7.527872   2.754987    -2.73   0.007    -12.94979   -2.105953
         sex |  -1.050771   1.989969    -0.53   0.598    -4.967106    2.865564
       grade |   .3519506   2.282883     0.15   0.878     -4.14085    4.844752
         sta |   .4516989   2.919159     0.15   0.877    -5.293317    6.196715
         dom |   .6303185    2.09929     0.30   0.764    -3.501164    4.761801
          ex |   2.349775   .4679602     5.02   0.000     1.428812    3.270739
      belief |  -1.488547   .4739794    -3.14   0.002    -2.421357   -.5557379
         inc |    .000056   .0000336     1.67   0.097    -.0000102    .0001222
          pv |  -.0012581   .0061199    -0.21   0.837    -.0133023    .0107861
         ttc |  -.0022699   .0151408    -0.15   0.881    -.0320675    .0275277
        t_tt |  -4.329307   2.405923    -1.80   0.073    -9.064255    .4056412
        t_st |   3.793012   1.273737     2.98   0.003     1.286249    6.299774
       costt |  -.0000676   .0000515    -1.31   0.190     -.000169    .0000338
      c_beer |   .0004575   .0001372     3.34   0.001     .0001876    .0007275
    z_other2 |   .0000254   .0000233     1.09   0.276    -.0000204    .0000712
    t_other2 |  -.0031929   .0013218    -2.42   0.016    -.0057943   -.0005916
       _cons |   15.36305   10.13627     1.52   0.131    -4.585514    35.31161
------------------------------------------------------------------------------
```

Poisson regression model

```
. poisson y program sex grade sta dom ex belief inc pv ttc t_tt t_st costt
c_beer z_other2 t_other2

Iteration 0:   log likelihood = -2728.4595
Iteration 1:   log likelihood = -2707.9118
Iteration 2:   log likelihood = -2707.8069
Iteration 3:   log likelihood = -2707.8069

Poisson regression                                Number of obs   =        312
                                                  LR chi2(16)     =    1871.54
                                                  Prob > chi2     =     0.0000
Log likelihood = -2707.8069                       Pseudo R2       =     0.2568

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     program |  -.6122153   .0478042   -12.81   0.000    -.7059098   -.5185208
         sex |  -.0946098   .0389511    -2.43   0.015    -.1709527    -.018267
       grade |  -.0632803   .0397681    -1.59   0.112    -.1412243    .0146637
         sta |   .0769079   .0562114     1.37   0.171    -.0332644    .1870801
         dom |   .1129061   .0405919     2.78   0.005     .0333474    .1924648
          ex |   .1841726   .0083343    22.10   0.000     .1678378    .2005075
      belief |  -.1156058   .0083492   -13.85   0.000    -.1319699   -.0992418
         inc |   4.11e-06   5.69e-07     7.23   0.000     3.00e-06    5.23e-06
          pv |  -.0002181   .0001203    -1.81   0.070    -.0004539    .0000177
         ttc |   .0002749   .0002735     1.01   0.315    -.0002612     .000811
        t_tt |   -.499294   .0573647    -8.70   0.000    -.6117268   -.3868612
        t_st |   .1915786   .0202995     9.44   0.000     .1517923    .2313649
       costt |  -5.32e-06   8.44e-07    -6.30   0.000    -6.97e-06   -3.66e-06
      c_beer |   .0000181   1.61e-06    11.22   0.000     .0000149    .0000213
    z_other2 |   2.04e-06   3.63e-07     5.61   0.000     1.33e-06    2.75e-06
    t_other2 |  -.0002584    .000027    -9.56   0.000    -.0003114   -.0002055
       _cons |   3.307107   .1826709    18.10   0.000     2.949079    3.665135
------------------------------------------------------------------------------

. estat gof

         Goodness-of-fit chi2  =  5530.444
         Prob > chi2(303)      =    0.0000

. mfx

Marginal effects after poisson
      y  = Predicted number of events (predict)
         =  9.6471698
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 program*|  -7.342441      .69422  -10.58   0.000  -8.70308  -5.9818   .830128
     sex*|  -.9069866      .37078   -2.45   0.014   -1.6337  -.18027   .429487
   grade |  -.6104762      .38331   -1.59   0.111  -1.36175  .140801   2.84212
     sta*|   .7201405      .51056    1.41   0.158  -.280529  1.72081   .891026
     dom*|   1.089014      .39086    2.79   0.005   .322952  1.85508    .50641
      ex |   1.776745      .07717   23.02   0.000   1.62549  1.92799   2.75962
  belief |  -1.115269      .07952  -14.03   0.000  -1.27112 -.959422   5.17949
     inc |   .0000397      .00001    7.25   0.000   .000029   .00005   64401.9
      pv |  -.0021043      .00116   -1.81   0.070  -.004377  .000169   366.275
     ttc |   .0026521      .00264    1.01   0.315  -.002517  .007821   88.5439
    t_tt |  -4.816774      .54623   -8.82   0.000  -5.88736 -3.74619   .730004
    t_st |   1.848192      .19635    9.41   0.000   1.46336  2.23302     3.553
   costt |  -.0000513      .00001   -6.31   0.000  -.000067 -.000035   47913.9
  c_beer |   .0001747      .00002   11.18   0.000   .000144  .000205   2480.15
z_other2 |   .0000196      .00000    5.63   0.000   .000013  .000026     33390
t_other2 |  -.0024933      .00026   -9.65   0.000     -.003 -.001987   2827.45
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Alternative command for Linear regression model

```
. glm y program sex grade sta dom ex belief inc pv ttc t_tt t_st costt c_beer
z_other2 t_other2
```

```
Iteration 0:   log likelihood =  -1291.564

Generalized linear models                          No. of obs      =       312
Optimization     : ML                              Residual df     =       295
                                                   Scale parameter =  244.0583
Deviance         =  71997.20924                    (1/df) Deviance =  244.0583
Pearson          =  71997.20924                    (1/df) Pearson  =  244.0583

Variance function: V(u) = 1                        [Gaussian]
Link function    : g(u) = u                        [Identity]

                                                   AIC             =  8.388231
Log likelihood   = -1291.564016                    BIC             =  70303.02

------------------------------------------------------------------------------
             |                 OIM
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     program |  -7.527872   2.754987    -2.73   0.006    -12.92755   -2.128197
         sex |  -1.050771   1.989969    -0.53   0.597    -4.951039    2.849497
       grade |   .3519506   2.282883     0.15   0.877    -4.122418    4.826319
         sta |   .4516989   2.919159     0.15   0.877    -5.269747    6.173145
         dom |   .6303185    2.09929     0.30   0.764    -3.484214    4.744851
          ex |   2.349775   .4679602     5.02   0.000      1.43259    3.266961
      belief |  -1.488547   .4739794    -3.14   0.002     -2.41753   -.5595649
         inc |    .000056   .0000336     1.67   0.096    -9.91e-06    .0001219
          pv |  -.0012581   .0061199    -0.21   0.837    -.0132529    .0107367
         ttc |  -.0022699   .0151408    -0.15   0.881    -.0319453    .0274055
        t_tt |  -4.329307   2.405923    -1.80   0.072    -9.044829    .3862155
        t_st |   3.793012   1.273737     2.98   0.003     1.296533     6.28949
       costt |  -.0000676   .0000515    -1.31   0.189    -.0001686    .0000334
      c_beer |   .0004575   .0001372     3.34   0.001     .0001887    .0007264
    z_other2 |   .0000254   .0000233     1.09   0.275    -.0000202     .000071
    t_other2 |  -.0031929   .0013218    -2.42   0.016    -.0057836   -.0006023
       _cons |   15.36305   10.13627     1.52   0.130    -4.503673    35.22977
------------------------------------------------------------------------------
```

<u>Alternative command for Poisson regression model</u>

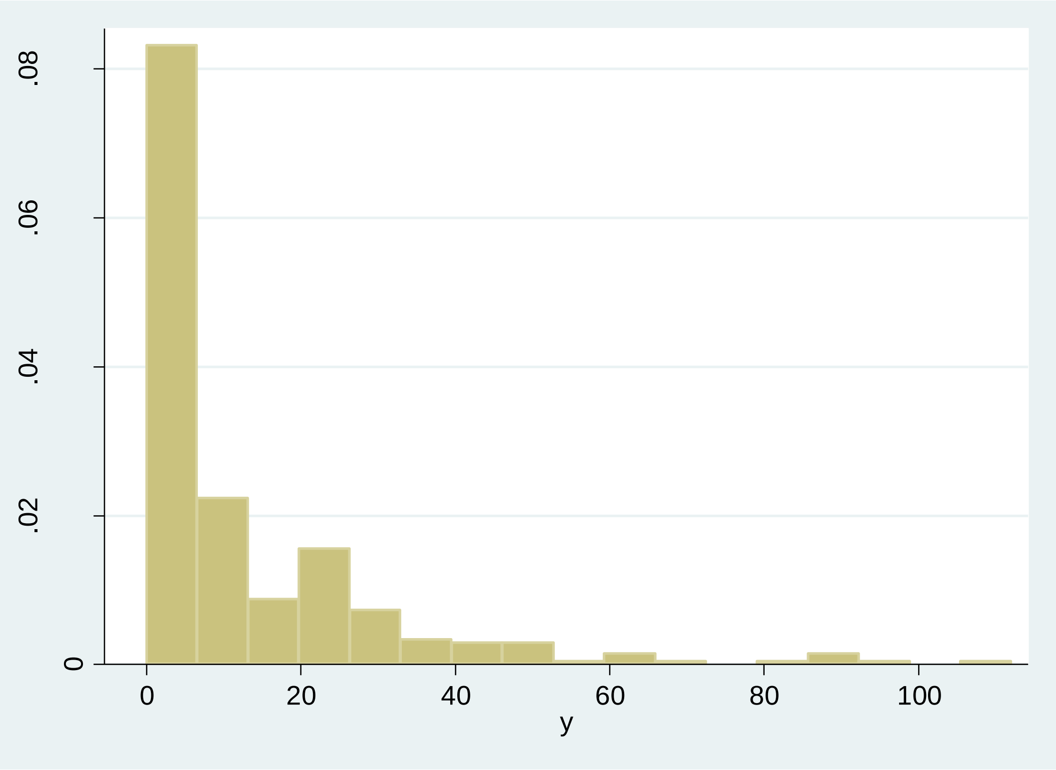
```
. glm y program sex grade sta dom ex belief inc pv ttc t_tt t_st costt c_beer
z_other2 t_other2, family(poisson) link(log)

Iteration 0:   log likelihood = -2860.0678
Iteration 1:   log likelihood = -2708.9239
Iteration 2:   log likelihood = -2707.8074
Iteration 3:   log likelihood = -2707.8069
Iteration 4:   log likelihood = -2707.8069

Generalized linear models                          No. of obs      =       312
Optimization     : ML                              Residual df     =       295
                                                   Scale parameter =         1
Deviance         =  4548.291553                    (1/df) Deviance =  15.41794
Pearson          =  5311.074618                    (1/df) Pearson  =  18.00364

Variance function: V(u) = u                        [Poisson]
Link function    : g(u) = ln(u)                    [Log]

                                                   AIC             =  17.46671
Log likelihood   = -2707.806856                    BIC             =  2854.106

------------------------------------------------------------------------------
             |                 OIM
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     program |  -.6122153   .0478042   -12.81   0.000    -.7059097   -.5185208
         sex |  -.0946098   .0389511    -2.43   0.015    -.1709527    -.018267
       grade |  -.0632803   .0397681    -1.59   0.112    -.1412243    .0146637
         sta |   .0769079   .0562114     1.37   0.171    -.0332644    .1870801
         dom |   .1129061   .0405919     2.78   0.005     .0333474    .1924648
          ex |   .1841726   .0083343    22.10   0.000     .1678378    .2005075
      belief |  -.1156058   .0083492   -13.85   0.000    -.1319699   -.0992418
         inc |   4.11e-06   5.69e-07     7.23   0.000     3.00e-06    5.23e-06
          pv |  -.0002181   .0001203    -1.81   0.070    -.0004539    .0000177
         ttc |   .0002749   .0002735     1.01   0.315    -.0002612     .000811
        t_tt |   -.499294   .0573647    -8.70   0.000    -.6117268   -.3868612
```

```
       t_st |   .1915786   .0202995     9.44   0.000     .1517923    .2313649
      costt |  -5.32e-06   8.44e-07    -6.30   0.000    -6.97e-06   -3.66e-06
     c_beer |   .0000181   1.61e-06    11.22   0.000     .0000149    .0000213
   z_other2 |   2.04e-06   3.63e-07     5.61   0.000     1.33e-06    2.75e-06
   t_other2 |  -.0002584    .000027    -9.56   0.000    -.0003114   -.0002055
      _cons |   3.307107   .1826709    18.10   0.000     2.949079    3.665135
------------------------------------------------------------------------------
```

Negative binomial regression model

```
. nbreg y program sex grade sta dom ex belief inc pv ttc t_tt t_st costt c_beer
z_other2 t_other2

Fitting Poisson model:

Iteration 0:   log likelihood = -2728.4595
...
Iteration 3:   log likelihood = -2707.8069

Fitting constant-only model:

Iteration 0:   log likelihood = -1103.2879
...
Iteration 4:   log likelihood = -1009.9302

Fitting full model:

Iteration 0:   log likelihood = -989.50145
...
Iteration 4:   log likelihood =  -969.5487

Negative binomial regression                      Number of obs   =        312
                                                  LR chi2(16)     =      80.76
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -969.5487                        Pseudo R2       =     0.0400

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     program |  -.3898534   .2760147    -1.41   0.158    -.9308322    .1511255
         sex |  -.1288375   .2122094    -0.61   0.544    -.5447602    .2870852
       grade |    .169173   .2551819     0.66   0.507    -.3309743    .6693203
         sta |   .0300347   .2971865     0.10   0.920    -.5524402    .6125095
         dom |  -.0102337   .2069429    -0.05   0.961    -.4158343    .3953669
          ex |   .4768634   .0735502     6.48   0.000     .3327077     .621019
      belief |  -.1058805   .0517931    -2.04   0.041    -.2073931   -.0043679
         inc |   4.17e-06   4.00e-06     1.04   0.297    -3.67e-06     .000012
          pv |  -.0009809   .0005992    -1.64   0.102    -.0021553    .0001935
         ttc |  -.0004837   .0012703    -0.38   0.703    -.0029734    .0020061
        t_tt |   -.142558   .1920595    -0.74   0.458    -.5189876    .2338716
        t_st |   .1466567   .1117141     1.31   0.189    -.0722989    .3656123
       costt |  -1.62e-06   5.65e-06    -0.29   0.774    -.0000127    9.46e-06
      c_beer |   .0000311   .0000162     1.92   0.055    -6.18e-07    .0000628
    z_other2 |   1.28e-06   2.03e-06     0.63   0.527    -2.69e-06    5.26e-06
    t_other2 |  -.0002077   .0001176    -1.77   0.077    -.0004382    .0000228
       _cons |    1.51107    1.00422     1.50   0.132    -.4571661    3.479306
-------------+----------------------------------------------------------------
    /lnalpha |   .8080556   .0977188                      .6165303    .9995809
-------------+----------------------------------------------------------------
       alpha |   2.243541   .2192362                      1.852489    2.717143
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) = 3476.52 Prob>=chibar2 = 0.000

. mfx

Marginal effects after nbreg
      y  = Predicted number of events (predict)
         =  8.5030321
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 program*|  -3.794166     3.08225   -1.23   0.218  -9.83526  2.24693   .830128
     sex*|  -1.086353     1.77792   -0.61   0.541  -4.57101   2.3983   .429487
   grade |   1.438484      2.1729    0.66   0.508  -2.82032  5.69728   2.84212
     sta*|   .2524134     2.46848    0.10   0.919  -4.58572  5.09055   .891026
```

```
     dom*|  -.0870237      1.7599    -0.05   0.961  -3.53637  3.36232    .50641
       ex |   4.054784      .70836     5.72   0.000   2.66643  5.44314   2.75962
   belief |  -.9003054      .44663    -2.02   0.044  -1.77569  -.02492   5.17949
      inc |   .0000355      .00003     1.04   0.299  -.000031  .000102   64401.9
       pv |  -.0083407      .00514    -1.62   0.105  -.018413  .001731   366.275
      ttc |  -.0041126      .01081    -0.38   0.704  -.025294  .017069   88.5439
     t_tt |  -1.212175     1.63612    -0.74   0.459  -4.41891  1.99456   .730004
     t_st |   1.247027      .95574     1.30   0.192  -.626191  3.12024     3.553
    costt |  -.0000138      .00005    -0.29   0.774  -.000108   .00008   47913.9
   c_beer |   .0002646      .00014     1.90   0.058  -8.8e-06  .000538   2480.15
 z_other2 |   .0000109      .00002     0.63   0.528  -.000023  .000045     33390
 t_other2 |  -.0017661      .00101    -1.75   0.081  -.003747  .000215   2827.45
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. histogram y
(bin=17, start=0, width=6.5882353)
```



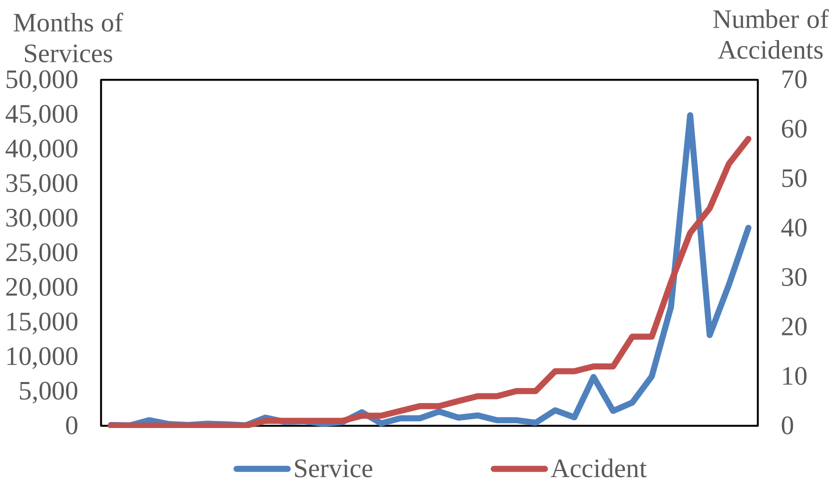
Zero inflated poisson regression model

```
. zip y program grade sta dom ex belief inc t_tt t_st costt c_beer z_other2
t_other2, inflate(dom ex pv) vuong

Fitting constant-only model:

Iteration 0:   log likelihood = -2476.4912
...
Iteration 5:   log likelihood = -2066.6059

Fitting full model:

Iteration 0:   log likelihood = -2066.6059
...
Iteration 4:   log likelihood = -1625.5224

Zero-inflated Poisson regression                  Number of obs   =        312
                                                  Nonzero obs     =        198
                                                  Zero obs        =        114

Inflation model = logit                           LR chi2(13)     =     882.17
```

```
Log likelihood  = -1625.522                       Prob > chi2     =     0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
     program |  -.4891019   .0463289   -10.56   0.000    -.5799048    -.398299
       grade |  -.0732467   .0403027    -1.82   0.069    -.1522385    .0057451
         sta |   .1413059   .0557201     2.54   0.011     .0320964    .2505153
         dom |  -.1053156   .0400339    -2.63   0.009    -.1837806   -.0268506
          ex |   .0450504   .0096875     4.65   0.000     .0260632    .0640376
      belief |  -.1119621   .0084534   -13.24   0.000    -.1285305   -.0953936
         inc |   4.04e-06   6.10e-07     6.62   0.000     2.84e-06    5.23e-06
        t_tt |  -.3349799   .0450537    -7.44   0.000    -.4232834   -.2466763
        t_st |   .2043139   .0180036    11.35   0.000     .1690274    .2396004
       costt |  -5.46e-06   8.43e-07    -6.47   0.000    -7.11e-06   -3.81e-06
      c_beer |   9.32e-06   1.54e-06     6.04   0.000     6.30e-06    .0000123
    z_other2 |   1.48e-06   3.31e-07     4.48   0.000     8.35e-07    2.13e-06
    t_other2 |  -.0002147   .0000271    -7.92   0.000    -.0002678   -.0001615
       _cons |   3.838944   .1848651    20.77   0.000     3.476615    4.201273
-------------+----------------------------------------------------------------
inflate      |
         dom |  -.5846571   .3032645    -1.93   0.054    -1.179045    .0097304
          ex |  -.8348106   .0989515    -8.44   0.000    -1.028752   -.6408693
          pv |   .0013891   .0008184     1.70   0.090     -.000215    .0029931
       _cons |   1.193482   .4236727     2.82   0.005     .3630984    2.023865
------------------------------------------------------------------------------
Vuong test of zip vs. standard Poisson:            z =     8.44  Pr>z = 0.0000

. mfx

Marginal effects after zip
      y  = Predicted number of events (predict)
         =  11.536871
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 program*|  -6.697791      .80233   -8.35   0.000  -8.27034 -5.12524   .830128
   grade |  -.8450383       .4661   -1.81   0.070  -1.75857  .068498   2.84212
     sta*|   1.543878      .58021    2.66   0.008   .406684  2.68107   .891026
     dom*|    .744751     1.11733    0.67   0.505  -1.44518  2.93468    .50641
      ex |   3.308446      .32428   10.20   0.000   2.67288  3.94402   2.75962
  belief |  -1.291692      .11296  -11.43   0.000   -1.5131 -1.07029   5.17949
     inc |   .0000466      .00001    6.38   0.000   .000032  .000061   64401.9
    t_tt |   -3.86462      .54495   -7.09   0.000  -4.93271 -2.79653   .730004
    t_st |   2.357143      .23783    9.91   0.000   1.89101  2.82328     3.553
   costt |   -.000063      .00001   -6.21   0.000  -.000083 -.000043   47913.9
  c_beer |   .0001075      .00002    5.83   0.000   .000071  .000144   2480.15
z_other2 |   .0000171      .00000    4.41   0.000   9.5e-06  .000025     33390
t_other2 |  -.0024765      .00033   -7.47   0.000  -.003126 -.001827   2827.45
      pv |  -.0046403      .00274   -1.70   0.090  -.010002  .000721   366.275
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Poisson Regression Model using Panel Data

Data set from STATA manual of the number of ship accidents for five different types of ships (McCullagh and Nelder 1989, 205). The study analyzes whether the *"incident"* rate is affected by the period in which the ship was constructed and operated. *The measure of exposure is months of service for the ship, and in this model, it is assumed that the exponentiated random effects are distributed as gamma with mean one and variance α.*




```
. xtset ship
       panel variable:  ship (balanced)
```

<u>Random-effects model</u>

```
. xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exposure(service) irr
Fitting Poisson model:
Iteration 0:   log likelihood = -147.37993
...
Iteration 5:   log likelihood = -74.811217
Random-effects Poisson regression               Number of obs      =        34
Group variable: ship                            Number of groups   =         5
Random effects u_i ~ Gamma                      Obs per group:
                                                               min =         6
                                                               avg =       6.8
                                                               max =         7
                                                Wald chi2(4)       =     50.90
Log likelihood  = -74.811217                    Prob > chi2        =    0.0000
------------------------------------------------------------------------------
    accident |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    op_75_79 |   1.466305   .1734005     3.24   0.001     1.162957    1.848777
    co_65_69 |   2.032543    .304083     4.74   0.000     1.515982     2.72512
    co_70_74 |   2.356853   .3999259     5.05   0.000     1.690033    3.286774
    co_75_79 |   1.641913   .3811398     2.14   0.033      1.04174     2.58786
       _cons |   .0013724   .0002992   -30.24   0.000     .0008952     .002104
 ln(service) |          1  (exposure)
-------------+----------------------------------------------------------------
    /lnalpha |  -2.368406   .8474597                     -4.029397   -.7074155
-------------+----------------------------------------------------------------
       alpha |   .0936298   .0793475                      .0177851    .4929165
------------------------------------------------------------------------------
LR test of alpha=0: chibar2(01) = 10.61               Prob >= chibar2 = 0.001
```

Fixed-effects model

```
. xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exposure(service) irr fe
Iteration 0:   log likelihood = -80.738973
...
Iteration 4:   log likelihood = -54.641859
Conditional fixed-effects Poisson regression    Number of obs      =        34
Group variable: ship                            Number of groups   =         5
                                                Obs per group:
                                                              min =         6
                                                              avg =       6.8
                                                              max =         7
                                                Wald chi2(4)       =     48.44
Log likelihood  = -54.641859                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
    accident |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    op_75_79 |   1.468831   .1737218     3.25   0.001     1.164926    1.852019
    co_65_69 |   2.008003   .3004803     4.66   0.000     1.497577    2.692398
    co_70_74 |    2.26693    .384865     4.82   0.000     1.625274    3.161912
    co_75_79 |   1.573695   .3669393     1.94   0.052     .9964273    2.485397
 ln(service) |          1  (exposure)
------------------------------------------------------------------------------
```

Random-effects model with random effects following normal distribution

```
. xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exposure(service) irr
normal

Fitting comparison Poisson model:

Iteration 0:   log likelihood = -147.37993
...
Iteration 5:   log likelihood = -74.780982

Random-effects Poisson regression               Number of obs      =        34
Group variable: ship                            Number of groups   =         5

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =         6
                                                              avg =       6.8
                                                              max =         7

Integration method: mvaghermite                 Integration pts.   =        12

                                                Wald chi2(4)       =     50.95
Log likelihood  = -74.780982                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
    accident |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    op_75_79 |   1.466677   .1734403     3.24   0.001     1.163259    1.849236
    co_65_69 |   2.032604   .3040933     4.74   0.000     1.516025    2.725205
    co_70_74 |   2.357045   .3998397     5.05   0.000     1.690338    3.286717
    co_75_79 |   1.646935   .3820235     2.15   0.031     1.045278    2.594905
       _cons |   .0013075   .0002775   -31.28   0.000     .0008625     .001982
 ln(service) |          1  (exposure)
-------------+----------------------------------------------------------------
    /lnsig2u |  -2.351868   .8586262    -2.74   0.006    -4.034745   -.6689918
-------------+----------------------------------------------------------------
     sigma_u |   .3085306   .1324562                      .1330045    .7156988
------------------------------------------------------------------------------
LR test of sigma_u=0: chibar2(01) = 10.67          Prob >= chibar2 = 0.001
```

Negative Binomial Regression Model using Panel Data

Example data from STATA manual revealed injury *"incidents"* incurred among 20 airlines in each of 4 years. (Incidents range from major injuries to exceedingly minor ones.) The government agency in charge of regulating airlines has run an experimental safety training program, and, in each of the years, some airlines have participated and some have not. The study intended to analyze whether the "incident" rate is affected by the program. Random-effects negative binomial regression was applied, as the dispersion might vary across the airlines for unidentified airline-specific reasons. The measure of exposure is passenger miles for each airline in each year.

Random-effects model

```
. xtnbreg i_cnt inprog, exposure(pmiles) irr
Fitting negative binomial (constant dispersion) model:
Iteration 0:   log likelihood = -293.57997
…
Iteration 5:   log likelihood = -265.38202
Random-effects negative binomial regression     Number of obs      =        80
Group variable: airline                         Number of groups   =        20
Random effects u_i ~ Beta                       Obs per group:
                                                               min =         4
                                                               avg =       4.0
                                                               max =         4
                                                Wald chi2(1)       =      2.04
Log likelihood  = -265.38202                    Prob > chi2        =    0.1532
------------------------------------------------------------------------------
       i_cnt |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      inprog |    .911673   .0590277    -1.43   0.153     .8030206    1.035027
       _cons |   .0367524   .0407032    -2.98   0.003     .0041936    .3220983
  ln(pmiles) |          1  (exposure)
-------------+----------------------------------------------------------------
       /ln_r |   4.794991    .951781                      2.929535    6.660448
       /ln_s |   3.268052   .4709033                      2.345098    4.191005
-------------+----------------------------------------------------------------
           r |   120.9033   115.0735                      18.71892    780.9007
           s |   26.26013   12.36598                       10.4343    66.08918
------------------------------------------------------------------------------
LR test vs. pooled: chibar2(01) = 19.03            Prob >= chibar2 = 0.000
```

Fixed-effects model

```
. xtnbreg i_cnt inprog, exposure(pmiles) irr fe
Iteration 0:   log likelihood = -176.86277
...
Iteration 4:   log likelihood = -174.25143
Conditional FE negative binomial regression     Number of obs      =        80
Group variable: airline                         Number of groups   =        20
                                                Obs per group:
                                                               min =         4
                                                               avg =       4.0
                                                               max =         4
                                                Wald chi2(1)       =      2.11
Log likelihood  = -174.25143                    Prob > chi2        =    0.1463
------------------------------------------------------------------------------
       i_cnt |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      inprog |   .9062669   .0613917    -1.45   0.146      .793587    1.034946
       _cons |   .0329025   .0331262    -3.39   0.001     .0045734    .2367111
  ln(pmiles) |          1  (exposure)
------------------------------------------------------------------------------
```