An Example of Poisson Regression

Consider the AML data from Topic 3 notes. We will look at this data as binned survival data and perform Poisson regression to model the log(rate of relapse) and compare the estimated treatment effect from several Poisson regression models to that obtained by the Cox PH model.

```
. * The analysis consists of the following steps:
.
. * Part a.  Input data, define as survival dataset
. * Part b.  Define survival variables:  stset
. * Part c.  Fit the Cox PH model
. * Part d.  Perform some Poisson regression models on the data
.
.
.
. * Part a.  Input data, define as a survival dataset
.
. *   id,  x(0=no maint 1=maint), t = time to relapse,  failed=(1=relaspsed
0=censored)
.
. input id x t failed

            id          x          t     failed
  1. 1   1  9   0
  2. 2   1  13  0
  3. 3   1  13  1
  4. 4   1  18  0
  5. 5   1  23  0
  6. 6   1  28  1
  7. 7   1  31  0
  8. 8   1  34  0
  9. 9   1  45  1
 10. 10  1  48  0
 11. 11  1  161 1
 12. 12  0  5   0
 13. 13  0  5   0
 14. 14  0  8   0
 15. 15  0  8   0
 16. 16  0  12  0
 17. 17  0  16  1
 18. 18  0  23  0
 19. 19  0  27  0
 20. 20  0  30  1
 21. 21  0  33  0
 22. 22  0  43  0
 23. 23  0  45  0
 24. end
```

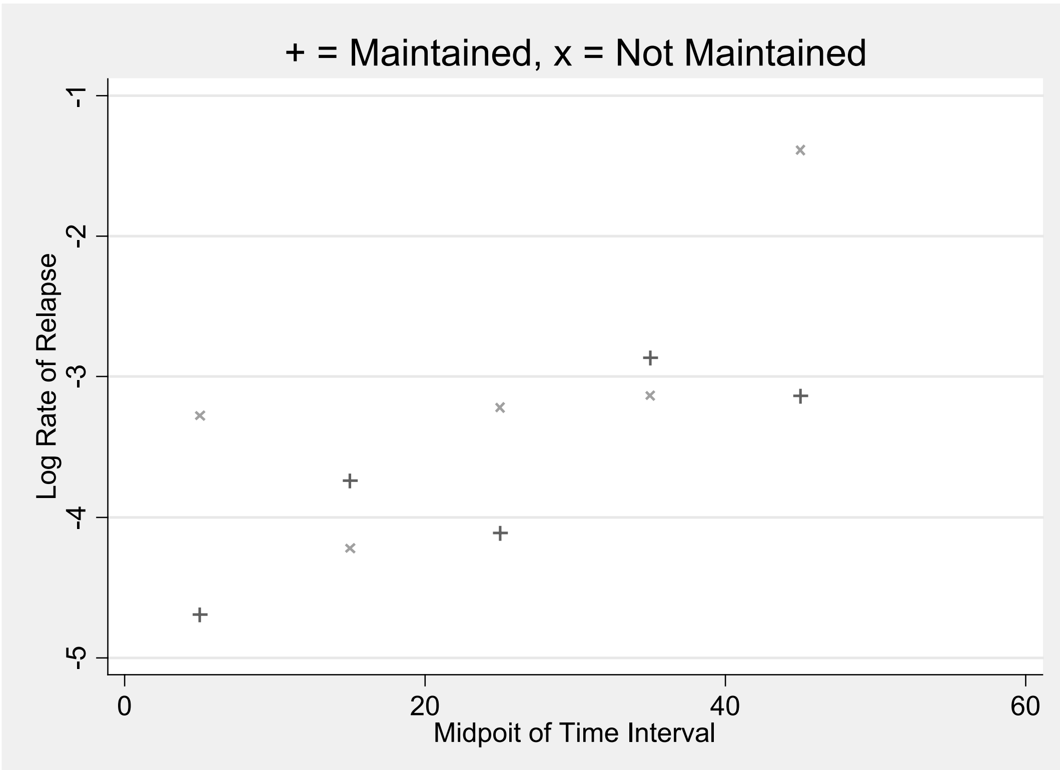
```
.
.
. * Part b.  Define survival variables:  stset
.
. stset t , failure(failed==1) id(id)

                id:  id
     failure event:  failed == 1
obs. time interval:  (t[_n-1], t]
 exit on or before:  failure

------------------------------------------------------------------------------
       23  total obs.
        0  exclusions
------------------------------------------------------------------------------
       23  obs. remaining, representing
       23  subjects
        6  failures in single failure-per-subject data
      678  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =       161
.
. * Save as Stata dataset
.
. save cl10ex1.dta , replace
file cl10ex1.dta saved


. * Part c.  Fit the Cox PH model
.
. use cl12ex1.dta , clear

. stcox x

         failure _d:  failed == 1
   analysis time _t:  t
                 id:  id

Cox regression -- Breslow method for ties

No. of subjects =           23                     Number of obs   =        23
No. of failures =           17
Time at risk    =          678
                                                   LR chi2(1)      =      2.52
Log likelihood  =   -39.438713                     Prob > chi2     =    0.1121

------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .4440875   .2316031    -1.56   0.120     .1597883    1.234219
------------------------------------------------------------------------------
```

From the Cox PH model, we estimate that the relative rate of relapse is 0.44 comparing the maintained and non-maintained patients (95% CI: 0.16 to 1.23). We estimate that the rate of relapse among the maintained patients is 56 percent lower than the rate among patients not maintained.

Now consider fitting several Poisson regression models to the same data. We will bin the data into 10-week periods and define our time variable to be the midpoint of each time interval.

```
.
. * Part d.  Perform some Poisson regression models on this data:
.
. * bin the survival data
.
. stsplit tbin , at( 10(10)50)
(45 observations (episodes) created)

. strate tbin x , output(binrates.dta,replace)

          failure _d:  failed == 1
   analysis time _t:  t
                  id:  id

Estimated rates and lower/upper bounds of 95% confidence intervals
(68 records included in the analysis)

  +--------------------------------------------------------------+
  | tbin   x   D          Y        Rate       Lower       Upper |
  |--------------------------------------------------------------|
  |    0   0   4   106.0000   0.0377358   0.0141629   0.1005437 |
  |    0   1   1   109.0000   0.0091743   0.0012923   0.0651291 |
  |   10   0   1    68.0000   0.0147059   0.0020715   0.1043981 |
  |   10   1   2    84.0000   0.0238095   0.0059547   0.0952009 |
  |   20   0   2    50.0000   0.0400000   0.0100039   0.1599375 |
  |--------------------------------------------------------------|
  |   20   1   1    61.0000   0.0163934   0.0023092   0.1163782 |
  |   30   0   1    23.0000   0.0434783   0.0061245   0.3086553 |
  |   30   1   2    35.0000   0.0571429   0.0142913   0.2284822 |
  |   40   0   2     8.0000   0.2500000   0.0625244   0.9996095 |
  |   40   1   1    23.0000   0.0434783   0.0061245   0.3086553 |
  |--------------------------------------------------------------|
  |   50   1   0   111.0000     0.0e+00           .           . |
  +--------------------------------------------------------------+


.
. * use some Poisson models to estimate the treatment effect
. use binrates.dta, clear

. gen midp = tbin + 5

.
. * Graph the observed log rates!
. gen lograte = log(_D/_Y)
(1 missing value generated)
```

```
. twoway (scatter lograte midp if x==1, ms(+) legend(off)) (scatter lograte midp
if x==0, ms(x) xtit
> le(Midpoit of Time Interval) ytitle(Log Rate of Relapse) title("+ =
Maintained, x = Not Maintained
> "))

. graph export "amldata.wmf", replace
(file C:\DATA\amldata.wmf written in Windows Metafile format)
```




```
.
. * Linear Time
. poisson _D x midp, exposure(_Y) irr

Iteration 0:   log likelihood = -17.984142
Iteration 1:   log likelihood = -17.984065
Iteration 2:   log likelihood = -17.984065

Poisson regression                                Number of obs   =         11
                                                  LR chi2(2)      =       3.45
                                                  Prob > chi2     =     0.1780
Log likelihood = -17.984065                       Pseudo R2       =     0.0876

------------------------------------------------------------------------------
          _D |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .3748694   .2030711    -1.81   0.070     .1296505    1.083891
        midp |   1.008972    .015746     0.57   0.567     .9785778     1.04031
          _Y | (exposure)
------------------------------------------------------------------------------
```

```
.
. * Non-Linear Time, change of slope at 25 weeks
. gen midpsp = 0

. replace midpsp = midp if midp>25
(5 real changes made)

. poisson _D x midp midpsp, exposure(_Y) irr

Iteration 0:   log likelihood = -17.876345
Iteration 1:   log likelihood = -17.876093
Iteration 2:   log likelihood = -17.876093

Poisson regression                                Number of obs   =         11
                                                  LR chi2(3)      =       3.67
                                                  Prob > chi2     =     0.2997
Log likelihood = -17.876093                       Pseudo R2       =     0.0930

------------------------------------------------------------------------------
          _D |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .3640791   .1997663    -1.84   0.066     .1242083    1.067188
        midp |    .993896   .0361748    -0.17   0.866     .9254646    1.067387
      midpsp |   1.013537   .0298692     0.46   0.648     .9566527    1.073803
          _Y | (exposure)
------------------------------------------------------------------------------

.
. * Non-proportional hazards?
. gen inter = x * midp

. poisson _D x midp inter, exposure (_Y) irr

Iteration 0:   log likelihood = -17.011002
Iteration 1:   log likelihood = -17.010532
Iteration 2:   log likelihood = -17.010532

Poisson regression                                Number of obs   =         11
                                                  LR chi2(3)      =       5.40
                                                  Prob > chi2     =     0.1448
Log likelihood = -17.010532                       Pseudo R2       =     0.1369

------------------------------------------------------------------------------
          _D |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   1.009246   .8828206     0.01   0.992     .1817287    5.604938
        midp |   1.038428   .0259255     1.51   0.131      .988838    1.090505
       inter |   .9561694   .0305253    -1.40   0.160     .8981742    1.017909
          _Y | (exposure)
------------------------------------------------------------------------------
```

```
.
. * Step function of time
. xi: poisson _D x i.midp, exposure(_Y) irr
i.midp            _Imidp_5-55         (naturally coded; _Imidp_5 omitted)

Iteration 0:   log likelihood = -13.530063
Iteration 1:   log likelihood = -13.347317
Iteration 2:   log likelihood = -13.343495
Iteration 3:   log likelihood = -13.343089
Iteration 4:   log likelihood =    -13.343
Iteration 5:   log likelihood =  -13.34298
Iteration 6:   log likelihood = -13.342976
Iteration 7:   log likelihood = -13.342976

Poisson regression                                Number of obs   =         11
                                                  LR chi2(6)      =      12.73
                                                  Prob > chi2     =     0.0475
Log likelihood = -13.342976                       Pseudo R2       =     0.3230

------------------------------------------------------------------------------
          _D |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |    .494325    .247489    -1.41   0.159      .185289     1.31879
   _Imidp_15 |   .8758902   .6399406    -0.18   0.856     .2091929     3.66735
   _Imidp_25 |    1.19683   .8743759     0.25   0.806     .2858676    5.010721
   _Imidp_35 |   2.380262   1.741827     1.19   0.236     .5671958    9.988874
   _Imidp_45 |    4.95249   3.669628     2.16   0.031     1.159076    21.16096
   _Imidp_55 |   3.78e-08   .0001483    -0.00   0.997            0           .
          _Y | (exposure)
------------------------------------------------------------------------------
```

Comparison of the estimated relative rate of relapse for the Cox and Poisson models:

Model:	RR (95% CI)
Cox PH	0.44 (0.16, 1.23)
Poisson Model: PH, Linear Time	0.37 (0.13, 1.08)
Poisson Model: PH, Non-Linear Time	0.36 (0.12, 1.07)
Poisson Model: PH, Step-function of Time	0.49 (0.19, 1.32)